



Microsatellite instability assessment is instrumental for Predictive, Preventive and Personalised Medicine: status quo and outlook

Jakub Styk^{1,2,3} · Zuzana Pös^{2,3,4} · Ondrej Pös^{2,3} · Jan Radvanszky^{2,4,5} · Evelina Hrcckova Turnova^{2,6} · Gergely Buglyó⁷ · Daniela Klimova¹ · Jaroslav Budis^{2,3,8} · Vanda Repiska^{1,9} · Bálint Nagy^{2,7} · Tomas Szemes^{2,3,5}

Received: 31 August 2022 / Accepted: 6 January 2023
© The Author(s) 2023

Abstract

A form of genomic alteration called microsatellite instability (MSI) occurs in a class of tandem repeats (TRs) called microsatellites (MSs) or short tandem repeats (STRs) due to the failure of a post-replicative DNA mismatch repair (MMR) system. Traditionally, the strategies for determining MSI events have been low-throughput procedures that typically require assessment of tumours as well as healthy samples. On the other hand, recent large-scale pan-tumour studies have consistently highlighted the potential of massively parallel sequencing (MPS) on the MSI scale. As a result of recent innovations, minimally invasive methods show a high potential to be integrated into the clinical routine and delivery of adapted medical care to all patients. Along with advances in sequencing technologies and their ever-increasing cost-effectiveness, they may bring about a new era of Predictive, Preventive and Personalised Medicine (3PM). In this paper, we offered a comprehensive analysis of high-throughput strategies and computational tools for the calling and assessment of MSI events, including whole-genome, whole-exome and targeted sequencing approaches. We also discussed in detail the detection of MSI status by current MPS blood-based methods and we hypothesised how they may contribute to the shift from conventional medicine to predictive diagnosis, targeted prevention and personalised medical services. Increasing the efficacy of patient stratification based on MSI status is crucial for tailored decision-making. Contextually, this paper highlights drawbacks both at the technical level and those embedded deeper in cellular/molecular processes and future applications in routine clinical testing.

Keywords Microsatellite instability · Cancer · Screening · Massively parallel sequencing · Liquid biopsy · Patient stratification · Predictive Preventive Personalised Medicine (PPPM / 3PM)

Introduction

Microsatellite instability (MSI) is defined as alterations in the length of short tandem repeats (STRs) or microsatellites (MSs) found across the whole genome. Simple, canonical MSs comprise repetitions of a short sequence (usually

defined as 1 to 6 nucleotides) that were associated with human cancer in the 1990s for the first time [1–3], compound MSs contain two or more different repeat units, while complex motifs display varied lengths of repeat units and may include interruptions [4].

✉ Jakub Styk
jakub.styk@gmail.com

¹ Institute of Medical Biology, Genetics and Clinical Genetics, Faculty of Medicine, Comenius University, 811 08 Bratislava, Slovakia

² Comenius University Science Park, 841 04 Bratislava, Slovakia

³ Geneton Ltd, 841 04 Bratislava, Slovakia

⁴ Institute of Clinical and Translational Research, Biomedical Research Centre, Slovak Academy of Sciences, 845 05 Bratislava, Slovakia

⁵ Department of Molecular Biology, Faculty of Natural Sciences, Comenius University, 841 04 Bratislava, Slovakia

⁶ Slovgen Ltd, 841 04 Bratislava, Slovakia

⁷ Department of Human Genetics, Faculty of Medicine, University of Debrecen, 4032 Debrecen, Hungary

⁸ Slovak Centre of Scientific and Technical Information, 811 04 Bratislava, Slovakia

⁹ Medirex Group Academy, NPO, 949 05 Nitra, Slovakia

MSs have long been viewed as an important aspect of human genetic variation that may contribute to the personalised management of many diseases including intellectual disability, cancer, diabetes and cardiovascular diseases [5, 6]. Accurate assessment of the role of MSs in human conditions displaying genetic heterogeneity is difficult due to the poorly defined copy numbers in human reference genomes [7].

In certain types of cancer and associated conditions, MSs display genome-wide instability, i.e. they tend to accumulate indels of several bases due to a failure of the DNA mismatch repair (MMR) machinery. MSI is detected in about 10–15% of colorectal (CRC), gastric and endometrial cancers (EC), while found less frequently in other solid tumours [8, 9]. This replication error phenotype is considered to be a hallmark of a hereditary cancer-prone syndrome, increasing individual risk for various malignancies. It is called Lynch syndrome (LS) [10] and MSI occurs in approximately 90% of cases [11, 12]. Tumours are conventionally classified according to the proportion of unstable STR markers into the following classes: (i) MSI-H (microsatellite instability-high), (ii) MSI-L (MSI-low) and (iii) MSS (microsatellite stable) [13]. Redford et al. [14] suggested replacing the above 3 classes by a binary assessment of MSI events. This would fit-in well with the findings that MSI-L CRCs are often clinically similar to MSS CRCs [15]. Furthermore, patients with MSI tumours show an upregulated expression of various immune checkpoints, currently under study as potential therapeutic targets. They provide better prognosis in the early stages of CRC and a limited benefit from 5-fluorouracil adjuvant chemotherapy in stages II and III [16]. We believe the core of personalised treatment depends on identifying cases where clinicians can effectively utilise the targeted therapeutic procedures and agents to increase the patient's lifetime follow-up. That is the way how the diagnosis and management of individuals with LS or their at-risk relatives should move from reactive testing to proactive Predictive, Preventive and Personalised Medicine (3PM). This shift is fundamental to improve patient outcomes and deliver socio-economic benefit [17, 18].

MSI detection methods evolve continuously, the mechanisms and relationship between MSI and tumours are better understood, and more possibilities for their clinical application are emerging. Earlier methods to determine MSI status, such as immunohistochemistry (IHC) and MSI-PCR [3, 19–22], are still widely used. Apart from certain technical limitations, the full informational potential of MSI status is limited mainly by the need for both tumour and healthy samples for analysis. As such, they are applicable when characterising known and surgically accessible tumours. Non-invasive techniques, however, have started to complement conventional and routinely used invasive methods, exploiting the vast potential of highly sensitive massively parallel

sequencing (MPS) coupled with liquid biopsy sampling. Their main selling point is the future possibility of analysing tumours with no need of a priori awareness of their existence regardless of surgical accessibility or the availability of matching normal tissue for parallel analysis. These will enhance both screening and diagnosis in oncogenetics, with a tremendous impact of 3PM across various levels of health-care [23, 24].

Acknowledging the limitations we have to face in a rapidly growing field, in this review, we would like to provide an up-to-date description of MSI detection strategies and computational tools which allow the processing of high-throughput MPS data for MSI screening and characterisation. According to fact-based evidence, we hypothesise how the routine MSI diagnosis could be advanced to a new age of novel 3PM-based procedures.

The origin of MSI events in cancer

Human genome is a highly dynamic structure, constantly accumulating DNA damage which leads to mutation. One of the essential mechanisms to ensure an acceptable level of whole-genome integrity is a robust and evolutionarily fine-tuned error detection and elimination mechanism called mismatch repair system (MMR). It consists of several proteins encoded by MMR genes. The four most well-known include *MLH1* (mutL homologue 1), *MSH2* (mutS homologue 2), *MSH6* (mutS homologue 6) and *PMS2* (PMS1 homologue 2) [25, 26].

Due to a relatively high complexity of this post-replication system, inter-individual and tumour-specific variability in MMR efficiency is considerable [27]. Loss-of-function mutations in MMR genes may lead to an impaired MMR system, i.e. MMR deficiency (dMMR), which may result in unique somatic alteration events in the length of MS loci, known as MSI [28]. In normal cells, the MMR verifies and maintains the repeat count of MSs. The results of its failure are seen in MSI-positive tissues: an accumulation of mutations that may affect the function of many oncogenes and tumour suppressor genes [29, 30]. This, in turn, is followed by a genome-wide accumulation of frameshift and missense mutations resulting from insufficient correction of base–base mismatches and insertion/deletion (indel) loops that occur in cancer genomes as a result of replication errors or slippage events. MMR deficiency also affects MS sequences, which represent, as pointed out above, highly unstable genomic regions [31].

It has been estimated that replicative polymerases make an error on average once every 10^4 – 10^5 nucleotides. Thus, in every single mammalian cell genome, approximately 100,000 replication errors occur during each cell division. Due to their 3'–5' exonuclease activity, human DNA

polymerases, Pol δ and ϵ , keep, check and reduce this error rate to 1 in 10^5 – 10^7 bases. However, the MMR system may further reduce replication mismatches to about 1 in 10^9 – 10^{10} synthesised nucleotides [32–34]. Surprisingly, MMR is more proficient in removing replication errors on the lagging strand [35, 36].

The MMR complex itself is composed of two types of heterodimers. The first one is MutS, built from MSH2/MSH6 or MSH2/MSH3 pairs, resulting in MutS α and MutS β complexes, respectively. The MutS α complex recognises mismatches and indels of 1–2 bp, while MutL β is able to identify indels of larger sizes [37]. Initiation of the repair mechanism by MutS α and MutL β is followed by the recruitment of MutL, the second heterodimer, composed of MLH1 in a complex with PMS2, PMS1 or MSH3, giving rise to the MutL α , MutL β or MutL γ complexes, respectively, which catalyse the excision of mismatches and error-free resynthesis of DNA [38, 39].

Chromatin organisation affects the frequency of MSI

It should be noted that MSI restricted to single loci or even to single alleles may result from the length of individual MS motifs themselves. This happens when MS repeats exceed certain stability thresholds (MMR deficiency is not required). Such single allele MSIs are typical for repeat expansion disorders, in which the instability of a locus may gradually increase with a growing number of repeats. Typically, one large unstable motif develops (premutation or a fully expanded allele) while other motifs of the same genome, or even the second allele of the same locus, remain stable [40–42]. In contrast, conventional MMR-deficiency-dependent MSI exerts genome-wide effects on MSs, although this effect seems highly uneven. It has been observed that the local density of MSI events is inversely correlated with that of mutations overrepresented in actively transcribed regions. The relative depletion of MSI at stable nucleosome positions also supports the notion that chromatin configuration is a major determinant of the genomic distribution of MSI, at least in CRCs and a subset of uterine corpus endometrial carcinomas [43]. A follow-up study evaluating whole genomes, most of them stated as MSI-H, revealed a high ratio of MSI events in DNA regulatory elements like promoters and enhancers, in actively transcribed areas (euchromatic regions), and within intergenic, intronic and 3' untranslated regions (UTR), contrary to non-transcribed constitutive heterochromatin [44].

Nevertheless, there is considerable inconsistency in the definition of which genome regions are more or less prone to presence of MSI events. The interpretation of affected domains largely depends on the approach used. Further

research is needed to understand whether the increased MSI frequency in open chromatin-like domains occurs during or after DNA replication. There are, however, some possible explanations for the increased MSI occurrence in open chromatin domains: (i) the proofreading activity of DNA polymerase may depend on the state of chromatin; (ii) replication in these domains may proceed more quickly, resulting in a higher error rate; (iii) chromatin-dependent DNA polymerase fidelity may significantly reduce the number of MSI events in nucleosome-occupied regions; and (iv) the base composition of the sequence around an MS site also significantly impacts the susceptibility of the repeat to expansion [45].

According to some evidence, MS' impact on chromatin state is frequently mediated by regulatory changes of histones. Histone tags that open up chromatin and increase transcriptional activity, such as H3K4me3, H3K36me3 and H3K9ac, are associated with an increased prevalence of MSI events, compared to H3K9me2, H3K9me3 and H3K27me2, which act as transcriptional repressors [43]. A genome-wide analysis of STRs affecting gene expression demonstrated that they were strongly associated with histone modifications responsible for transcriptional activation — H3K4me3, H3K27ac, H3K36me3 and H3K9ac. Conversely, they were found to be depleted near the H3K27me3 mark linked to transcriptional repression [46].

Hause et al. [47] published data showing that in certain tumours, MSI accumulates in some functional areas of the genome, such as ion-binding genes. The activin A receptor type 2A (*ACVR2A*) [48, 49] and ring finger protein 43 (*RNF43*) [50] with coding repeat sequence A(8) or C(7) were found to be frequently affected in MSI-positive tumours compared to MSS tumours. To sum it up briefly, studies suggest relevant differences between the MSI-H and MSS groups, indicating the importance of MSI patterns across two levels (within and/or between tumour types). This suggests that they could be used to develop new pan-cancer panels with much greater detection capability. In addition, in 67–100% of evaluated samples classified as MSI-H, the homopolymer repeat located within the splicing region of the *ORC4* gene was also found to be unstable [47].

Role of homopolymeric sequences in MSI

It has been almost two decades since researchers uncovered the vast potential of homopolymeric marker sites as the most interesting type of STRs. Despite the conserved nature of coding homopolymer repeats [51], the non-coding repeats may vary significantly among different individuals. However, if no polymorphisms are known and the number of recurring units at a locus is identical and constant in almost all individuals, even non-coding homopolymers could be

seen as quasi-monomorphic. In CRC patients, three of the five well-known quasi-monomorphic monomeric tracts (BAT-25, BAT-26 and NR-21) are relevant in evaluating MSI events [52].

From all types of STR sequences, the most copious ones in the human genome are homopolymers, predominantly poly(A) and poly(T) tracts [53, 54]. A recent regression analysis proved that motifs composed of A/T have a higher error rate compared to G/C, making them the most mutated MS sequences in humans [55]. These results suggest that the association between A/T stretches and cancer may have developed due to the occurrence of MSI.

Indels in polypyrimidine tracts located in the 5' immediate neighbourhood of a donor splice site may cause exon skipping [56, 57], whereas changes in the 5' and 3' UTR affect transcriptional efficiency and mRNA stability, respectively [58]. Evidence suggests that a poly(A) tract of 30 bp in the promoter of the *HGF* gene is associated with a compactly packed promoter in normal tissue and thus inaccessible to DNA cleaving enzymes. However, deletions in this poly(A) repeat loosen the chromatin structure, modify protein binding and transcriptionally stimulate the promoter in breast cancer tissue [59]. In this context, a comprehensive Selective Targets database, called *Seltarbase*, was created to collect all known mutations in coding, non-coding, UTR and intronic homopolymeric tracts studied in MSI-positive tumours of different organs (<http://www.seltarbase.org/>). It enables identification of relevant genes for tumorigenesis based on their mutation frequency [60]. However, its relevance is declining since the database content has not been updated for almost 10 years.

Clinical significance of MSI detection in cancer management and therapeutic decision-making

Detection of MSI events has broad clinical applications during the lifetime follow-up of CRC, enabling (i) a better management of MSI-positive tumours, and therefore, of LS patients and their relatives at risk; (ii) a better prognosis for patients with MSI-H CRC compared to those with MSS lesions; (iii) omission of fluorouracil, carboplatin or cisplatin adjuvant chemotherapy in cases exhibiting MSI (as they tend to respond poorly) [61, 62]; and (iv) a prediction for the efficacy of immune checkpoint inhibitors (ICIs) in order to identify patients in whom such personalised treatment may be beneficial [63]. The consequence of impaired efficacy of the MMR system is a tumour mutation burden (TMB), which may lead to a high neoantigen load. These are capable of triggering an anti-tumour immune response, which accounts for ICI sensitivity [64–66]. Nonetheless, in recent years, the development

of personalised ICI therapies that target neoantigens has been drawing attention in the field of customised cancer vaccines [67]. This type of specific immunotherapy might deliver clinical benefits. Cost optimisation and time-to-production are as yet challenging; nevertheless, there is hope for laboursaving personalised treatment with a tremendous impact on the socio-economic burden in developing 3PM-based oncology.

In a study analysing MSI status in exomes, Hause et al. observed that the number of unstable MS loci in a tumour exome correlates with patient survival [68, 69]. Such distributions of global instability may be more informative than conventional approaches in clinical management of cancer patients and their treatment. Previous studies have shown that MSI-positive tumours are burdened with neo-antigens easily recognised by the immune system. The effectiveness of PD-1 inhibitors against solid MSI-H tumours exceeds its efficiency compared to MSI-L/MSS cases. In addition to inactivating the genes involved in tumour suppressor pathways as a consequence of dMMR, tumour cells produce neo-peptides known as immunogenic frameshift peptides caused by indels occurring in coding MSs [70, 71].

Abida et al. [72] showed that patients with MSI-H prostate cancer have a good prognosis with the anti-PD-1/PD-L1 treatment, similarly to MSI-H CRC patients. Among the various antibodies, pembrolizumab and nivolumab, two agents that target PD-1/PD-L1 and CTLA-4 immune checkpoint components, have been shown to be efficient across different types of cancer [73, 74]. Currently, a consensus concerning the benefit of such neo-adjuvant therapies seems to have been reached for patients with MSI-H tumours; however, there are still many open questions regarding treatment duration, the timing of surgery and the efficiency of anti-PD-1 therapy alone versus a combination with anti-CTLA-4 therapy [75]. Such clinical studies also prove that MSI testing is adequate in malignancies that are not routinely recommended for MSI screening. Surprisingly, almost 35% of all MSI-H tumours were found to be non-CRC [76] and these patients may benefit from early diagnosis as well as personalised medical services and 3PM.

Apart from the patients' response to immunotherapy, it is also important to note that MSI-H tumours are enriched in altered signalling pathways, rendering the patient less likely to respond. Factors that may impair response include disrupted JAK/STAT signalling [77, 78] and activation of WNT signalling [79–81]. Studies focusing on such pathways may inform about combination therapies and identify not only patients who are likely to respond well to immunotherapy due to dMMR, but also patients who are equally likely to respond for a different reason: a high neoantigen burden unrelated to dMMR. This may help improve the stratification of patients and treatment regimens, thus delivering higher 3PM efficiency.

On a side note, the phenomenon of MSI was reported in diseases other than cancer, such as neurodegenerative disorders [82] (which are now recognised as repeat expansion disorders), atherosclerosis [83, 84] and in spontaneously aborted embryos [85]. In these disorders, however, MSI is not evaluated as part of the diagnostic process as it is in cancer.

High-throughput strategies in MSI screening

Methods suitable for MSI detection range from conventional assays [8], such as capillary gel electrophoresis based on comparing the fragment profile of an amplified set of five quasi-monomorphic homopolymer loci (see Table 1) and IHC [86], through high-resolution melting analysis (HRM) [87, 88] and droplet digital PCR (ddPCR) to high-throughput sequencing platforms [89] (see Table 1 and Table 2). The power of MPS in research and routine clinical settings has been well established and extensively discussed in the literature, at least for SNVs (single nucleotide variants) and CNVs (copy number variants) [90, 91]. Compared to the above, the MPS-based detection and characterisation of TRs took longer to catch on, most likely due to high allelic variability in individual TR loci and certain initial technical limitations. Numerous other challenges also need to be overcome, including issues related to the PCR-induced stutter effect producing artefacts that appear as minor peaks of one or more repeat units, preferential amplification of shorter alleles, overall amplification of larger alleles and the existence of complex motifs. Following the first studies exploiting MPS data for TR characterisation, large-scale genome-wide TR detection began to spread across different biomedical areas, from monogenic repeat-expansion disorders [92, 93] to complex multifactorial traits and diseases [94, 95]. A few years ago, it was confirmed that tumour sequencing, as a first step in the diagnosis of cancer predisposition syndromes such as LS, could serve as an alternative screening assay even for patients with CRC [96].

Currently, the gold standard MSI-PCR assay designed by Promega (Madison, USA), MSI Analysis System, Version 1.2, uses a consensus set of five relatively long homopolymeric MS sites in the human genome (Table 1) [89, 97, 98]. The limitations of conventional tissue-based methods, analysing a limited number of loci, along with the prognostic and therapeutic intervention value of MSI status, urge the necessity of more precise and rapid detection of MSI for better survival outcomes in CRC. Exome and whole-genome analyses under The Cancer Genome Atlas (TCGA) project [99] have revealed the potential of MPS technologies in MSI detection, based on the sequencing of 224 paired tumour vs healthy samples focusing on homopolymers ranging from 6 up to 10 nucleotides. Genome sequencing is the current

state-of-the-art technology to screen for MSI events. At present, there are numerous MPS-based MSI detection methods, some designed for paired healthy and tumour samples, and some more recent ones aiming to reveal MSI events using no control specimens (Table 2). Moreover, these strategies represent an appropriate alternative, especially in non-Lynch-associated tumours, where the driving mechanisms of MSI can be more complex and involve more genes and unusual pathways [100].

Targeted sequencing

MPS has mostly replaced IHC and PCR approaches in MSI status detection, as it enables to analyse dMMR and MSI status simultaneously from tumour specimens [101]. Thus, various MPS-based approaches have begun to emerge, including MSIplus, designed to evaluate MSI status together with mutations in *KRAS*, *NRAS* and *BRAF* genes. This method does not require a matched normal sample and instead uses an amplicon sequencing strategy. Such an approach enables effective targeting and a high depth of coverage for the loci of interest [102]. The ColonCore panel (Burning Rock, Guangzhou, China) was designed to detect the MSI status together with mutations in 36 CRC-related genes (including *KRAS*, *NRAS*, *BRAF*, hereditary CRC genes and other genes related to carcinogenesis and tumour development). In combination with this sequencing panel, Zhu et al. developed an algorithm suitable for MSI detection, revealing the MSI status of a sample by determining the percentage of unstable loci, without a paired specimen [103]. They validated this approach with a gold standard MSI-PCR and IHC, reaching a high concordance rate. Compared to MSIsensor [104] and mSINGS [105], it achieved a better result and seems to represent a robust method based on read-count distribution [103].

So far, many methods have been developed for the detection of MSI from targeted MPS data and have demonstrated their reliability, sensitivity and efficiency across various cancers, even those not typically screened, without the need for matched normal tissue [76, 81, 106–110]. Apart from developing one of these methods, Trabucco et al. [81] analysed additional alterations that should be common for MSI-H and MSS tumours. They identified genes enriched in MSI-H tumours that fell into common pathways, such as the phosphatidylinositol 3-kinase (PI3K), NOTCH and WNT pathways. Janus kinase/signal transducers and activators of transcription (JAK-STAT) and Hedgehog pathways were also found to be enriched, but only when regions of polymerase slippage were included. On the other hand, MSS tumours were consistently enriched in alterations of genes involved in the cell cycle across all tumour types [81]. Currently, TruSight™ Oncology 500 (TSO500), the innovative

Table 1 List of assays for the assessment of MSI

Assay	Methods/Lab. equip	Analysis	Markers (MNRs/DNRs)	Classification of MSI status	Sample/input	Paired samples	Ref.
MSI Analysis System, I.2 (Promega)	FM-PCR, Capillary Electrophoresis/compatible Genetic Analyzer	Fragment Analysis Software (e.g. GeneMapper®)	BAT-25, BAT-26, NR-21, NR-24, MONO-27/PentaC, PentaD	Ternary: MSI-H ($2 \geq 5$); MSI-L (1/5); MSS (0/5)	FFPE/1–2 ng DNA	R	[169]
OncoMate™ MSI Dx Analysis System (Promega)	FM-PCR, Capillary Electrophoresis/compatible Genetic Analyzers	Fragment Analysis Software	BAT-25, BAT-26, NR-21, NR-24, MONO-27/PentaC, PentaD	Binary: MSI-H ($2 \geq 5$); MSS ($< 2/5$)	FFPE/1 ng DNA	R	[170]
PlentiPlex™ MSI (PentaBase)	RT-PCR, Capillary Electrophoresis/compatible instruments (e.g. Base-Caller™ 48.4 RT-PCR system)	Fragment Analysis Software	BAT-25, BAT-26, NR-21, NR-24, MONO-27/D17S250, D2S123, D5S346	Ternary: MSI-H ($2 \geq 5$); MSI-L (1/5); MSS (0/5)	FFPE; FF/1–40 ng	R	[170, 171]
TrueMark™ MSI Assay (ThermoFisher)	RT-PCR, Capillary Electrophoresis/SeqStudio or ABI 3500 Series Genetic Analyzers	TrueMark™ MSI Analysis Software	BAT-25, NR-24, NR-21, BAT-40, CAT-25, NR-22, NR-27, ABI-19, ABI-20B, ABI-17, ABI-16, BAT-26, ABI-20A/TH01, PentaD	Binary: MSI-H vs MSS	FFPE/2 ng DNA	R	[170, 172]
ProDx® MSI Analysis System	FM-PCR, Capillary Electrophoresis/ABI 3500Dx Genetic Analyzer	Fragment Analysis Software	BAT-52, BAT-56, BAT-59, BAT-60, BAT-25, BAT-26, NR-21, MONO-27	Ternary: MSI-H ($2 \geq 5$); MSI-L (1/5); MSS (0/5)	FFPE/5–10 ng DNA	R	[123]
Bio-Rad® ddPCR MSI Assay	Droplet Digital PCR/Bio-Rad QX200 Instrument System	QuantaSoft software v.1.7.4 (Bio-Rad)	BAT-25, BAT-26, NR-21, NR-24, MONO-27	Binary: MSI-H ($2 \geq 5$); MSS (0 or 1/5)	FFPE; FF; plasma or serum samples/NA	NR	[89]
Idylla™ MSI Test (Biocartis)	RT-PCR, High-Resolution Melting Curve Analysis/Biocartis Idylla™ System	Idylla™ Explore	ACVR2A, BTBD7, D1D01, MRE11, RYR3, SEC31A, SULF2	Binary: MSI-H ($2 \geq 7$); MSS ($< 2/7$)	5 μ m (50–600 mm ²) or 10 μ m (25–300 mm ²) FFPE tissue section/no DNA extraction	NR	[173]
QiaXcel Advanced System/pilot phase	PCR/QiaXcel Advanced System	QiaXcel System	BAT-25, BAT-26, BAT-40/D2s123, D10s197, D13s153, D17s250, D18s58, D5s346, Mycl	Binary: MSI-H vs MSS	FFPE/5 ng DNA	R	[174]
HSP110 (T17) PCR	E-ice-COLD-PCR/LightCycler, ABI 3500Dx Genetic Analyzer	NA	T(17) in HSP110 gene	Suitable for improving analytical sensitivity of MSI testing; not for MSI status classification	FFPE/20 ng DNA	NR	[175–177]
Drop-off ddPCR	ddPCR/Bio-Rad System	NA	BAT26, ACVR2A/DEFB105A/B	Bimodal distribution of MSI allelic frequencies	FFPE; plasma; serum/NA	NR	[144]

Table 1 (continued)

Assay	Methods/Lab. equip	Analysis	Markers (MNRs/DNRs)	Classification of MSI status	Sample/input	Paired samples	Ref.
Titano MSI Test (Diatech Pharmacogenetics)	FM-PCR, Capillary Electrophoresis/compatible Genetic Analyzers	Fragment Analysis Software	BAT25, BAT26, D2S123, D17S250, D5S346, BAT40, D18S58, NR21, NR24, TGFβRIII/TPOX, TH01	Binary: MSI-H vs MSS	FFPE; FF; WB (only CRC)/20 ng DNA	R	[178]
Agilent TapeStation 4200 platform	Capillary Electrophoresis/Bioanalyzer	TapeStation Analysis Software	BAT25, BAT26/D5S346, D2S123, D17S250	Binary: MSI-H vs MSS	FFPE/20 ng DNA	R	[178]
VarTracer® MSI qPCR Assay (NuProbe)	Multi-blocker displacement amplification (mBDA) & qPCR/BioRad CFX96	VarTrace MSI Software	BAT-25, BAT-26, NR-21, NR-24, MONO-27	Ternary: According to Ct value for each out of the marker: MSI-H (> 2/5); MSI-L(1/5); MSS (0/5)	FFPE; WB/1.5–15 ng DNA per reaction (3–30 ng per specimen); ideally 20 ng (10 ng DNA per tube)	NR	[179]

Explanatory notes: *Lab. equip.*, laboratory equipment; *FM-PCR*, fluorescent multiplex PCR; *ddPCR*, droplet digital PCR; *Ref.*, references; *NA*, not available; *R*, required; *NR*, not required; *MNRs*, mononucleotide repeats; *DNRs*, dinucleotide repeats; *MSI-H*, microsatellite instability - high; *MSI-L*, microsatellite instability - low; *MSS*, microsatellite stable; *FFPE*, formalin-fixed and paraffin-embedded; *FF*, fresh frozen; *WB*, whole blood

and accurate tumour-only pan-cancer workflow proposed by Illumina, is widely used for MSI screening. It may be regarded as a straightforward and cost-effective assay, allowing the detection of various DNA variant types [111, 112].

Whole-exome sequencing

Many known MSI characteristics were identified using conventional strategies. However, these observations came from analysing only a few dozen loci in cohorts lacking sufficient size. A robust strategy using tumour exomes from TCGA Research Network to predict and examine MSI status across 18 cancer types was introduced [68]. It worked with 2.7% of all MSs identified through the human genome (more than 500,000 loci), representing 95.9% of all coding MSs and 98.4% of MSs located in splice sites. Based on this approach, MOSAIC, a classifier for predicting MSI status, was introduced (see section “[Bioinformatics: essential element for MPS-based MSI assessment](#)”). It reclassified MSI-L tumours as MSS, supporting previous assumptions that MSI-L is not a distinct class [43, 113]. Another tool, MSIpred, was developed and trained on MSS (encompassing MSI-L) and MSI-H tumours. This tool also seems to be reliable for MSI classification and capable of distinguishing between different tumour types as MSS and MSI-H tumours possess distinct somatic mutational loads [114].

Since whole-exome sequencing (WES) is becoming increasingly used in clinical care and 3PM, there is a growing need to use this data for MSI status detection. Some of the emerging methods are turning out as valid, adequate and sensitive [115]. New computational tools have been developed to extract MSI status from whole-exome data, but each method, process, tool and even the whole pipeline should be validated before it is integrated into routine usage [116]. Accordingly, MANTIS, one of the tools developed for the above purpose [117], was compared with mSINGS [105] and MSIsensor [104], and their performance and accuracy were validated. MANTIS displayed high accuracy and robustness across many cancer types and loci. Despite some discrepancies [68, 102, 114], the above tools may reliably complement the conventional MSI-PCR method, and may soon become stand-alone gold standards for MSI status identification.

Whole-genome sequencing: a tool for MSI-related cancer detection

Genome-wide genotyping of STRs offers extensive potential to transform applications in cancer research and 3PM beyond the study of conventional MSI. In whole-genome sequencing (WGS), STRs have presented a broad range of challenges as diagnostic markers. The screening of MSI

Table 2 List of microsatellite panels using tissue-based MPS strategies

Study	Panel/No. of tested loci	Seq. Platform	Paired samples	Tool	Tumour type	Analytical performance/reference methods
[180]	3 In-house panels/3 154, 230, 23	NextSeq 500	R	mSINGS, mSILICO	CRC	NA
[112]	TruSight™ Oncology 500/130 nc-MNRs	NextSeq™ 550Dx	NR	In-house	6 cancer types	98.0% (106) specificity/pentaplex MSI-PCR
[108]	OncoPanel AMC version 3/85 MNRs	MiSeq	NR	NA	CRC	92.1% sensitivity; 100.0% specificity for MSI-H/NA
[181]	5 MNRs; 2 DNRs	MiSeq	R	In-house	CRC	100.0% sensitivity; 100.0% specificity/pentaplex MSI-PCR or 9 loci NCI/Bethesda panel
[109]	40 unspecified loci	MiSeq	NR	In-house	7 cancer types	98.0% (98/100) concordance/pentaplex MSI-PCR
[14]	17 MNRs	MiSeq	NR	In-house	CRC	100.0% sensitivity; 100.0% specificity/pentaplex MSI-PCR
[81]	1 880 MNRs	In-house FoundationOne assay	NR	In-house	Large cohort	Combined concordance (con.) 97.0% (65/67)/pentaplex MSI-PCR and IHC; 95.0% con./MSI-PCR; 100.0% con./IHC; 97.0% con. with 95.0% sensitivity and 98.0% specificity to comparable methods [101]
[182]	BROCA/146 MNRs	HiSeq 2500	R	mSINGS	Prostate cancer	NA/pentaplex MSI-PCR
[107]	smMIPs panel/111 MNRs	NextSeq 500	NR	mSINGS	CRC, prostate, endometrial cancer	95.8–100.0% sensitivity; 100.0% specificity/pentaplex MSI-PCR
[106]	Caris MI TumourSeek/7 317	NextSeq	NR	In-house	26 cancer types	100.0% sensitivity; 99.9% specificity (CRC), 95.8% sensitivity; 99.4% specificity (all cancer types)/pentaplex MSI-PCR
[105]	UW-OncoPlex/15 MNRs	NA	NA	mSINGS	NA	100.0% sensitivity; 100.0% specificity/pentaplex MSI-PCR
[183]	Caris MI TumourSeek	NA	NA	In-house	12 cancer types	95.8% sensitivity; 99.4% specificity/pentaplex MSI-PCR
[103]	ColonCore/22 MNRs	NA	R	MSI-ColonCore	CRC	97.9% sensitivity (47/48); 100.0% specificity (37/37)/pentaplex MSI-PCR and IHC
[117]	NA	NA	R	MANTIS, MSIsensor, mSINGS	6 cancer types	97.2% sensitivity; 99.7% specificity for MANTIS, 96.5%; 98.7% for MSIsensor, 76.06%; 99.7% for mSINGS/NA
[76]	MSK-IMPACT/1000–1500 unspecified loci	Targeted NGS dataset	R	MSIsensor	66 cancer types	96.1% sensitivity; 98.5% specificity (MSS vs MSI-H); 100.0% sensitivity; 99.3% specificity (MSI-H in CRCs and uterine endometrial cancers), 96.6% sensitivity; 100.0% specificity (MSI-H in other tumour types)/pentaplex MSI-PCR

Table 2 (continued)

Study	Panel/No. of tested loci	Seq. Platform	Paired samples	Tool	Tumour type	Analytical performance/reference methods
[101]	NA	HiSeq 2500	NR	In-house	CRC	100.0% sensitivity; 99.0% specificity in detection of MSI-H/pentaplex MSI-PCR
[102]	17 MNRS	MiSeq	NR	mSINGs	78 cancer types	97.1% sensitivity (34/35) in MSI positive samples; 100.0% specificity (42/42) in MSI negative samples/pentaplex MSI-PCR

Explanatory notes: *Seq. Platform*, sequencing platform; *Ref.*, references; *NA*, not available; *R*, required; *NR*, not required; *nc*, non-coding; *MNRs*, mononucleotide repeats; *DNRs*, dinucleotide repeats

using a large-scale WGS approach may yield a continuously valued MSI score that may have greater utility in determining genomic MSI events [117]. One of the common issues for robust MSI detection on the whole-genome scale is the usage of sequencing data obtained from low-coverage analyses [76]. However, the performance of the software for the assessment of MSI from WGS data is severely affected by the depth of sequencing. Thus, determining the minimum sequencing depth of coverage for making MSI calls and characterising the loss of detection power in low-coverage regions seems helpful [118]. A genome-wide survey of MSI presented by Cortes-Ciriano et al. revealed that MSI events may affect as many as 300,000 loci. Non-coding MSI events occurring in regulatory elements may function as cancer drivers [44]. Although there are numerous computational tools for screening of somatic MSI, only few use low-coverage WGS data. MSIClass, developed by Maruvka et al. [119] using ultra-low coverage data (as low as 0.05×), could accurately detect somatic MSI events.

As the era of MPS began, this method was also recognised as useful in the field of cancer assessment, especially for detecting MSI status in tumours characterised by low MSI frequency, since MSI testing is not performed systematically. It has been shown that each cancer type may have its own MSI signature, so an important benefit of MPS is the ability to include numerous MS loci. The number and selection of the analysed loci may affect the accuracy of the MSI calling method. It may therefore be helpful to find an optimal number of loci for analysis with each tool to ensure its optimal performance [117]. There are many MSs found across the human genome [44, 68], and their behaviour may vary across different malignancies. The cancer-specific MSI landscape promises potential predictive power with huge implications for clinical diagnosis but it may only be uncovered by WES or WGS. Most MSs are located in non-coding regions of the genome and may affect gene regulation, consequently leading to carcinogenesis.

Technical and biological limitations of MPS-based MSI detection approaches

MSI detection through MPS is a very complex strategy involving many obstacles and challenges. Technical issues in MSI calling involve uneven genome coverage and artefacts in sequencing technology. It has been known for a long time that the frequency of errors (in both PCR or sequencing) tends to increase with homopolymer length [120]. Among all available sequencing platforms, Illumina has been shown to be the most suitable for these stretches due to its low error rate within monomeric tracts. The average error rate has been shown to range from 0.002% for homopolymers of 2 bp in length to 2% for 17-bp regions [121]. All in all,

using homopolymeric tracts as diagnostic markers in MSI screening shows a significantly higher sensitivity in MSI-positive cancers in contrast to polymorphic dinucleotide repeat markers. In general, mononucleotide markers are preferred over higher repeat stretches. It seems a bit confusing, as multiple studies have shown that dinucleotide repeats are better for detecting MSI-L cancers, most of which are MMR-proficient, while mononucleotide MSI markers are far more effective in detecting dMMR tumours [122].

A major challenge in most strategies based on the analysis of homopolymeric markers is the detection of small indels [118]. Some evidence has suggested the number of repetitions under seven as a poor indicator; thus, longer mononucleotide repeats (> 15 bp) have been preferred [58]. Such repeats, however, are more prone to errors due to PCR or sequencing. Therefore, short repeats seem to offer an alternative, as they are also more monomorphic than longer variants, not requiring analysis of matched normal tissue [14]. There are hypotheses that longer homopolymeric repetitions could easily increase the sensitivity of the monomeric tract, thus improving the performance of MSI screening. A study by Wu et al. [123] performed on CRC patients demonstrated that the ProDx® MSI system containing four additional markers (BAT-52, BAT-56, BAT-59 and BAT-60) (Table 1) brought a minor improvement in detection sensitivity, similar to the prior finding of enhanced detection sensitivity by long mononucleotide repeat markers in colon polyps. Despite the above, it may be argued that long homopolymers might be ultimately better than shorter ones as more extensive changes are much more reliable when calculating MSI scores [123].

Detecting MSI events using not only homopolymer markers may also be difficult due to technical artefacts associated with FFPE specimens, routinely used in translational research settings [124]. The preparation of FFPE blocks may result in remarkable changes in DNA structure, such as deamination of C to T or transition of G to A, which may be partially resolved by enzymatic treatment with uracil DNA glycosylase and/or 5-methylcytosine DNA glycosylase, thus reducing background noise and false positive results [125].

Several factors pose a computational challenge for MSI detection by MPS, such as the errors induced during the sequencing-by-synthesis, the difficulties of alignment of repetitive DNA motifs due to the short read length and the low accuracy of indel calling [126]. Further improvements are expected to help reduce the high risk of interpretation errors and the chance of incorrect indication of patients for immunotherapy. Reliable algorithms successfully addressing this problem will likely take longer to develop.

Targeted sequencing allows efficient testing for MSI while screening for additional relevant genomic information. Moreover, MSI status may be reliably determined regardless of the specific baits on the MPS panel, as long as there is

sufficient genomic coverage. Validation studies have shown the method to be sensitive and specific, with high concordance to traditional methods, enabling confident assessment of MSI status without the requirement for matched normal tissue [105]. However, dependence on read coverage represents the main limitation of MPS-based MSI detection compared to MSI-PCR and IHC [117]. We should also keep in mind that, similarly to PCR, MPS reveals the effects of dMMR without revealing the underlying cause.

Bioinformatics as an essential element for MPS-based MSI assessment

In terms of 3PM objectives, bioinformatics offers improved diagnosis at the genomic level with earlier detection and more targeted therapy for effective personal disease monitoring [127]. Computational tools needed for MSI analysis may be easily integrated into the existing pipelines commonly used to detect other mutations (e.g. SNPs, CNVs) from sequenced genomic data. A retrospective analysis of previously generated MPS data is also possible. Various tools for detecting MSI status from MPS data are available (Fig. 1): some assess MS loci directly in DNA, while others indirectly, e.g. through the analysis of somatic mutations [128, 129]. There are other tools designed to determine MSI status from RNA-seq data (Table 4) [130]. Since these tools do not use the same sequencing data for input, it is not possible to compare their performance.

In the context of MPS-based algorithms usually focused on comparing the repeat lengths of several homopolymers, the majority of MSI detection tools are intended for a limited number of cancer types. There are two main types of MSI analysis. The mutation burden approach is limited by the need for large and costly sequencing panels, ranging from targeted to WGS panels, as the tumour mutation burden (TMB) calculated from small panels may deviate from reality. The second limitation is that cut-off values between MSS, MSI-L and MSI-H have to be defined for every sequencing panel. The second type of analysis is the read-count distribution approach [103], which is more compatible with smaller and cheaper sequencing panels compared to the mutation burden approach. With regard to cut-off values, this approach could be more versatile because cut-offs can be defined for any given panel using the percentage of unstable loci.

Several issues must be considered when selecting an appropriate tool for MSI analysis (Table 4). First, the type of biological material for nucleic acid extraction. Since tissue samples were commonly used for MSI detection, most tools were designed with them in mind. However, cell-free nucleic acids from liquid biopsy are steadily gaining ground and tissue-based approaches may not be compatible, as sensitivity

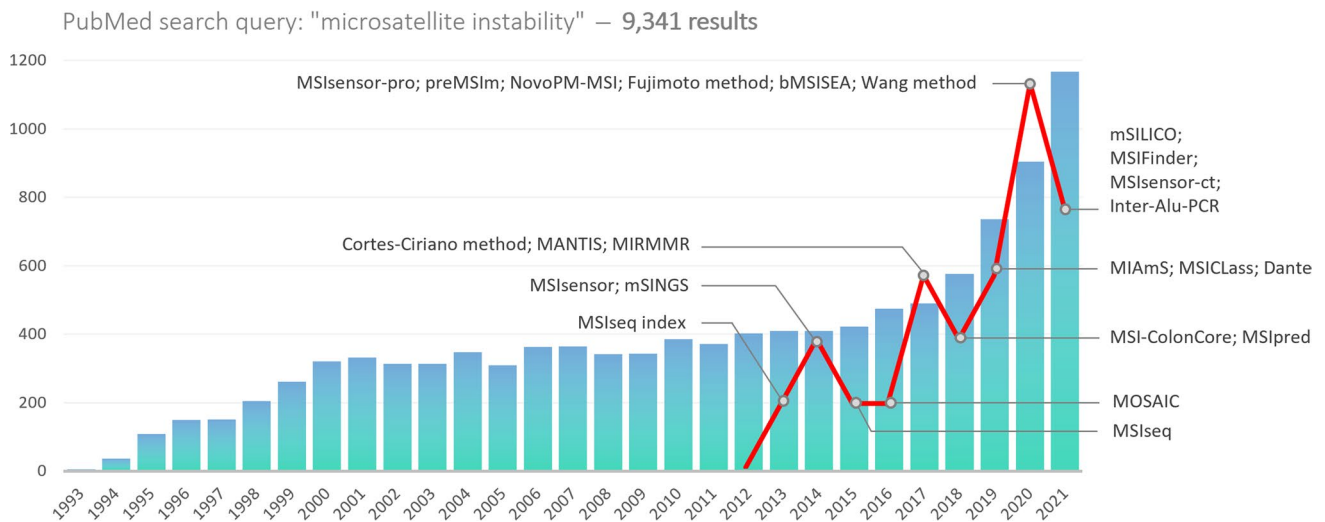


Fig. 1 Annual distribution of PubMed search results for the term “microsatellite instability” as of December 21, 2022. The red curve represents the number of emerging bioinformatics tools for MPS-

based MSI analysis. There seems to be a correlation between the increasing number of publications and the emergence of novel tools

and other metrics differ between these types of samples. The second important issue is related to the sequencing approach used for data generation. With the decreasing cost of DNA sequencing [131], targeted or panel sequencing, WES and even WGS strategies are becoming available. They differ in the range of the analysed regions and consequently in the amount of generated data that need to be processed in the downstream analysis. Some of the available tools can adequately work with all types of such data (MSIsensor-pro, MSIseq, Cortes-Ciriano method, mSING), while others are validated only on specific datasets (MSIsensor, MOSAIC, MANTIS, MSI-ColonCore, mSILICO, MIMcall, etc.) (see Table 4). Such tools may not process other datasets correctly, or data size itself may represent a limitation if the given tool has been designed for smaller panels. Many tools detect MSI, but are also capable of detecting TMB and may be also applied to predict the ICI efficacy [64–66].

Despite all the talk about decreasing costs, larger applications such as WGS are still too expensive for most labs worldwide. Low-coverage sequencing is a cheaper solution; however, to our knowledge, not many tools support the analysis of low-coverage data [119]. Technical differences between tools should be considered, including the programming language, availability, interface (command line and/or graphical user interface), up-to-date service and support. The choice of the right tool should be carefully considered based on the input data and the requirements of the downstream analysis. Ongoing development will lead to improved performance, thus offering new tools to enhance early prevention, diagnosis, prognosis and patient stratification regarding proper anti-cancer treatment and follow-up.

Liquid biopsy: an instrument of MSI-related cancer screening

Multiple tissue-based tests are available for MSI status assessment, mainly using IHC and PCR [3, 19–22]. Although they are regarded as a gold standard, they come with inherent limitations, mainly due to the invasive nature of tissue biopsy. Disadvantages include the difficulty or impossibility of repeated sampling, the requirement of an adequate amount of tumour tissue with a high concentration of cancer cells and the routine need for paired control healthy tissue that is unavailable or infeasible to obtain in some patients. Secondly, tumours typically display genetic heterogeneity, so test results may be ambiguous due to the heterogeneous expression of MMR proteins, and tissue biopsy may not fully capture tumour diversity. Moreover, test performance often requires specialised equipment, but also an expert pathologist who is needed to interpret the results (in the case of IHC), which, in turn, may be a subjective and error-prone process [132–134]. Thirdly, there is the possibility of false positive or false negative results due to artificial loss of expression, missense mutations in MMR genes or polymorphisms in MSs, which may be termed as technical limitations [100, 135]. Traditional sampling may not be applicable in patients with inadequate tumour tissue or tumour localisation with limited accessibility [66, 136].

Therefore, an increasing number of clinical laboratories are seeking to implement circulating tumour DNA (ctDNA) sequencing into their routine to provide 3PM care [137, 138]. It is expected that liquid biopsy-based methods

will develop further in the near future and will be widely applied in the detection of early cancer and other diseases, disease monitoring and personalised therapy [139]. Thus, using liquid biopsy in combination with MPS in MSI assessment is an attractive and promising approach with many advantages in the field of 3PM. Contrary to conventional tissue biopsy, the main benefit of liquid biopsy in MSI testing is the minimally invasive procedure without any pre-requirement and with the possibility of repeated sampling, which is essential to monitor disease progression, treatment response or resistance to therapy. ctDNA comprises genomic changes that are hallmarks of cancer and thus may be used as a potential surrogate for the entire tumour, allowing diagnosticians to deal with its genetic heterogeneity (Fig. 2) and capture the mutational landscape of the entire tumour or the TMB [140–142]. This approach seems to yield a high concordance rate with tissue biopsy-based detection and high specificity, precision and sensitivity, as reported by several authors [126, 143–145]. Liquid biopsy-based methods have the potential to enhance the utility of tumour detection assays to help

clinicians decide on immunotherapeutic treatment [143, 146].

Although liquid biopsy offers great potential in patient management [147–149], the assessment of MSI through ctDNA or cell-free DNA (cfDNA) is still in development due to various limitations. Several challenges need to be addressed before it is implemented into routine clinical practice and personalised medical service. One of the challenges is related to liquid biopsy in general because there is a lack of standard procedures, protocols and guidelines for sample collection, preparation, processing and storage, as well as for validation of studies. This could significantly affect the analytical stage and lead to difficulty in comparing results across different approaches [150–152]. Although liquid biopsy seems to be a sensitive procedure with high precision and accuracy [126, 143–145, 153], there is limited evidence regarding its sensitivity and accuracy in a diagnostic setting [154]. Next, there are cases with a low tumour fraction in the circulation (low concentration of cfDNA, ctDNA or CTCs). Artefacts or technical noise may also occur due

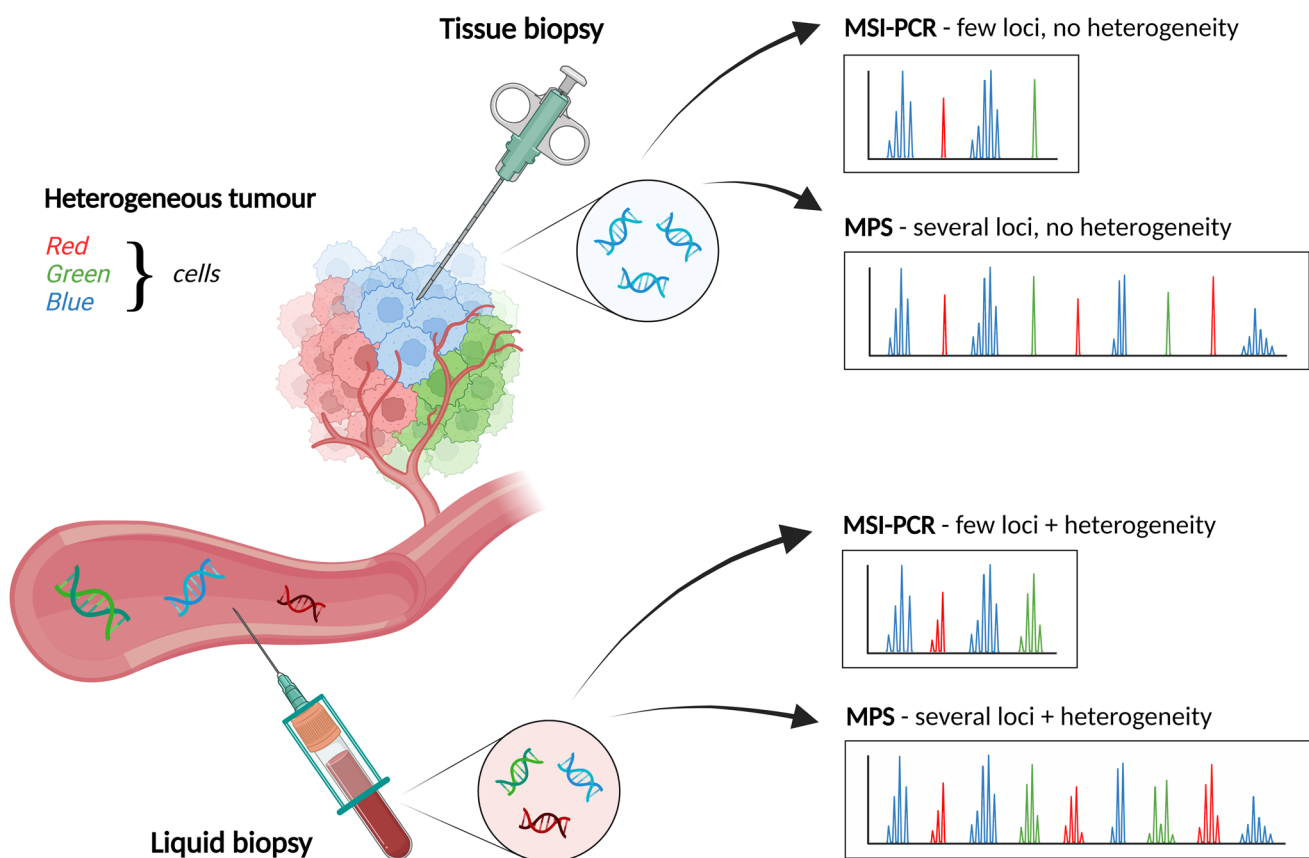


Fig. 2 Since tissue biopsy analyses only a selected portion of the tumour, it may not capture its overall heterogeneity (only blue cells). As DNA from the entire tumour enters the bloodstream, liquid biopsy is able to capture ctDNA from all populations of tumour cells (red,

green and blue cells). Given that MPS strategies can capture several MS markers in a single assay, these methods, combined with liquid biopsy, are the most representative approach for the personalisation of MSI detection (Created with BioRender.com)

to polymerase slippage resulting in false positive or false negative findings, while low signal-to-noise ratio may indicate contamination by non-tumour cells [151, 153, 155–157]. The final challenge to be addressed is related to the PCR- or MPS-based methodology used for analysis (see the previous chapters).

To summarise, detection of MSI status through liquid biopsy using MPS seems to be a promising approach, not only in early detection, predictive diagnostic and targeted prevention, but also in therapy management and consecutive monitoring of treatment response. Such approaches also offer diagnostic possibilities to patients at risk of developing MSI-associated disease even without a family history or to patients whose tumour tissue is unavailable, difficult to obtain or its acquisition is contraindicated. However, due to various technological and bioinformatic challenges, MSI detection through liquid biopsy using MPS is still in development, and further progress is needed before implementation into routine clinical practice and personalised medical services.

Current methods based on liquid biopsy for MSI status detection

To date, there is limited knowledge about the utility of liquid biopsy in determining the MSI status, but some approaches and methods are emerging (Table 3), mainly thanks to the significant progress made towards improving the resolution of existing tissue-based approaches and adapting them to

liquid biopsy. In recent years, many groups attempted to validate MPS-based approaches using plasma ctDNA for MSI detection with promising results [158, 159]. Using a ctDNA-based approach will also allow increased access to checkpoint inhibitors in a pan-tumour setting. This would be relevant in targeted prevention of cancers where routine MSI testing is contraindicated or when tissue is not available or accessible [159].

The potential of cfDNA in cancer diagnosis leads to the development of various tests, such as the OncoLBx test designed as a combination of a single-molecule sequencing (SMSEQ) platform with a pan-cancer MPS panel targeting 75 genes and five homopolymer marker sites. During its validation, OncoLBx demonstrated that it could detect variant allele frequencies of $\geq 0.1\%$ for SNVs and indels, $\geq 0.5\%$ for fusions, $\geq 4.5\%$ copies for CNVs and $\geq 2\%$ for MSI, displaying a specificity of $\geq 99.999\%$ for all variant types [160]. Willis et al. demonstrated robust analytic performance for MSI detection on the Guardant360® CDx test (Guardant Health, Redwood City, CA, USA), an FDA-approved liquid biopsy, which had been previously validated for the detection of four other variant types. This approach proved to be highly sensitive compared to tissue-based methods while maintaining very high specificity [143]. To our knowledge, it is the first ctDNA-based landscape analysis of MSI in a large advanced pan-cancer cohort. Georgiadis et al. developed a protocol for simultaneous MSI and TMB detection using a non-invasive method which is also effective in predicting response to immune checkpoint therapy. They achieved 100% sensitivity and specificity during the

Table 3 List of blood panel-based MPS strategies

Study	Panel/method	Sequencing platform	No. of tested loci	cfDNA input	Analytical performance/reference methods
[143]	Guardant360®CDx	NextSeq500/550 or HiSeq 2500	90 MS loci (MNRs, DNRs)	5–30 ng	LOD: 0.09–0.1% ctDNA content 86.6% sensitivity; 99.5% specificity/pentaplex MSI-PCR, IHC, MSI-NGS on tissues
[161]	FoundationOne® Liquid CDx	In-house MPS-based device	~2000 MS (MNRs, DNRs, TNRs)	20–30 ng	LOD: 0.8% unstable loci/NA
[184]	NA	MiSeq	24 MNRs	100 ng	98.0–100.0% sensitivity; 98.0–100.0% specificity/tissue MSI, MMR genes sequencing
[160, 161]	OncoLBX	MiSeq	5 MNRs	20–30 ng	LOD: 2.0% tumour fraction/NA
[146]	In-house	HiSeq 2000/2500	5 MNRs	5–250 ng	78.3% sensitivity (18/23); 100.0% specificity (6/6)/pentaplex MSI-PCR, IHC, MSI-NGS on tissues

Explanatory notes: *NA*, not available; *MNRs*, mononucleotide repeats; *DNRs*, dinucleotide repeats; *TNRs*, trinucleotide repeats; *LOD*, limit of detection; *ctDNA*, circulating tumour DNA; *MMR*, mismatch repair; *IHC*, immunohistochemistry

Table 4 List of computational tools

Sample	Tool	Strategy	Prog. Language	Description	Ref.
Tissue	Cortes-Ciriano method	WGS, WES	NA	The tool allows assessing the difference in reading length distribution at each locus using the statistical Kolmogorov–Smirnov test	[44]
	MSIsensor	WES	C + + , R	This software tool is suitable for deriving MSI status in standard paired tumour/normal sequencing data and automatically detecting somatic changes at MNRs. It computes length distributions of MS per site and subsequently uses these to compare observed distributions in both samples statistically	[104]
	MSIsensor-pro	WGS, WES, TGS	C + +	MSIsensor-pro is an updated freely available version of the MSIsensor tool that quantifies polymerase slippages for each tumour sample with no need for control	[185]
	MSIseq	WGS, WES, TGS	R	MSIseq is an R statistical language package for classifying MSI status based on four machine-learning frameworks operating on MPS data	[128]
	MSIseq index	RNAseq	NA	MSI prediction method utilises RNA sequencing data comparing indels in defined MSs vs all MSs	[130]
	MOSAIC	WES	Python, Perl, R	MOSAIC is a command-line interface tool whose algorithm is based on the average gain in the number of MS alleles and locus instability. It corrects for class imbalance in its cross-validation training procedure (with an approximately 3:1 MSS-to-MSI-H ratio), making predictions in new cancer types without any prior assumption about the expected prevalence of MSI-H tumours	[68]
	mSINGs	WES, TGS	Python	mSINGs is an MSI-detection tool based on the detection of altered read count distribution which evaluates each MS locus within a homopolymer marker set and reports MSI status by specifying certain cut-offs	[105]
	MANTIS	WES	Python	This tool could be incorporated into the existing pipelines, applicable to a variety of tumours with varying numbers of MNRs. It analyses the instability of paired samples by aggregating loci instead of individual locus differences to evaluate general instability present in a tumour sample since the data from the paired sample acts as an error-correcting baseline	[117]
	MSI-ColonCore	TGS	NA	This tool predicts the MSI status according to Z-score, sorting samples into MSI-H, MSI-L or MSS groups	[103]
	mSILICO	TGS	NA	mSILICO is a TGS-based computational method using a custom marker panel to assess MSI	[180]
	NA	WGS	Perl	Fujimoto method implements MIMcall (somatic indel caller) in calculating microsatellite error rates for the various MS motifs on the X chromosome. A sample is considered to be MSI if selected microsatellites show instability in $\geq 3\%$ loci	[55]
	MIAmS	TGS	NA	MIAmS (Microsatellites Instability by Amplicon Sequencing) is a command-line tool for routinely predicting MSI status without analysing normal tissue samples	[186]
	MSIpred	WES	Python	Microsatellite Instability Predictor, implementing SVM classifiers, is a robust tool for pan-tumour MSI classification, freely available as a Python 2 package	[114]
	preMSIm	TGS	R	preMSIm (Predicting MSI from mRNA) is a publicly available tool for predicting MSI from the expression profiling of a 15-gene panel	[187]
	NovoPM-MSI	TGS	NA	NovoPM-MSI is an algorithm that uses a target panel consisting of a set of enriched MNRs. MSI score is reported as a fraction of unstable loci by comparing length distribution in paired samples	[110]

Table 4 (continued)

Sample	Tool	Strategy	Prog. Language	Description	Ref.
	MSIFinder	TGS	Python	MSIFinder, a python package for automatic MSI classification that uses a random forest classifier (RFC)-based genome sequencing, a machine learning technology. The length of analysed MNRs ranged from 10 to 34 bp	[188]
	MIRMMR	WGS	R	MIRMMR predicts MSI status using methylation and mutation information and highlights genetic alterations significantly contributing to MSI screening	[189]
	Dante	WGS, WES, TGS	Python	A tool for analysing STRs from MPS data, originally designed for TGS data. It does not require the mapping of reads to a reference genome and is designed on the specific sequencing characteristics of STR alleles	[92]
Blood	bMSISEA	TGS/ColonCore panel	R	The blood MSI signature enrichment analysis is a method for blood-based MSI screening/100.0% (27/27) specificity; 94.1% sensitivity (16/17) for ctDNA content > 0.4%; 98.8% concordance between bMSISEA and IHC	[162]
	MSIsensor-ct	TGS	NA	A tool based on machine learning, dedicated to detecting MSI status using cfDNA sequencing data with a potential stable MSIScore threshold of 20%, with no requirement for pre-constructed baseline control/100.0% sensitivity; 100.0% specificity; LOD: 0.05% ctDNA with a minimum coverage of 3000×	[163]
	Wang method	WES	NA	This method enables detection of the difference in reading length distribution at each locus. The proportion of unstable loci among the selected 100 loci is used to calculate the blood-based MSI score/82.5% sensitivity (33/40); 92.6% specificity (201/209); 94.0% overall concordance (234/249); LOD: 0.5% ctDNA at an input of 30 ng DNA	[190]
	MSIClass	WGS	NA	A tool for MSI status classification from shallow WGS data which automatically detects somatic MSI events. The sequencing coverage as low as 0.05× is sufficient for accurately calling the MSI status	[191]
	Inter-Alu-PCR	TGS	NA	This tool combines MNR enrichment and a custom MSI-tracer algorithm to compare the MNR's read length. It can distinguish the cfDNA from cancer patients with MSI and samples from healthy individuals	[192]

Explanatory notes: *Prog. Language*, programming language; *Ref.*, references; *NA*, not available; *TGS*, targeted sequencing; *WES*, whole exome sequencing; *WGS*, whole genome sequencing; *MSI-H*, microsatellite instability-high; *MSI-L*, microsatellite instability-low; *MSS*, microsatellite stable; *cfDNA*, cell-free DNA; *ctDNA*, circulating tumour DNA; *STRs*, short tandem repeats; *MNRs*, mononucleotide repeats; *LOD*, limit of detection; *IHC*, immunohistochemistry

evaluation of their matched tumour and normal tissue samples [146].

A few months later, a different approach utilising a quantitative drop-off ddPCR was introduced (see Table 1). This method is able to target long and short repetitive genome regions, so it may be applied in MSI status determination and accurate ctDNA quantification. Although it is a 3-marker MSI-ddPCR assay, results suggest that it is possible to develop assays targeting additional markers identified by cancer exome sequencing for improving the clinical sensitivity of MSI testing across different cancer types. The method can also be applied to screen formalin-fixed and paraffin-embedded (FFPE) tumour samples and does not require paired tumour-normal samples for reliable MSI identification [144]. As it aims for absolute

quantification of ctDNA or MSI sequences, it has the potential to monitor ctDNA in patients longitudinally and may significantly improve 3PM strategies. Apart from the above-mentioned Guardant360® CDx, there is another FDA-approved liquid biopsy test, for tumour profiling, including MSI detection, FoundationOne® Liquid CDx (Foundation Medicine, Cambridge, MA, USA). In a validation study, the proposed test achieved high sensitivity and specificity and proved to be sufficiently robust and comparable to other MPS-based broad molecular profiling liquid biopsy assays [161].

The strategies applied by Guardant360® CDx, FoundationOne® Liquid CDx, OncoLBx and Georgiadis et al. (Table 3) all employ hybrid-capture enrichment of target regions and use molecular barcoding to avoid false positive

events arising due to technical PCR errors. The error correction process they use reduces background noise and enables variant calling with high sensitivity, specificity and accuracy. Along with blood MSI status, many of these approaches offer detection of other variant types, including the determination of blood-based tumour mutation burden (bTMB), a complementary predictive biomarker that may inform ICI treatment.

However, detection of MSI events from blood adds another specific challenge. Since the highly fragmented nature of ctDNA and the low ctDNA fraction in body fluids (as low as 0.01% in early stages of cancer) along with technical artefacts generated during library preparation, amplification or sequencing lessen the analytical sensitivity, determination of MSI from plasma samples requires methods that are, above all, highly sensitive [153]. The sensitivity of 98% and specificity of 100% in deep sequencing of samples with a minimum of 1% ctDNA by the ColonCore panel were observed using bMSISEA (blood MSI signature enrichment analysis) detection algorithm. This tool also showed high sensitivity of 94.1% and specificity of 100% with a minimum ctDNA fraction at 0.4% (Table 4). Moreover, it efficiently distinguished between MSI-H and MSS samples [162]. MSIsensor-ct [163] deployed on the NovaSeq platform (Illumina) using cfDNA sequencing data yielded 100% sensitivity and 100% specificity on 39 plasma specimens. The minimal ctDNA fraction for reliable assessment of the MSI status should be at least 0.05% with a sequencing depth over $3000\times$.

Conclusions and outlook in the framework of 3PM

End-of-life care for advanced cancer patients is generally considered of poor quality [164]. Moreover, huge expenditures are required during the patient's last year of life to increase the quality of palliative care. Since generalised approaches in cancer management come with an enormous socio-economic burden, personalised medicine is a vital aspect of palliative care. Various 3PM concepts are designed to develop maximum efforts to improve individual outcomes at the initial care level [165]. Implementing such a strategy is crucial to save lives while ensuring cost efficacy of medical services provided to the population [166]. Innovations in less invasive diagnosis and high-throughput technologies are accelerating, promoting and revolutionising a shift from decades-old testing strategies to reactive 3PM approaches [167, 168]. Since the role of MSI status as a predictive biomarker for cancer immunotherapy had already been clarified, the era of personalised medicine has dawned in regard to screening MSI across different levels of the human genome, from individual marker sites (numbers ranging from several to

hundreds of sites) to comprehensive large-scale pan-cancer analyses comprising thousands of evaluated loci.

MSI status as a promising biomarker with strong clinical significance used in the early diagnosis of LS and CRC patients positively correlates with survival outcomes and enables predicting the efficacy of immune checkpoint blockade therapy in solid tumours. This might improve individually tailored treatment decisions and follow-up strategies in MSI tumours. We hypothesise that comprehensive panels could now advocate universal MSI screening for all relevant individuals to provide better cancer prevention opportunities for them and their relatives. More effective treatment and surveillance for these individuals will result in increased survival and reduced costs of medical care. In addition, for the LS detection routine, MSI testing should be used as a universal MPS-based tumour screening approach. It turns out that most experts endorse the possibility of using it and recommend it as the first step in diagnosing LS and other diseases. Therefore, 3PM-based diagnostic approach using MPS technology is expected to significantly revolutionise the screening approach of this cancer-prone syndrome. Accurate diagnosis, precise management and accurate prevention are what the patient population urgently needs.

As their implementation into clinical practice is already in progress, 3PM-based approaches are becoming increasingly accessible to patients. However, to date, only a small number of them are applicable for the detection of MSI events in plasma, blood or tumour tissue with a low ctDNA fraction. As discussed in this review, technical and biological difficulties slowed down the introduction of these methods to clinical diagnostic laboratories. Therefore, newly available bioinformatics and computational algorithms must meet specific criteria, such as the ability to deal with unique characteristics of MS motifs, DNA slippage events or PCR errors introduced into homopolymeric sites during sequencing as well as alignment errors due to short read length.

We compiled a comprehensive list of high-throughput strategies and computational tools for MSI evaluation based on tissue and liquid biopsy. In the context of less invasive methods so far not used to the extent of their potential, there is growing evidence that ctDNA provides a highly concordant source of MSI-affected alleles. Liquid biopsy-derived methods thus increase patient access to standard-of-care-targeted therapies by quickly identifying the cases that are otherwise difficult to find due to tissue sampling limitations. With an increasing number of studies promoting the use of ctDNA and/or low-coverage sequencing data, many new in-house MSI screening tools are likely to be designed soon. In addition, using the MPS data enables analysis of other cancer-related signatures, facilitating the adoption of MSI clinical testing and supporting personalised cancer therapy. When the third generation of sequencing gains a foothold

in cancer research, it will bring along its own tools for data analysis. Moreover, in the age of individualised and precise treatment, these strategies can no longer be ignored.

Abbreviations 3PM/PPPM: Predictive, Preventive and Personalised Medicine; MSI: Microsatellite instability; MSI-H: Microsatellite instability-high; MSI-L: Microsatellite instability-low; MSS: Microsatellite stable; TR: Tandem repeat; MS: Microsatellite; STR: Short tandem repeat; MMR: Mismatch repair system; dMMR: Mismatch repair deficiency; MPS: Massively parallel sequencing; CRC: Colorectal cancer; EC: Endometrial cancer; LS: Lynch syndrome; IHC: Immunohistochemistry; FFPE: Formalin-fixed and paraffin-embedded; TMB: Tumour mutation burden; bTMT: Blood-based tumour mutation burden; ICI: Immune checkpoint inhibitor; WES: Whole-exome sequencing; WGS: Whole-genome sequencing; SNV: Single nucleotide variant; CNV: Copy number variant; ctDNA: Circulating tumour DNA; cfDNA: Cell-free DNA

Author contribution All authors contributed to the manuscript conception and design. Conceptualisation, investigation, writing — original draft preparation, visualisation and supervision by JS and ZP. Conceptualisation, investigation, writing — review and editing and visualisation by OP. Writing — review and editing, conceptualisation and supervision by JR and EHT. Writing — review and editing by GB. Writing — review and editing, supervision by JB. Investigation, writing — review and editing by DK. Supervision by VR and BN. Supervision and funding acquisition by TS. With contribution from ZP, JS prepared a first draft of the manuscript and all authors commented on previous version of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding was provided by the Operational Programme Integrated Infrastructure for the projects Biomedires II (ITMS: 313011W428), PreveLynch (ITMS: 313011V578) and LISPER (ITMS: 313011V446), co-financed by the European Regional Development Fund. Financial support was also obtained from GenoMicrosat, a grant from the Slovak Research and Development Agency (APVV-18-0319) and the PANGAIA project H2020-MSCA-RISE-2019 (Grant agreement ID: 872539) funded under H2020-EU.1.3.3. Programme.

Data availability Not applicable

Code availability Not applicable

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publications Not applicable.

Competing interests JS, ZP, OP, JR, JB and TS are employees of Geneton Ltd., participating in research and development activity that does not conflict with the content of the manuscript. VR is an employee of Medirex Group Academy n.p.o., participating in research and project activity with no potential conflict of interest to the proposed manuscript. EHT is an employee of Slovgen Ltd. and declares no conflict of interest. BN is an editorial board member of the EPMA journal, but had no involvement in, influence over, or access to the details of the peer

review process of this work. All remaining authors have no relevant financial or non-financial interests to disclose.

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/>.

References

1. Aaltonen LA, Peltomäki P, Leach FS, Sistonen P, Pylkkänen L, Mecklin JP, et al. Clues to the pathogenesis of familial colorectal cancer. *Science*. 1993;260:812–6. <https://doi.org/10.1126/science.8484121>.
2. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*. 1993;363:558–61. <https://doi.org/10.1038/363558a0>.
3. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science*. 1993;260:816–9. <https://doi.org/10.1126/science.8484122>.
4. Chambers GK, MacAvoy ES. Microsatellites: consensus and controversy. *Comp Biochem Physiol B Biochem Mol Biol*. 2000;126:455–76. [https://doi.org/10.1016/s0305-0491\(00\)00233-9](https://doi.org/10.1016/s0305-0491(00)00233-9).
5. Lindstrand A, Eisfeldt J, Pettersson M, Carvalho CMB, Kvarnang M, Grigelioniene G, et al. From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. *Genome Med*. 2019;11:68. <https://doi.org/10.1186/s13073-019-0675-1>.
6. Press MO, Carlson KD, Queitsch C. The overdue promise of short tandem repeat variation for heritability. *Trends Genet*. 2014;30:504–12. <https://doi.org/10.1016/j.tig.2014.07.008>.
7. Xiao W, Wu L, Yavas G, Simonyan V, Ning B, Hong H. Challenges, solutions, and quality metrics of personal genome assembly in advancing precision medicine. *Pharmaceutics*. 2016;8:15. <https://doi.org/10.3390/pharmaceutics8020015>.
8. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res*. 1998;58:5248–57.
9. Boland CR, Goel A. Microsatellite instability in colorectal cancer. *Gastroenterology*. 2010;138:2073–87.e3. <https://doi.org/10.1053/j.gastro.2009.12.064>.
10. Buglyó G, Styk J, Pös O, Csók Á, Repiska V, Soltész B, et al. Liquid biopsy as a source of nucleic acid biomarkers in the diagnosis and management of Lynch syndrome. *Int J Mol Sci*. 2022;23:18. <https://doi.org/10.3390/ijms23084284>.
11. Bacher JW, Sievers CK, Albrecht DM, Grimes IC, Weiss JM, Matkowskyj KA, et al. Improved detection of

- microsatellite instability in early colorectal lesions. *PLoS One*. 2015;10:e0132727. <https://doi.org/10.1371/journal.pone.0132727>.
12. Sepulveda AR, Hamilton SR, Allegra CJ. Molecular biomarkers for the evaluation of colorectal cancer: guideline from the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and the American Society of Clinical Oncology. *J Clin Oncol*. 2017;147:221–60. <https://doi.org/10.1093/ajcp/aqw209>.
 13. Hirotsu Y, Nagakubo Y, Amemiya K, Oyama T, Mochizuki H, Omata M. Microsatellite instability status is determined by targeted sequencing with MSICall in 25 cancer types. *Clin Chim Acta*. 2020;502:207–13. <https://doi.org/10.1016/j.cca.2019.11.002>.
 14. Redford L, Alhilal G, Needham S, O'Brien O, Coaker J, Tyson J, et al. A novel panel of short mononucleotide repeats linked to informative polymorphisms enabling effective high volume low cost discrimination between mismatch repair deficient and proficient tumours. *PLoS One*. 2018;13:e0203052. <https://doi.org/10.1371/journal.pone.0203052>.
 15. Phipps AI, Limburg PJ, Baron JA, Burnett-Hartman AN, Weisenberger DJ, Laird PW, et al. Association between molecular subtypes of colorectal cancer and patient survival. *Gastroenterology*. 2015;148:77–87.e2. <https://doi.org/10.1053/j.gastro.2014.09.038>.
 16. Battaglin F, Naseem M, Lenz H-J, Salem ME. Microsatellite instability in colorectal cancer: overview of its clinical significance and novel perspectives. *Clin Adv Hematol Oncol*. 2018;16:735–45.
 17. Janssens JP, Schuster K, Voss A. Preventive, predictive, and personalized medicine for effective and affordable cancer care. *EPMA J*. 2018;9:113–23. <https://doi.org/10.1007/s13167-018-0130-1>.
 18. Bizzarri M, Fedeli V, Monti N, Cucina A, Jalouli M, Alwasel SH, et al. Personalization of medical treatments in oncology: time for rethinking the disease concept to improve individual outcomes. *EPMA J*. 2021;12:545–58. <https://doi.org/10.1007/s13167-021-00254-1>.
 19. Modica I, Soslow RA, Black D, Tornos C, Kauff N, Shia J. Utility of immunohistochemistry in predicting microsatellite instability in endometrial carcinoma. *Am J Surg Pathol*. 2007;31:744–51. <https://doi.org/10.1097/01.pas.0000213428.61374.06>.
 20. Shia J. Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. *J Mol Diagn*. 2008;10:293–300. <https://doi.org/10.2353/jmol2008.080031>.
 21. Samaison L, Grall M, Staroz F, Uguen A. Microsatellite instability diagnosis using the fully automated Idylla platform: feasibility study of an in-house rapid molecular testing ancillary to immunohistochemistry in pathology laboratories. *J Clin Pathol*. 2019;72:830–5. <https://doi.org/10.1136/jclinpath-2019-205935>.
 22. Dietmaier W, Hofstädter F. Detection of microsatellite instability by real time PCR and hybridization probe melting point analysis. *Lab Invest*. 2001;81:1453–6. <https://doi.org/10.1038/labinvest.3780358>.
 23. Bodrova TA, Kostyushev DS, Antonova EN, Slavin S, Gnatenko DA, Bocharova MO, et al. Introduction into PPPM as a new paradigm of public health service: an integrative view. *EPMA J*. 2012;3:16. <https://doi.org/10.1186/1878-5085-3-16>.
 24. Tieng FYF, Abu N, Lee L-H, Ab Mutalib N-S. Microsatellite instability in colorectal cancer liquid biopsy-current updates on its potential in non-invasive detection, prognosis and as a predictive marker. *Diagnostics*. 2021;11:544. <https://doi.org/10.3390/diagnostics11030544>.
 25. Reyes GX, Schmidt TT, Kolodner RD, Hombauer H. New insights into the mechanism of DNA mismatch repair. *Chromosoma*. 2015;124:443–62. <https://doi.org/10.1007/s00412-015-0514-0>.
 26. Kunkel TA, Erie DA. Eukaryotic mismatch repair in relation to DNA replication. *Annu Rev Genet*. 2015;49:291–313. <https://doi.org/10.1146/annurev-genet-112414-054722>.
 27. Nagel ZD, Chaim IA, Samson LD. Inter-individual variation in DNA repair capacity: a need for multi-pathway functional assays to promote translational DNA repair research. *DNA Repair*. 2014;19:199–213. <https://doi.org/10.1016/j.dnarep.2014.03.009>.
 28. Tamura K, Kaneda M, Futagawa M, Takeshita M, Kim S, Nakama M, et al. Genetic and genomic basis of the mismatch repair system involved in Lynch syndrome. *Int J Clin Oncol*. 2019;24:999–1011. <https://doi.org/10.1007/s10147-019-01494-y>.
 29. Pino MS, Mino-Kenudson M, Wildemore BM, Ganguly A, Batten J, Sperduti I, et al. Deficient DNA mismatch repair is common in Lynch syndrome-associated colorectal adenomas. *J Mol Diagn*. 2009;11:238–247. <https://doi.org/10.2353/jmol2009.080142>.
 30. Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895–2015. *Nat Rev Cancer*. 2015;15:181–94. <https://doi.org/10.1038/nrc3878>.
 31. De' Angelis GL, Bottarelli L, Azzoni C, De' Angelis N, Leandro G, Di Mario F, et al. Microsatellite instability in colorectal cancer. *Acta Biomed*. 2018;89:97–101. <https://doi.org/10.23750/abm.v89i9-S.7960>.
 32. Kunkel TA. Evolving views of DNA replication (in)fidelity. *Cold Spring Harb Symp Quant Biol*. 2009;74:91–101. <https://doi.org/10.1101/sqb.2009.74.027>.
 33. Ganai RA, Johansson E. DNA replication—a matter of fidelity. *Mol Cell*. 2016;62:745–55. <https://doi.org/10.1016/j.molcel.2016.05.003>.
 34. Liu D, Keijzers G, Rasmussen LJ. DNA mismatch repair and its many roles in eukaryotic cells. *Mutat Res*. 2017;773:174–87. <https://doi.org/10.1016/j.mrrev.2017.07.001>.
 35. Andrianova MA, Bazykin GA, Nikolaev SI, Seplyarskiy VB. Human mismatch repair system balances mutation rates between strands by removing more mismatches from the lagging strand. *Genome Res*. 2017;27:1336–43. <https://doi.org/10.1101/gr.219915.116>.
 36. Haradhvala NJ, Kim J, Maruvka YE, Polak P, Rosebrock D, Livitz D, et al. Distinct mutational signatures characterize concurrent loss of polymerase proofreading and mismatch repair. *Nat Commun*. 2018;9:1746. <https://doi.org/10.1038/s41467-018-04002-4>.
 37. Gupta D, Heinen CD. The mismatch repair-dependent DNA damage response: mechanisms and implications. *DNA Repair*. 2019;78:60–9. <https://doi.org/10.1016/j.dnarep.2019.03.009>.
 38. Groothuizen FS, Sixma TK. The conserved molecular machinery in DNA mismatch repair enzyme structures. *DNA Repair*. 2016;38:14–23. <https://doi.org/10.1016/j.dnarep.2015.11.012>.
 39. do Amaral-Silva G, do Amaral-Silva GK, Martins MD, Pontes HAR, Fregnani ER, Lopes MA, et al. Mismatch repair system proteins in oral benign and malignant lesions. *J Oral Pathol Med*. 2017;46:241–5. <https://doi.org/10.1111/jop.12484>.
 40. Radvanszky J, Hyblova M, Radvanska E, Spalek P, Valachova A, Magyarova G, et al. Characterisation of non-pathogenic premutation-range myotonic dystrophy type 2 alleles. *J Clin Med Res*. 2021;10:3934. <https://doi.org/10.3390/jcm10173934>.
 41. Radvanszky J, Surovy M, Polak E, Kadasi L. Uninterrupted CCTG tracts in the myotonic dystrophy type 2 associated locus. *Neuromuscul Disord*. 2013;23:591–8. <https://doi.org/10.1016/j.nmd.2013.02.013>.
 42. Depienne C, Mandel J-L. 30 years of repeat expansion disorders: what have we learned and what are the remaining challenges? *Am J Hum Genet*. 2021;108:764–85. <https://doi.org/10.1016/j.ajhg.2021.03.011>.

43. Kim T-M, Laird PW, Park PJ. The landscape of microsatellite instability in colorectal and endometrial cancer genomes. *Cell*. 2013;155:858–68. <https://doi.org/10.1016/j.cell.2013.10.015>.
44. Cortes-Ciriano I, Lee S, Park W-Y, Kim T-M, Park PJ. A molecular portrait of microsatellite instability across multiple cancers. *Nat Commun*. 2017;8:15180. <https://doi.org/10.1038/ncomms15180>.
45. Chung H, Lopez CG, Holmstrom J, Young DJ, Lai JF, Ream-Robinson D, et al. Both microsatellite length and sequence context determine frameshift mutation rates in defective DNA mismatch repair. *Hum Mol Genet*. 2010;19:2638–47. <https://doi.org/10.1093/hmg/ddq151>.
46. Gymrek M, Willems T, Guilmatre A, Zeng H, Markus B, Georgiev S, et al. Abundant contribution of short tandem repeats to gene expression variation in humans. *Nat Genet*. 2016;48:22–9. <https://doi.org/10.1038/ng.3461>.
47. Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Corrigendum: classification and characterization of microsatellite instability across 18 cancer types. *Nat Med*. 2018;24:525. <https://doi.org/10.1038/nm0418-525a>.
48. Jung B, Doctolero RT, Tajima A, Nguyen AK, Keku T, Sandler RS, et al. Loss of activin receptor type 2 protein expression in microsatellite unstable colon cancers. *Gastroenterology*. 2004;126:654–9. <https://doi.org/10.1053/j.gastro.2004.01.008>.
49. Yuza K, Nagahashi M, Ichikawa H, Hanyu T, Nakajima M, Shimada Y, et al. Activin a receptor type 2A mutation affects the tumor biology of microsatellite instability-high gastric cancer. *J Gastrointest Surg*. 2021;25:2231–41. <https://doi.org/10.1007/s11605-020-04889-9>.
50. Giannakis M, Hodis E, Jasmine MuX, Yamauchi M, Rosenbluh J, Cibulskis K, et al. RNF43 is frequently mutated in colorectal and endometrial cancers. *Nat Genet*. 2014;46:1264–6. <https://doi.org/10.1038/ng.3127>.
51. Metzgar D, Byttof J, Wills C. Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res*. 2000;10:72–80.
52. Nouri Nojaded J, Hashemzadeh S, SamadiKafil H, Behrouz Sharif S, Eftekharsadat A, Ghasemnejad T, et al. Evaluation of microsatellite instability in tumor and tumor marginal samples of sporadic colorectal cancer using mononucleotide markers. *Excli J*. 2018;17:945–51. <https://doi.org/10.17179/excli.2018-1455>.
53. Lai Y, Sun F. The relationship between microsatellite slippage mutation rate and the number of repeat units. *Mol Biol Evol*. 2003;20:2123–31. <https://doi.org/10.1093/molbev/msg228>.
54. Wren JD, Forgaes E, Fondon JW 3rd, Pertsemliadis A, Cheng SY, Gallardo T, et al. Repeat polymorphisms within gene regions: phenotypic and evolutionary implications. *Am J Hum Genet*. 2000;67:345–56. <https://doi.org/10.1086/303013>.
55. Fujimoto A, Fujita M, Hasegawa T, Wong JH, Maejima K, Oku-Sasaki A, et al. Comprehensive analysis of indels in whole-genome microsatellite regions and microsatellite instability across 21 cancer types. *Genome Res*. 2020;30:334–46. <https://doi.org/10.1101/gr.255026.119>.
56. Ejima Y, Yang L, Sasaki MS. Aberrant splicing of the ATM gene associated with shortening of the intronic mononucleotide tract in human colon tumor cell lines: a novel mutation target of microsatellite instability. *Int J Cancer*. 2000;86:262–8. [https://doi.org/10.1002/\(sici\)1097-0215\(20000415\)86:2%3c262::aid-ijc17%3e3.0.co;2-i](https://doi.org/10.1002/(sici)1097-0215(20000415)86:2%3c262::aid-ijc17%3e3.0.co;2-i).
57. Giannini G, Rinaldi C, Ristori E, Ambrosini MI, Cerignoli F, Viel A, et al. Mutations of an intronic repeat induce impaired MRE11 expression in primary human cancer with microsatellite instability. *Oncogene*. 2004;23:2640–7. <https://doi.org/10.1038/sj.onc.1207409>.
58. Suraweera N, Iacopetta B, Duval A, Compoin A, Tubacher E, Hamelin R. Conservation of mononucleotide repeats within 3' and 5' untranslated regions and their instability in MSI-H colorectal cancer. *Oncogene*. 2001;20:7472–7. <https://doi.org/10.1038/sj.onc.1204952>.
59. Ma J, DeFrances MC, Zou C, Johnson C, Ferrell R, Zarnegar R. Somatic mutation and functional polymorphism of a novel regulatory element in the HGF gene promoter causes its aberrant expression in human breast cancer. *J Clin Invest*. 2009;119:478–91. <https://doi.org/10.1172/JCI36640>.
60. Woerner SM, Yuan YP, Benner A, Korff S, von Knebel DM, Bork P. SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. *Nucleic Acids Res*. 2010;38:D682–9. <https://doi.org/10.1093/nar/gkp839>.
61. Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM, et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N Engl J Med*. 2003;349:247–57. <https://doi.org/10.1056/NEJMoa022289>.
62. Bertagnolli MM, Niedzwiecki D, Compton CC, Hahn HP, Hall M, Damas B, et al. Microsatellite instability predicts improved response to adjuvant therapy with irinotecan, fluorouracil, and leucovorin in stage III colon cancer: Cancer and Leukemia Group B Protocol 89803. *J Clin Oncol*. 2009;27:1814–21. <https://doi.org/10.1200/JCO.2008.18.2071>.
63. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol*. 2019;16:361–75. <https://doi.org/10.1038/s41575-019-0126-x>.
64. André T, Shiu K-K, Kim TW, Jensen BV, Jensen LH, Punt C, et al. Pembrolizumab in microsatellite-instability-high advanced colorectal cancer. *N Engl J Med*. 2020;383:2207–18. <https://doi.org/10.1056/NEJMoa2017699>.
65. Overman MJ, Lonardi S, Wong KYM, Lenz H-J, Gelsomino F, Aglietta M, et al. Durable clinical benefit with nivolumab plus ipilimumab in DNA mismatch repair-deficient/microsatellite instability-high metastatic colorectal cancer. *J Clin Oncol*. 2018;36:773–9. <https://doi.org/10.1200/JCO.2017.76.9901>.
66. Lenz H-J, Van Cutsem E, Luisa Limon M, Wong KYM, Hendlisz A, Aglietta M, et al. First-line nivolumab plus low-dose ipilimumab for microsatellite instability-high/mismatch repair-deficient metastatic colorectal cancer: the phase II CheckMate 142 study. *J Clin Oncol*. 2021;40:161–70. <https://doi.org/10.1200/JCO.21.01015>.
67. Nakajima H, Nakatsura T. Towards the era of immune checkpoint inhibitors and personalized cancer immunotherapy. *Immunol Med*. 2021;44:10–5. <https://doi.org/10.1080/25785826.2020.1785654>.
68. Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med*. 2016;22:1342–50. <https://doi.org/10.1038/nm.4191>.
69. Samowitz WS, Curtin K, Ma KN, Schaffer D, Coleman LW, Leppert M, et al. Microsatellite instability in sporadic colon cancer is associated with an improved prognosis at the population level. *Cancer Epidemiol Biomarkers Prev*. 2001;10:917–23.
70. Kloor M, von Knebel DM. The immune biology of microsatellite-unstable cancer. *Trends Cancer Res*. 2016;2:121–33. <https://doi.org/10.1016/j.trecan.2016.02.004>.
71. Kloor M, Reuschenbach M, Pauligk C, Karbach J, Rafiyan M-R, Al-Batran S-E, et al. A frameshift peptide neoantigen-based vaccine for mismatch repair-deficient cancers: a phase I/IIa clinical trial. *Clin Cancer Res*. 2020;26:4503–10. <https://doi.org/10.1158/1078-0432.CCR-19-3517>.

72. Abida W, Cheng ML, Armenia J, Middha S, Autio KA, Vargas HA, et al. Analysis of the prevalence of microsatellite instability in prostate cancer and response to immune checkpoint blockade. *Jama Oncol.* 2019;5:471–8. <https://doi.org/10.1001/jamaoncol.2018.5801>.
73. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med.* 2016;375:1823–33. <https://doi.org/10.1056/NEJMoa1606774>.
74. Hellmann MD, Paz-Ares L, Bernabe Caro R, Zurawski B, Kim S-W, Carcereny Costa E, et al. Nivolumab plus ipilimumab in advanced non-small-cell lung cancer. *N Engl J Med.* 2019;381:2020–31. <https://doi.org/10.1056/NEJMoa1910231>.
75. André T, Cohen R, Salem ME. Immune checkpoint blockade therapy in patients with colorectal cancer harboring microsatellite instability/mismatch repair deficiency in 2022. *Am Soc Clin Oncol Educ Book.* 2022;42:1–9. https://doi.org/10.1200/EDBK_349557.
76. Middha S, Zhang L, Nafa K, Jayakumaran G, Wong D, Kim HR, et al. Reliable pan-cancer microsatellite instability assessment by using targeted next-generation sequencing data. *JCO Precis Oncol.* 2017;2017:PO.17.00084. <https://doi.org/10.1200/PO.17.00084>.
77. Albacker LA, Wu J, Smith P, Warmuth M, Stephens PJ, Zhu P, et al. Loss of function JAK1 mutations occur at high frequency in cancers with microsatellite instability and are suggestive of immune evasion. *PLoS One.* 2017;12:e0176181. <https://doi.org/10.1371/journal.pone.0176181>.
78. Sveen A, Johannessen B, Tengs T, Danielsen SA, Eilertsen IA, Lind GE, et al. Multilevel genomics of colorectal cancers with microsatellite instability-clinical impact of JAK1 mutations and consensus molecular subtype 1. *Genome Med.* 2017;9:46. <https://doi.org/10.1186/s13073-017-0434-0>.
79. Luke JJ, Bao R, Sweis RF, Spranger S, Gajewski TF. WNT/ β -catenin pathway activation correlates with immune exclusion across human cancers. *Clin Cancer Res.* 2019;25:3074–83. <https://doi.org/10.1158/1078-0432.CCR-18-1942>.
80. Pai SG, Carneiro BA, Mota JM, Costa R, Leite CA, Barroso-Sousa R, et al. Wnt/ β -catenin pathway: modulating anticancer immune response. *J Hematol Oncol.* 2017;10:101. <https://doi.org/10.1186/s13045-017-0471-6>.
81. Trabucco SE, Gowen K, Maund SL, Sanford E, Fabrizio DA, Hall MJ, 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;21:1053–66. <https://doi.org/10.1016/j.jmol.2019.06.011>.
82. Snell RG, MacMillan JC, Cheadle JP, Fenton I, Lazarou LP, Davies P, et al. Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nat Genet.* 1993;4:393–7. <https://doi.org/10.1038/ng0893-393>.
83. Spandidos DA, Ergazaki M, Arvanitis D, Kiaris H. Microsatellite instability in human atherosclerotic plaques. *Biochem Biophys Res Commun.* 1996;220:137–40. <https://doi.org/10.1006/bbrc.1996.0370>.
84. Hatzistamou J, Kiaris H, Ergazaki M, Spandidos DA. Loss of heterozygosity and microsatellite instability in human atherosclerotic plaques. *Biochem Biophys Res Commun.* 1996;225:186–90. <https://doi.org/10.1006/bbrc.1996.1151>.
85. Spandidos DA, Koumantakis E, Sifakis S, Sourvinos G. Microsatellite mutations in spontaneously aborted embryos. *Fertil Steril.* 1998;70:892–5. [https://doi.org/10.1016/s0015-0282\(98\)00279-9](https://doi.org/10.1016/s0015-0282(98)00279-9).
86. McCarthy AJ, Capo-Chichi J-M, Spence T, Grenier S, Stockley T, Kamel-Reid S, et al. Heterogenous loss of mismatch repair (MMR) protein expression: a challenge for immunohistochemical interpretation and microsatellite instability (MSI) evaluation. *Hip Int.* 2019;5:115–29. <https://doi.org/10.1002/cjp2.120>.
87. Janavicius R, Matiukaite D, Jakubauskas A, Griskevicius L. Microsatellite instability detection by high-resolution melting analysis. *Clin Chem.* 2010;56:1750–7. <https://doi.org/10.1373/clinchem.2010.150680>.
88. Susanti S, Fadhil W, Ebili HO, Asiri A, Nestarenkaite A, Hadjimichael E, et al. N_LyST: a simple and rapid screening test for Lynch syndrome. *J Clin Pathol.* 2018;71:713–20. <https://doi.org/10.1136/jclinpath-2018-205013>.
89. Gilson P, Levy J, Rouyer M, Demange J, Husson M, Bonnet C, et al. Evaluation of 3 molecular-based assays for microsatellite instability detection in formalin-fixed tissues of patients with endometrial and colorectal cancers. *Sci Rep.* 2020;10:16386. <https://doi.org/10.1038/s41598-020-73421-5>.
90. Kurian AW, Hare EE, Mills MA, Kingham KE, McPherson L, Whittemore AS, et al. Clinical evaluation of a multiple-gene sequencing panel for hereditary cancer risk assessment. *J Clin Oncol.* 2014;32:2001–9. <https://doi.org/10.1200/JCO.2013.53.6607>.
91. Shen T, de Pajaro-Van Stadt SH, Yeat NC, Lin JCH. Clinical applications of next generation sequencing in cancer: from panels, to exomes, to genomes. *Front Genet.* 2015;6:215. <https://doi.org/10.3389/fgene.2015.00215>.
92. Budiš J, Kucharík M, Ďuriš F, Gazdarica J, Zrubcová M, Ficek A, et al. Dante: genotyping of known complex and expanded short tandem repeats. *Bioinformatics.* 2019;35:1310–7. <https://doi.org/10.1093/bioinformatics/bty791>.
93. Dolzhenko E, Bennett MF, Richmond PA, Trost B, Chen S, van Vugt JJFA, et al. ExpansionHunter Denovo: a computational method for locating known and novel repeat expansions in short-read sequencing data. *Genome Biol.* 2020;21:102. <https://doi.org/10.1186/s13059-020-02017-z>.
94. Mojarad BA, Engchuan W, Trost B, Backstrom I, Yin Y, Thiruvahindrapuram B, et al. Genome-wide tandem repeat expansions contribute to schizophrenia risk. *Mol Psychiatry.* 2022;27:3692–8. <https://doi.org/10.1038/s41380-022-01575-x>.
95. Hannan AJ. Tandem repeats mediating genetic plasticity in health and disease. *Nat Rev Genet.* 2018;19:286–98. <https://doi.org/10.1038/nrg.2017.115>.
96. Hampel H, Pearlman R, Beightol M, Zhao W, Jones D, Frankel WL, et al. Assessment of tumor sequencing as a replacement for lynch syndrome screening and current molecular tests for patients with colorectal cancer. *JAMA Oncol.* 2018;4:806–13. <https://doi.org/10.1001/jamaoncol.2018.0104>.
97. Goel A, Nagasaka T, Hamelin R, Boland CR. An optimized pentaplex PCR for detecting DNA mismatch repair-deficient colorectal cancers. *PLoS One.* 2010;5:e9393. <https://doi.org/10.1371/journal.pone.0009393>.
98. Boyle TA, Bridge JA, Sabatini LM, Nowak JA, Vasalos P, Jennings LJ, et al. Summary of microsatellite instability test results from laboratories participating in proficiency surveys: proficiency survey results from 2005 to 2012. *Arch Pathol Lab Med.* 2014;138:363–70. <https://doi.org/10.5858/arpa.2013-0159-CP>.
99. Network CGA. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012;487:330–7. <https://doi.org/10.1038/nature11252>.
100. Luchini C, Bibeau F, Ligtenberg MJL, Singh N, Nottagegar A, Bosse T, 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;30:1232–43. <https://doi.org/10.1093/annonc/mdz116>.

101. Nowak JA, Yurgelun MB, Bruce JL, Rojas-Rudilla V, Hall DL, Shivdasani P, et al. Detection of mismatch repair deficiency and microsatellite instability in colorectal adenocarcinoma by targeted next-generation sequencing. *J Mol Diagn*. 2017;19:84–91. <https://doi.org/10.1016/j.jmol.2016.07.010>.
102. Hempelmann JA, Scroggins SM, Pritchard CC, Salipante SJ. MSIplus for integrated colorectal cancer molecular testing by next-generation sequencing. *J Mol Diagn*. 2015;17:705–14. <https://doi.org/10.1016/j.jmol.2015.05.008>.
103. Zhu L, Huang Y, Fang X, Liu C, Deng W, Zhong C, et al. A novel and reliable method to detect microsatellite instability in colorectal cancer by next-generation sequencing. *J Mol Diagn*. 2018;20:225–31. <https://doi.org/10.1016/j.jmol.2017.11.007>.
104. Niu B, Ye K, Zhang Q, Lu C, Xie M, McLellan MD, et al. MSIsensor: microsatellite instability detection using paired tumor-normal sequence data. *Bioinformatics*. 2014;30:1015–6. <https://doi.org/10.1093/bioinformatics/btt755>.
105. Salipante SJ, Scroggins SM, Hampel HL, Turner EH, Pritchard CC. Microsatellite instability detection by next generation sequencing. *Clin Chem*. 2014;60:1192–9. <https://doi.org/10.1373/clinchem.2014.223677>.
106. Vanderwalde A, Spetzler D, Xiao N, Gatalica Z, Marshall J. Microsatellite instability status determined by next-generation sequencing and compared with PD-L1 and tumor mutational burden in 11,348 patients. *Cancer Med*. 2018;7:746–56. <https://doi.org/10.1002/cam4.1372>.
107. Waalkes A, Smith N, Penewit K, Hempelmann J, Konnick EQ, Hause RJ, et al. Accurate pan-cancer molecular diagnosis of microsatellite instability by single-molecule molecular inversion probe capture and high-throughput sequencing. *Clin Chem*. 2018;64:950–8. <https://doi.org/10.1373/clinchem.2017.285981>.
108. Kim JE, Chun S-M, Hong YS, Kim K-P, Kim SY, Kim J, et al. Mutation burden and I index for detection of microsatellite instability in colorectal cancer by targeted next-generation sequencing. *J Mol Diagn*. 2019;21:241–50. <https://doi.org/10.1016/j.jmol.2018.09.005>.
109. Pabla S, Andreas J, Lenzo FL, Burgher B, Hagen J, Giamo V, et al. Development and analytical validation of a next-generation sequencing based microsatellite instability (MSI) assay. *Oncotarget*. 2019;10:5181–93. <https://doi.org/10.18632/oncotarget.27142>.
110. Zhao L, Shan G, Li L, Yu Y, Cheng G, Zheng X. A robust method for the rapid detection of microsatellite instability in colorectal cancer. *Oncol Lett*. 2020;20:1982–8. <https://doi.org/10.3892/ol.2020.11702>.
111. Liu L, et al. Microsatellite instability testing and lynch syndrome screening for colorectal cancer patients through tumour sequencing. *Ann Oncol*. 2019;30(5):v574. <https://doi.org/10.1093/annonc/mdz257.001>.
112. Zhao C, Jiang T, Ju JH, Zhang S, Tao J, Fu Y, et al. TruSight oncology 500: enabling comprehensive genomic profiling and biomarker reporting with targeted sequencing. *BioRxiv*. 2020. <https://doi.org/10.1101/2020.10.21.349100v1.abstract>
113. Pawlik TM, Raut CP, Rodriguez-Bigas MA. Colorectal carcinogenesis: MSI-H versus MSI-L. *Dis Markers*. 2004;20:199–206. <https://doi.org/10.1155/2004/368680>.
114. Wang C, Liang C. MSIPred: a python package for tumor microsatellite instability classification from tumor mutation annotation data using a support vector machine. *Sci Rep*. 2018;8:17546. <https://doi.org/10.1038/s41598-018-35682-z>.
115. Ebili HO, Agboola AO, Rakha E. MSI-WES: a simple approach for microsatellite instability testing using whole exome sequencing. *Future Oncol*. 2021;17:3595–606. <https://doi.org/10.2217/fon-2021-0132>.
116. Kubiritova Z, Gyuraszova M, Nagyova E, Hyblova M, Harsanyova M, Budis J, et al. On the critical evaluation and confirmation of germline sequence variants identified using massively parallel sequencing. *J Biotechnol*. 2019;298:64–75. <https://doi.org/10.1016/j.jbiotec.2019.04.013>.
117. Kautto EA, Bonneville R, Miya J, Yu L, Krook MA, Reeser JW, et al. Performance evaluation for rapid detection of pan-cancer microsatellite instability with MANTIS. *Oncotarget*. 2017;8:7452–63. <https://doi.org/10.18632/oncotarget.13918>.
118. Maruvka YE, Mouw KW, Karlic R, Parasuraman P, Kamburov A, Polak P, et al. Analysis of somatic microsatellite indels identifies driver events in human tumors. *Nat Biotechnol*. 2017;35:951–9. <https://doi.org/10.1038/nbt.3966>.
119. Maruvka YE, Frazer R, Grimsby J, Van Seventer EE, Cibulskis C, Adalsteinsson V, et al. Abstract A094: MSIClass - Identification and classification of microsatellite unstable tumors using cell-free DNA from ultra low pass sequencing. *DNA Repair and Modulation*. 2019;18(12):A094. <https://doi.org/10.1158/1535-7163.targ-19-a094>.
120. Fazekas A, Steeves R, Newmaster S. Improving sequencing quality from PCR products containing long mononucleotide repeats. *Biotechniques*. 2010;48:277–85. <https://doi.org/10.2144/000113369>.
121. Minoche AE, Dohm JC, Himmelbauer H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol*. 2011;12:R112. <https://doi.org/10.1186/gb-2011-12-11-r112>.
122. Takehara Y, Nagasaka T, Nyuya A, Haruma T, Haraga J, Mori Y, et al. Accuracy of four mononucleotide-repeat markers for the identification of DNA mismatch-repair deficiency in solid tumors. *J Transl Med*. 2018;16:5. <https://doi.org/10.1186/s12967-017-1376-4>.
123. Wu S, Liu X, Wang J, Zhou W, Guan M, Liu Y, et al. DNA mismatch repair deficiency detection in colorectal cancer by a new microsatellite instability analysis system. *Interdiscip Sci*. 2020;12:145–54. <https://doi.org/10.1007/s12539-020-00358-8>.
124. Do H, Dobrovic A. Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin Chem*. 2015;61:64–71. <https://doi.org/10.1373/clinchem.2014.223040>.
125. Kim S, Park C, Ji Y, Kim DG, Bae H, van Vrancken M, et al. Deamination effects in formalin-fixed, paraffin-embedded tissue samples in the era of precision medicine. *J Mol Diagn*. 2017;19:137–46. <https://doi.org/10.1016/j.jmol.2016.09.006>.
126. Deng A, Yang J, Lang J, Jiang Z, Wang W, Yuan D, et al. Monitoring microsatellite instability (MSI) in circulating tumor DNA by next-generation DNA-seq. *J Clin Oncol*. 2018;36(15):12025–12025. https://doi.org/10.1200/JCO.2018.36.15_suppl.12025.
127. Krishnan GS, Sunil Krishnan G, Joshi A, Kaushik V. Bioinformatics in personalized medicine. In: Singh V, Kumar A (eds) *Advances in bioinformatics*. Singapore: Springer; 2021. pp. 303–15. https://doi.org/10.1007/978-981-33-6191-1_15
128. Huang MN, McPherson JR, Cutcutache I, Teh BT, Tan P, Rozen SG. MSIsq: software for assessing microsatellite instability from catalogs of somatic mutations. *Sci Rep*. 2015;5:13321. <https://doi.org/10.1038/srep13321>.
129. Stadler ZK, Battaglin F, Middha S, Hechtman JF, Tran C, Cercek A, et al. Reliable detection of mismatch repair deficiency in colorectal cancers using mutational load in next-generation sequencing panels. *J Clin Oncol*. 2016;34:2141–7. <https://doi.org/10.1200/JCO.2015.65.1067>.
130. Lu Y, Soong TD, Elemento O. A novel approach for characterizing microsatellite instability in cancer cells. *PLoS One*. 2013;8:e63056. <https://doi.org/10.1371/journal.pone.0063056>.
131. DNA Sequencing Costs: Data. 2021 <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>. Accessed 11 Feb 2022.
132. Tieng FYF, Abu N, Lee LH, Mutalib NSA. Microsatellite instability in colorectal cancer liquid biopsy-current updates on its

- potential in non-invasive detection, prognosis and as a predictive marker. *Diagnostics*. 2021;11:544. <https://doi.org/10.3390/diagnostics11030544>.
133. Gilson P, Merlin J-L, Harlé A. Detection of microsatellite instability: state of the art and future applications in circulating tumour DNA (ctDNA). *Cancers*. 2021;13:1491. <https://doi.org/10.3390/cancers13071491>.
 134. Overbeek LIH, Ligtenberg MJL, Willems RW, Hermens RPMG, Blokk WAM, Dubois SV, et al. Interpretation of immunohistochemistry for mismatch repair proteins is only reliable in a specialized setting. *Am J Surg Pathol*. 2008;32:1246–51. <https://doi.org/10.1097/pas.0b013e31816401bb>.
 135. Cohen R, Hain E, Buhard O, Guilloux A, Bardier A, Kaci R, et al. Association of primary resistance to immune checkpoint inhibitors in metastatic colorectal cancer with misdiagnosis of microsatellite instability or mismatch repair deficiency status. *JAMA Oncol*. 2019;5:551–5. <https://doi.org/10.1001/jamaoncol.2018.4942>.
 136. Marcus L, Lemery SJ, Keegan P, Pazdur R. FDA approval summary: pembrolizumab for the treatment of microsatellite instability-high solid tumors. *Clin Cancer Res*. 2019;25:3753–8. <https://doi.org/10.1158/1078-0432.CCR-18-4070>.
 137. Styk J, Buglyó G, Pös O, Csók Á, Soltész B, Lukasz P, et al. Extracellular nucleic acids in the diagnosis and progression of colorectal cancer. *Cancers*. 2022;14:3712. <https://doi.org/10.3390/cancers14153712>
 138. Gahan PB. Circulating nucleic acids in plasma and serum: diagnosis and prognosis in cancer. *EPMA J*. 2010;1:503–12. <https://doi.org/10.1007/s13167-010-0021-6>.
 139. Liskova A, Samec M, Koklesova L, Giordano FA, Kubatka P, Golubnitschaja O. Liquid biopsy is instrumental for 3PM dimensional solutions in cancer management. *J Clin Med Res*. 2020;9:2749. <https://doi.org/10.3390/jcm9092749>.
 140. Osumi H, Shinozaki E, Yamaguchi K, Zembutsu H. Clinical utility of circulating tumor DNA for colorectal cancer. *Cancer Sci*. 2019;110:1148–55. <https://doi.org/10.1111/cas.13972>.
 141. Kasi PM. Mutational burden on circulating cell-free tumor-DNA testing as a surrogate marker of mismatch repair deficiency or microsatellite instability in patients with colorectal cancers. *J Gastrointest Oncol*. 2017;8:747–8. <https://doi.org/10.21037/jgo.2017.06.05>.
 142. Gargalionis AN, Papavassiliou AG. Liquid biopsies in colorectal cancer: monitoring genetic heterogeneity. *Trends Cancer Res*. 2017;3:166–8. <https://doi.org/10.1016/j.trecan.2017.01.003>.
 143. Willis J, Lefterova MI, Artyomenko A, Kasi PM, Nakamura Y, Mody K, et al. Validation of microsatellite instability detection using a comprehensive plasma-based genotyping panel. *Clin Cancer Res*. 2019;25:7035–45. <https://doi.org/10.1158/1078-0432.CCR-19-1324>.
 144. Silveira AB, Bidard F-C, Kasperek A, Melaabi S, Tanguy M-L, Rodrigues M, et al. High-accuracy determination of microsatellite instability compatible with liquid biopsies. *Clin Chem*. 2020;66:606–13. <https://doi.org/10.1093/clinchem/hvaa013>.
 145. Odegaard JI, Vincent JJ, Mortimer S, Vowles JV, Ulrich BC, Banks KC, et al. Validation of a plasma-based comprehensive cancer genotyping assay utilizing orthogonal tissue- and plasma-based methodologies. *Clin Cancer Res*. 2018;24:3539–49. <https://doi.org/10.1158/1078-0432.CCR-17-3831>.
 146. Georgiadis A, Durham JN, Keefer LA, Bartlett BR, Zielonka M, Murphy D, et al. Noninvasive detection of microsatellite instability and high tumor mutation burden in cancer patients treated with PD-1 blockade. *Clin Cancer Res*. 2019;25:7024–34. <https://doi.org/10.1158/1078-0432.CCR-19-1372>.
 147. Kubiritova Z, Radvanszky J, Gardlik R. Cell-free nucleic acids and their emerging role in the pathogenesis and clinical management of inflammatory bowel disease. *Int J Mol Sci*. 2019;20:3662. <https://doi.org/10.3390/ijms20153662>.
 148. Szilágyi M, Pös O, Márton É, Buglyó G, Soltész B, Keserű J, et al. Circulating cell-free nucleic acids: main characteristics and clinical application. *Int J Mol Sci*. 2020;21:6827. <https://doi.org/10.3390/ijms21186827>.
 149. Pös O, Biró O, Szemes T, Nagy B. Circulating cell-free nucleic acids: characteristics and applications. *Eur J Hum Genet*. 2018;26:937–45. <https://doi.org/10.1038/s41431-018-0132-4>.
 150. Pös Z, Pös O, Styk J, Mocova A, Strieskova L, Budis J, et al. Technical and methodological aspects of cell-free nucleic acids analyses. *Int J Mol Sci*. 2020;21:8634. <https://doi.org/10.3390/ijms21228634>.
 151. Fleischhacker M, Schmidt B. Pre-analytical issues in liquid biopsy – where do we stand? *J Lab Med*. 2020;44:117–42. <https://doi.org/10.1515/labmed-2019-0167>.
 152. Neumann MHD, Bender S, Krahn T, Schlange T. ctDNA and CTCs in liquid biopsy - current status and where we need to progress. *Comput Struct Biotechnol J*. 2018;16:190–5. <https://doi.org/10.1016/j.csbj.2018.05.002>.
 153. Wang L, Ajani JA. Ushering in liquid biopsy for the microsatellite status: advantages and caveats. *Clin Cancer Res*. 2019;25:6887–9. <https://doi.org/10.1158/1078-0432.CCR-19-2585>.
 154. Ding Y, Li W, Wang K, Xu C, Hao M, Ding L. Perspectives of the application of liquid biopsy in colorectal cancer. *Biomed Res Int*. 2020;2020:6843180. <https://doi.org/10.1155/2020/6843180>.
 155. Castro-Giner F, Gkountela S, Donato C, Alborelli I, Quagliata L, Ng CKY, et al. Cancer diagnosis using a liquid biopsy: challenges and expectations. *Diagnostics*. 2018;8:31. <https://doi.org/10.3390/diagnostics8020031>.
 156. Bettgowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6:224ra24. <https://doi.org/10.1126/scitranslmed.3007094>.
 157. Makrigiorgos G, Ladas I, Mamon HJ, Ng K, Yu F, Leong CK, et al. Sensitive detection of microsatellite instability (MSI) in liquid biopsies from early stage colon cancer patients using nuclease-based enrichment and standard-marker or nGS based approaches. *Int J Radiat Oncol Biol Phys*. 2019;105:S24–5. <https://doi.org/10.1016/j.ijrobp.2019.06.427>.
 158. Maron SB, Chase LM, Lomnicki S, Kochanny S, Moore KL, Joshi SS, et al. Circulating tumor DNA sequencing analysis of gastroesophageal adenocarcinoma. *Clin Cancer Res*. 2019;25:7098–112. <https://doi.org/10.1158/1078-0432.CCR-19-1704>.
 159. Gowen K, Clark TA, Gregg JP, Greene MZ, Murphy A, White J, et al. MSI-H testing via hybrid capture based NGS sequencing of liquid biopsy samples. *J Clin Oncol*. 2019;37(4):504–504. https://doi.org/10.1200/JCO.2019.37.4_suppl.504.
 160. Atkins A, Gupta P, Zhang BM, Tsai W-S, Lucas J, Javey M, et al. Detection of circulating tumor DNA with a single-molecule sequencing analysis validated for targeted and immunotherapy selection. *Mol Diagn Ther*. 2019;23:521–35. <https://doi.org/10.1007/s40291-019-00406-0>.
 161. Woodhouse R, Li M, Hughes J, Delfosse D, Skoletsky J, Ma P, et al. Clinical and analytical validation of FoundationOne Liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin. *PLoS One*. 2020;15:e0237802. <https://doi.org/10.1371/journal.pone.0237802>.
 162. Cai Z, Wang Z, Liu C, Shi D, Li D, Zheng M, et al. Detection of microsatellite instability from circulating tumor DNA by targeted deep sequencing. *J Mol Diagn*. 2020;22:860–70. <https://doi.org/10.1016/j.jmol.2020.04.210>.
 163. Han X, Zhang S, Zhou DC, Wang D, He X, Yuan D, et al. MSIsensor-ct: microsatellite instability detection using cfDNA

- sequencing data. *Brief Bioinform.* 2021;22:bbaa402. <https://doi.org/10.1093/bib/bbaa402>
164. Davis MP, Vanenkevort EA, Elder A, Young A, Correa Ordóñez ID, Wojtowicz MJ, et al. The financial impact of palliative care and aggressive cancer care on end-of-life health care costs. *Am J Hosp Palliat Care.* 2023;40:52–60. <https://doi.org/10.1177/10499091221098062>.
 165. Ellinger J, Alajati A, Kubatka P, Giordano FA, Ritter M, Costigliola V, et al. Prostate cancer treatment costs increase more rapidly than for any other cancer-how to reverse the trend? *EPMA J.* 2022;13:1–7. <https://doi.org/10.1007/s13167-022-00276-3>.
 166. Crigna AT, Samec M, Koklesova L, Liskova A, Giordano FA, Kubatka P, et al. Cell-free nucleic acid patterns in disease prediction and monitoring-hype or hope? *EPMA J.* 2020;11:603–27. <https://doi.org/10.1007/s13167-020-00226-x>.
 167. Golubnitschaja O. Time for new guidelines in advanced healthcare: the mission of The EPMA Journal to promote an integrative view in predictive, preventive and personalized medicine. *EPMA J.* 2012;3:5. <https://doi.org/10.1186/1878-5085-3-5>.
 168. Golubnitschaja O, Costigliola V. Common origin but individual outcomes: time for new guidelines in personalized healthcare. *Per Med.* 2010;7:561–8. <https://doi.org/10.2217/pme.10.42>.
 169. Buhard O, Cattaneo F, Wong YF, Yim SF, Friedman E, Flejou J-F, et al. Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. *J Clin Oncol.* 2006;24:241–51. <https://doi.org/10.1200/JCO.2005.02.7227>.
 170. Piciotti R, Venetis K, Sajjadi E, Fusco N. Mismatch repair status characterization in oncologic pathology: taking stock of the real-world possibilities. *J Mol Pathol.* 2021;2:93–100. <https://doi.org/10.3390/jmp2020009>.
 171. PlentiPlex™ MSI . Pentabase A/S. PentaBase. 2022. <https://pentabase.com/plentiplex-msi/>. Accessed 5 Aug 2022.
 172. TrueMark MSI Assay for Microsatellite Instability Analysis. 2022. <https://www.thermofisher.com/uk/en/home/clinical/clinical-genomics/molecular-oncology-solutions/microsatellite-instability-assay.html>. Accessed 5 Aug 2022.
 173. Pécriaux A, Favre L, Calderaro J, Charpy C, Derman J, Pujals A. Detection of microsatellite instability in a panel of solid tumours with the Idylla MSI Test using extracted DNA. *J Clin Pathol.* 2021;74:36–42. <https://doi.org/10.1136/jclinpath-2020-206581>.
 174. Förster I, Brockmann M, Schildgen O, Schildgen V. Microsatellite instability testing in colorectal cancer using the QiaXcel advanced platform. *BMC Cancer.* 2018;18:484. <https://doi.org/10.1186/s12885-018-4400-z>.
 175. How-Kit A, Daunay A, Buhard O, Meiller C, Sahbatou M, Collura A, et al. Major improvement in the detection of microsatellite instability in colorectal cancer using HSP110 T17 E-ice-COLD-PCR. *Hum Mutat.* 2018;39:441–53. <https://doi.org/10.1002/humu.23379>.
 176. Berardinelli GN, Scapulatempo-Neto C, Durães R, de Antônio Oliveira M, Guimarães D, Reis RM. Advantage of (T17) marker inclusion for microsatellite instability (MSI) detection in colorectal cancer patients. *Oncotarget.* 2018;9:28691–701.
 177. Buhard O, Lagrange A, Guilloux A, Colas C, Chouchène M, Wanherdrick K, et al. HSP110 T17 simplifies and improves the microsatellite instability testing in patients with colorectal cancer. *J Med Genet.* 2016;53:377–84. <https://doi.org/10.1136/jmedgenet-2015-103518>.
 178. Malapelle U, Parente P, Pepe F, De Luca C, Cerino P, Covelli C, et al. Impact of pre-analytical factors on MSI test accuracy in mucinous colorectal adenocarcinoma: a multi-assay concordance study. *Cells.* 2020;9:2019. <https://doi.org/10.3390/cells9092019>.
 179. VarTrace MSI qPCR assay. NuProbe. 2021. <https://nuprobe.com/vartrace-msi/>. Accessed 5 Aug 2022.
 180. Lee Y, Lee JA, Park HE, Han H, Kim Y, Bae JM, et al. Targeted next-generation sequencing-based detection of microsatellite instability in colorectal carcinomas. *PLoS One.* 2021;16:e0246356. <https://doi.org/10.1371/journal.pone.0246356>.
 181. Gan C, Love C, Beshay V, Macrae F, Fox S, Waring P, et al. Applicability of next generation sequencing technology in microsatellite instability testing. *Genes.* 2015;6:46–59. <https://doi.org/10.3390/genes6010046>.
 182. Pritchard CC, Morrissey C, Kumar A, Zhang X, Smith C, Coleman I, et al. Complex MSH2 and MSH6 mutations in hypermutated microsatellite unstable advanced prostate cancer. *Nat Commun.* 2014;5:4988. <https://doi.org/10.1038/ncomms5988>.
 183. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science.* 2017;357:409–13. <https://doi.org/10.1126/science.aan6733>.
 184. Gallon R, Mühlegger B, Wenzel S-S, Sheth H, Hayes C, Aretz S, et al. A sensitive and scalable microsatellite instability assay to diagnose constitutional mismatch repair deficiency by sequencing of peripheral blood leukocytes. *Hum Mutat.* 2019;40:649–55. <https://doi.org/10.1002/humu.23721>.
 185. Lomate PA, Mane MV. Efficacy of multimodal analgesia with perineural buprenorphine or dexmedetomidine for surgeries performed under ultrasound-guided infraclavicular brachial plexus block. *J Anaesthesiol Clin Pharmacol.* 2020;36:66–71. https://doi.org/10.4103/joacp.JOACP_30_18.
 186. Escudié F, Van Goethem C, Grand D, Vendrell J, Vigier A, Brousset P, et al. MIAMs: microsatellite instability detection on NGS amplicons data. *Bioinformatics.* 2019;36:1915–6. <https://doi.org/10.1093/bioinformatics/btz797>.
 187. Li L, Feng Q, Wang X. PreMSIm: An R package for predicting microsatellite instability from the expression profiling of a gene panel in cancer. *Comput Struct Biotechnol J.* 2020;18:668–75. <https://doi.org/10.1016/j.csbj.2020.03.007>.
 188. Zhou T, Chen L, Guo J, Zhang M, Zhang Y, Cao S, et al. MSI-Finder: a python package for detecting MSI status using random forest classifier. *BMC Bioinformatics.* 2021;22:185. <https://doi.org/10.1186/s12859-021-03986-z>.
 189. Foltz SM, Liang W-W, Xie M, Ding L. MIRMMR: binary classification of microsatellite instability using methylation and mutations. *Bioinformatics.* 2017;33:3799–801. <https://doi.org/10.1093/bioinformatics/btx507>.
 190. Wang Z, Zhao X, Gao C, Gong J, Wang X, Gao J, et al. Plasma-based microsatellite instability detection strategy to guide immune checkpoint blockade treatment. *J Immunother Cancer.* 2020;8:e001297. <https://doi.org/10.1136/jitc-2020-001297>.
 191. Maruvka YE, Frazer R, Grimsby J, Cibulskis C, Adalsteinsson V, Corcoran R, et al. Abstract 2677: *MSIClass*—Identification and classification of microsatellite unstable using cell-free DNA from ultra-low-pass sequencing. *Cancer Res.* 2019;25:2677–2677. <https://doi.org/10.1158/1538-7445.am2019-2677>.
 192. Yu F, Leong KW, Makrigiorgos A, Adalsteinsson VA, Ladas I, Ng K, et al. NGS-based identification and tracing of microsatellite instability from minute amounts DNA using inter-Alu-PCR. *Nucleic Acids Res.* 2021;49:e24. <https://doi.org/10.1093/nar/gkaa1175>.

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