REVIEW ARTICLE

Genetics and Genomics

Integrative analysis of multi-omics data for liquid biopsy

Geng Chen [1,2✉], Jing Zhang[1], Qiaoting Fu[3], Valerie Taly [4✉] and Fei Tan [1,3✉]

The innovation of liquid biopsy holds great potential to revolutionise cancer management through early diagnosis and timely treatment of cancer. Integrative analysis of different tumour-derived omics data (such as genomics, epigenetics, fragmentomics, and proteomics) from body fluids for cancer detection and monitoring could outperform the analysis of single modality data alone. In this review, we focussed on the discussion of early cancer detection and molecular residual disease surveillance based on multi-omics data of blood. We summarised diverse types of tumour-derived components, current popular platforms for profiling cancer-associated signals, machine learning approaches for joint analysis of liquid biopsy data, as well as multi-omics-based early detection of cancers, molecular residual disease monitoring, and treatment response surveillance. We also discussed the challenges and future directions of multi-omics-based liquid biopsy. With the development of both experimental protocols and computational methods dedicated to liquid biopsy, the implementation of multi-omics strategies into the clinical workflow will likely benefit the clinical management of cancers including decision-making guidance and patient outcome improvement.

BACKGROUND

Cancer is one of the leading causes of death worldwide and a wide range of strategies have been developed to reduce cancer-related morbidity and mortality. In particular, liquid biopsy is becoming increasingly popular for early cancer detection and monitoring based on the detection of tumour biomarkers circulating in body fluids (e.g., blood) [1, 2]. Compared to traditional tissue biopsy, liquid biopsy is less burdensome and could enable faster and more cost-efficient evaluation of cancer profiles in a real-time and non-invasive manner [1, 3]. Moreover, conventional tissue biopsies may have sampling bias unless different tumour areas are investigated simultaneously, while liquid biopsy could provide more comprehensive cross-sectional information on tumour heterogeneity [4, 5]. Furthermore, a diversity of cancer-related components can be explored in body fluids, such as circulating tumour DNA (ctDNA), circulating tumour cells (CTCs), circulating tumour RNAs (ctRNAs), tumour-educated platelets (TEPs), as well as tumour-derived extracellular vesicles (EVs) (e.g., exosomes) (Fig. 1). These tumour-associated signals are mainly shed from primary and/or metastatic cancer sites. As reviewed previously, blood tests based on different types of circulating biomarkers have different advantages and limitations in cancer management [6–8]. Among them, ctDNA was the most profiled component, which was mainly derived from apoptosis, necrosis, and secretion of tumour cells, and was widely used as the signal for cancer detection and monitoring [9–11]. For instance, qualitative and quantitative ctDNA analysis of pre-operation could effectively predict the prognosis and postoperative recurrence of cancer patients, which may benefit the selection of appropriate patients for treatment [12]. Mutation profiling based on ctDNA before treatment can also provide useful guidance for the choice of targeted drugs or immunotherapy [13–15]. Due to the advantages of liquid biopsy, a shift from tumour biopsy to liquid biopsy has begun in the clinical management of cancers [6–8].

The applications of liquid biopsy can span the entire journey of tumour patients, including early cancer screening and diagnosis [16–18], molecular residual disease (MRD) detection and monitoring [19–21], surveillance of treatment response and resistance [22], and various others. Early cancer detection could provide a valuable time window for curative treatment and long-term survival of patients [2], which may become a routine part of health checks in the future. Moreover, MRD monitoring would allow the detection of disease recurrence and treatment response earlier than traditional methods (e.g., radiographic imaging), enabling the guidance and earlier intervention of cancer management [6, 23]. However, liquid biopsy approaches must be accurate, sensitive, and robust, because the concentrations of ctDNA and other types of tumour-related components in peripheral blood are typically low, particularly in patients with early-stage cancer or after definitive treatment [24].

Initial liquid biopsy developments mainly focussed on the detection of driver or pathogenic mutations in consideration of the leveraged signals highly specific to cancers. But the performance of those methods in the scenarios of early cancer

[1]Shanghai Skin Disease Hospital, School of Medicine, Tongji University, 200443 Shanghai, China. [2]Center for Bioinformatics and Computational Biology, School of Life Sciences, East China Normal University, 200241 Shanghai, China. [3]Shanghai Skin Disease Clinical College, The Fifth Clinical Medical College, Anhui Medical University, 200443 Shanghai, China. [4]Université de Paris, UMR-S1138, CNRS SNC5096, Équipe labélisée Ligue Nationale contre le cancer, Centre de Recherche des Cordeliers, Paris, France. ✉email: chengeng66666@outlook.com; valerie.taly@parisdescartes.fr; tanfeitrue@126.com
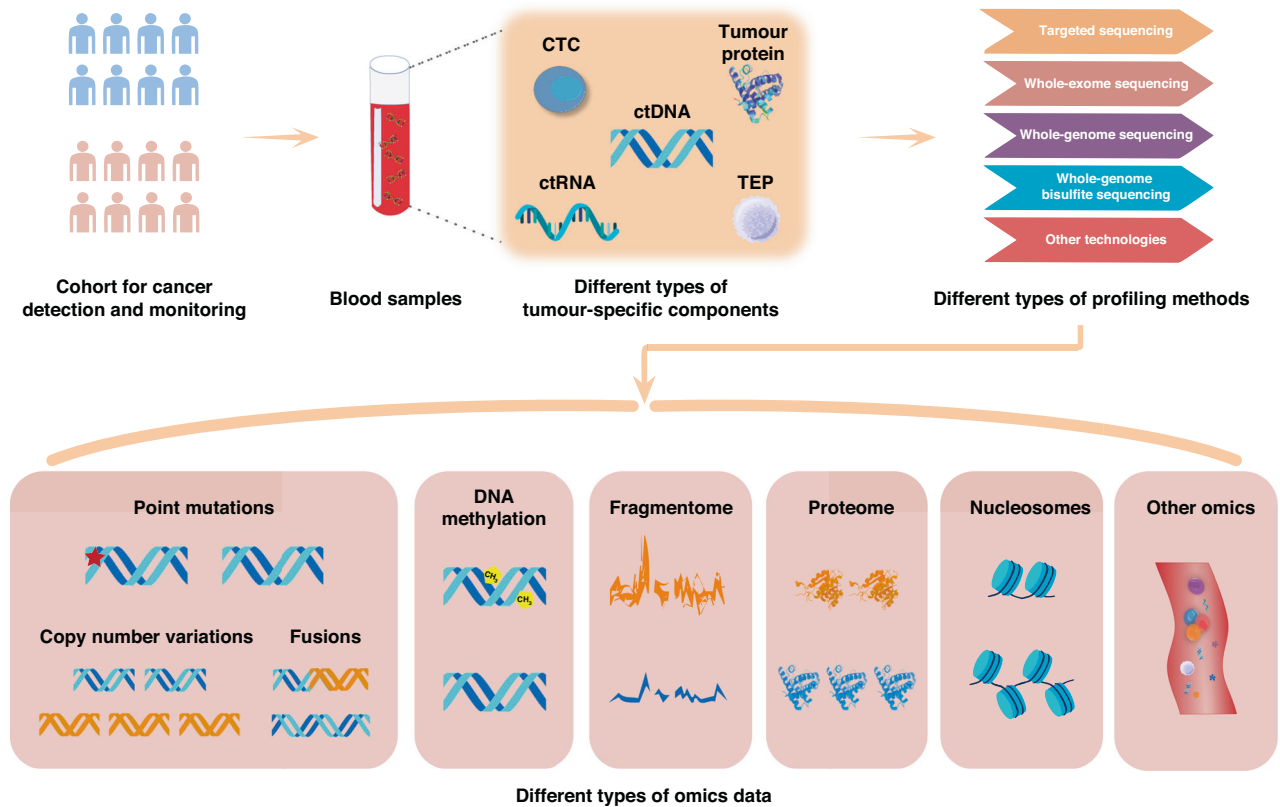
Fig. 1 Overview of diverse tumour-associated components in the blood of cancer patients. These tumour-related signals can be profiled with different approaches (e.g., targeted sequencing, WES, WGS, WGBS, among others), which are valuable for early cancer detection and MRD monitoring.

detection and MRD monitoring is limited due to insufficient biomarkers and related confounding factors (such as clonal haematopoiesis of indeterminate potential (CHIP)) [1, 8]. Unlike cancer-specific genetic variants, epigenetic alterations (e.g., DNA methylation changes, post-translational histone modifications, and nucleosome position changes) of tumours could be more prevalent, providing great potential for liquid biopsy [17, 25–30]. Other types of circulating biomarkers in blood, such as serum protein biomarkers [31], DNA fragment sizes [32–34], end motifs of plasma DNA [35], cell-free transcriptome [36], and even the microbiome signatures [37] are also important cancer-related signals. Furthermore, recent studies also showed that the inclusion of clinical risk factors (e.g., age and smoking history), and CT imaging can further boost the sensitivity of blood-based assays [16, 29, 38]. Since different omics layers are complementary for interrogating cancers as discussed previously [39, 40], integrative analysis of distinct modalities to improve the performance of liquid biopsy is increasingly popular [8]. However, no reviews have comprehensively summarised and discussed the progress of multi-omics-based liquid biopsies for early cancer detection and monitoring.

Considering that blood-based biopsy is the most studied among different body fluids, this Review mainly focusses on the discussion of multi-omics-based liquid biopsy by integrative analysis of ctDNA and other types of tumour-associated components in blood. We first summarise different tumour-related signals in plasma and currently popular platforms used in liquid biopsy. Then we describe different machine learning (ML) approaches employed to process blood-based biopsy data. Next, we discuss the progress in multi-omics-based early cancer detection, MRD monitoring, and treatment response surveillance through joint analysis of multi-omics data. We also outline the challenges and future directions of liquid biopsy.

TUMOUR-ASSOCIATED BIOMARKERS USED IN BLOOD-BASED LIQUID BIOPSY
Peripheral blood contains a variety of circulating biomarkers, including ctDNA, CTCs, ctRNAs, TEPs, circulating EVs, proteins, metabolites, and even viral sequences (e.g., HBVs) (Fig. 1). At the genomic level, various genetic alterations can be obtained from ctDNA through DNA sequencing, including point mutations, indels, copy number variations (CNVs), gene fusions, and other types of structural variations (SVs) [2]. For epigenetic alterations, DNA methylation is stable and highly frequent for specific genomic regions [41], providing abundant signals for early cancer detection and the determination of tissue of origin (TOO) [17]. The cell-free transcriptome is valuable for gene expression analysis and the identification of cancer-associated fusion transcripts and alternative splicing events [36, 42]. Furthermore, the fragmentation differences between tumour-derived and normal cfDNA (e.g., DNA fragment size, end-motif of fragments, and start-stop positions) were valuable biomarkers for liquid biopsy [8]. For example, Cristiano et al. achieved high performance in tumour detection and TOO identification for seven cancer types based on the fragmentation feature inferred from low-coverage whole-genome sequencing (WGS) [32]. Besides, some studies showed that tumour-related proteome [43, 44] and metabolome [45] in body fluids could provide important signals as well. These different types of components in the blood are promising for cancer detection and monitoring.

Initial liquid biopsy studies mainly focussed on the identification of known driver gene mutations, since they likely have a causative role in tumourigenesis, by employing PCR-based assays on plasma cfDNA from cancer patients [46–48]. Subsequently, the application of next-generation sequencing (NGS) technologies to targeted regions or the whole exome or genome enable parallel detection

of abundant cancer-specific alterations [5, 49, 50]. One main challenge is that the biological confounding factors (e.g., CHIP) [51] and related technical noise may generate a marked effect on mutation calling. By contrast, DNA methylation alterations in cancer are much more prevalent than canonical mutations and could promise a higher limit of detection [17]. Thus the epigenomic analysis of ctDNA is becoming increasingly popular in the liquid biopsy field [52, 53]. However, the ctDNA concentration is considerably low at the early stages of cancer and after definitive treatment, the performance of those assays only based on one type of data could be limited. Since different omics layers are complimentary for cancer profiling, one effective way to improve the performance of blood-based tests is the joint analysis of multiple types of cancer-related signals [8].

CURRENT POPULAR PLATFORMS AND MACHINE LEARNING APPROACHES FOR LIQUID BIOPSY
At present, PCR-based and NGS-based approaches are commonly used for liquid biopsy. PCR-based assays include sensitive qPCR-based methods (e.g., ARMS [54], Intplex [55], and COLD-PCR [56]), digital PCR-based techniques (e.g., ddPCR [57] and BEAMing [58]), among others [59, 60], having advantages in terms of very high sensitivity (0.01%~0.001%), rapidness, and cost-effectiveness. However, these approaches largely depend on the known mutations of primary tumours and can only track a limited number of variants. Contrariwise, NGS-based strategies allow the interrogation of both hot-spot and non-hot-spot areas, as well as different omics layers of the tumour. Targeted and untargeted sequencing are two main categories. Targeted approaches generally focus on a set of predefined genes/regions (e.g., gene panels regarding mutations or DNA methylation) [61], while untargeted ones can profile the entire exome/genome/methy-lome, such as whole-exome sequencing (WES), WGS, and whole-genome bisulfite sequencing (WGBS). With the help of prior knowledge of known cancer-specific markers, targeted methods may achieve higher sensitivity in a cost-effective way than untargeted ones [1]. But they cannot identify de novo tumour alterations like untargeted strategies. However, untargeted approaches often require higher DNA input and are more expensive than targeted strategies [62]. Thus, targeted and untargeted strategies have their own strength and weakness for liquid biopsy.

It has been suggested that the assays based on DNA methylation could outperform mutation-based sequencing strategies (e.g., WGS and targeted sequencing panels) in early cancer detection [63, 64]. One main reason could be that genome-wide DNA methylation profile enables to interrogate more abundant tumour-related signals compared to the relatively limited number of cancer-specific mutations [52]. Of note, both WGS and WGBS can provide other informative modalities besides mutations and DNA methylation level changes, such as DNA fragment sizes, nucleosome position maps, fragment end motifs, etc [8, 32, 34]. Joint analysis of those different types of omics data from the plasma to enhance the assay performance holds great promise for improving cancer detection and monitoring [8, 59].

Liquid biopsy studies often involve a large number of samples and many features, generally producing large-scale, high-dimensional, and complicated data sets. ML approaches can easily identify the trends and patterns within large volumes of data, which are very suitable for the setting of liquid biopsy. Many studies have shown that traditional ML algorithms, such as linear models [65], support vector machine (SVM) [66], and random forest (RF) [67], performed well in the data analysis of cancer detection [18, 34, 68–71]. However, traditional ML algorithms often heavily rely on the representation of the selected informative features which are difficult to obtain in certain scenarios [72]. Furthermore, most of the previous studies regarding liquid biopsy were mainly based on one type of data. Integrative analysis of multi-omics data with ML methods is increasingly popular to enhance the performance of liquid biopsy [8]. Advanced methods like ensemble learning can combine the outputs resulting from multiple ML models to improve decisions, it is appealing to improve the model performance and robustness by compensating the weaknesses of one learning model with the strengths of another one [73]. Besides, deep learning (such as the multilayer perceptron, also known as neural network) approaches can efficiently learn features from unprocessed, unstructured, and high-dimensional data [74, 75], holding promise to process the increasing scale of liquid biopsy data. Nevertheless, it is worth noting that if the sample size of the dataset is small, ML models are vulnerable to overfitting, especially for deep learning methods. A series of detailed suggestions regarding the preparations before model building, model evaluation, model comparison, and result reporting were proposed recently [76]. Such guidelines could be valuable for avoiding the ML pitfalls in liquid biopsy. Thus, the implementation of ML models for integrating multi-modal, multi-scale, as well as longitudinal data from liquid biopsies may greatly benefit the detection of weak cancer-related signals in the blood [77].

EARLY CANCER DETECTION BASED ON MULTI-OMICS DATA
Detecting cancers as early as possible is crucial and valuable for reducing tumour-associated morbidity and mortality when the malignancies are easy to be cured. Compared to traditional strategies (e.g., radiological imaging), liquid biopsy offers a non-invasive means for improving the adherence of clinicians, detection sensitivity, and cost-effectiveness of cancer diagnosis and screening [78–80]. Moreover, blood-based assays for early cancer detection are also promising to identify the tumour when it is small and before symptoms appear, which could make the disease more curable and provide a better chance for the long-term survival of patients [81]. Generally, early cancer detection can be further divided into single-cancer early detection (SCED) and multi-cancer early detection (MCED) according to the number of cancers investigated (Fig. 2 and Table 1).

Multi-omics-based SCED could effectively improve the sensitivity of cancer detection than the approaches using single modality data (such as somatic mutations) by integrating at least two types of data (Table 1). For example, Chen et al. combined four genomic features (e.g., 5-Hydroxymethylcytosine, NF, motif, and fragmentation) for early detection of HCC in cirrhotic patients, yielding an AUC > 0.99 [18]. They first randomly divided the 3,204 enrolled participants into training, validation, and test cohorts, then built the model using SVM and integrated the results with the LR approach. Qu et al. achieved 85% sensitivity and 93% specificity in discriminating (HCC) from an asymptomatic cohort with unknown HCC status by integrating cfDNA features (e.g., SNV/indel mutations and cfDNA concentration), protein marker levels (e.g., AFP and DCP), and clinical information (e.g., age and gender) [70]. They constructed the model with the LR algorithm and evaluated the model performance both in the training (65 HCC and 70 non-HCC individuals) and validation (331 individuals) cohorts. Moreover, Chalasani developed a blood-based multi-target HCC panel containing several methylated DNA (e.g., HOXA1, EMX1, TSPYL5, and B3GALT6) and protein markers (e.g., AFP and AFP-L3), their analysis based on LR model showed an AUC of 0.92 for detecting any stage of HCC [68]. These three studies integrated different types of tumour-associated signals in blood using LR algorithms for early detection of HCC, and all of them demonstrated that multi-omics strategies could outperform single-omics methods. Besides, Chabon et al. robustly detected early lung cancers by joint analysis of cfDNA mutations, background artifacts, cfDNA fragment size, the likelihood of clonal haematopoiesis, and copy-number variants (CNVs) with a multi-tiered ML approach [82]. They
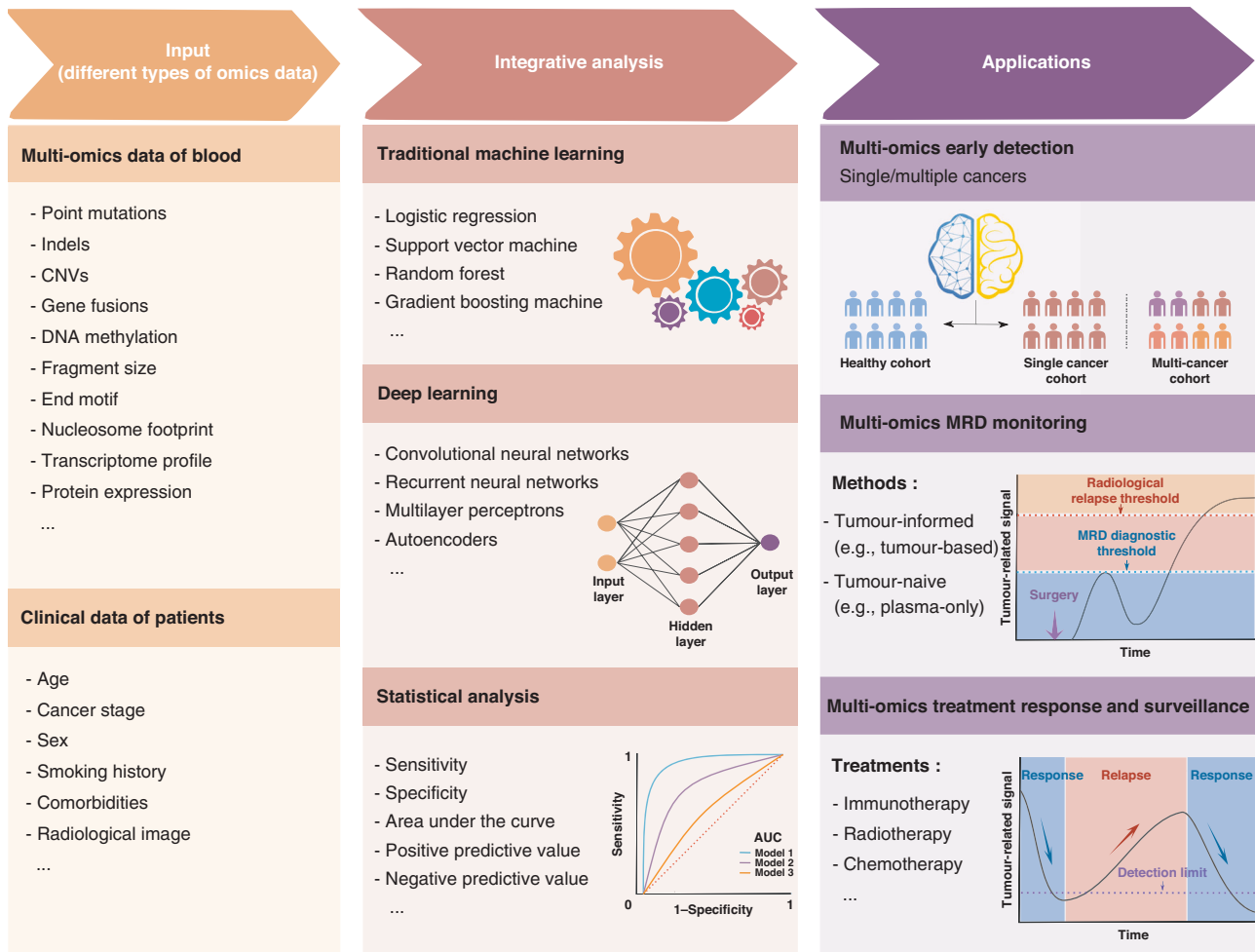
Fig. 2 Schematic plot of different applications for liquid biopsy based on multi-omics strategies. The multi-omics data and different clinical features of investigated cohorts can be integrated with bioinformatics approaches for early cancer detection, MRD monitoring, and treatment response surveillance.

performed model training in the discovery cohort (160 subjects) with a leave-one-out cross-validation framework and tested the model performance in the validation cohort (94 subjects) in terms of sensitivity, specificity, and AUC. In comparison, Mathios et al. showed that the integration of fragmentation features, CEA levels, clinical risk factors (e.g., age and smoking history), and CT imaging through an LR model with a LASSO penalty can significantly improve the non-invasive detection of lung cancer [38]. Their predictive ML model was tested with fivefold cross-validation for discriminating lung cancer patients from non-cancer individuals and the model performance was assessed with a fifth held-out fold. Uehiro et al. successfully discriminated breast cancer patients from healthy volunteers by integrating methylation markers (e.g., RASGRF1, ST3GAL6, CPXM1, SHF, JAK3, DACH1, P2RX3, DNM3, CAV2, HOXA10, B3GNT5, and chr8:23572595) and DNA concentration (e.g., CREM, GLYATL3, ELMOD3, and KLF9) with SVM model based on the testing of different variable combinations [69]. Additionally, Putcha et al. achieved high sensitivity and specificity for early detection of colorectal cancer (CRC) by combining the signals from WGS, bisulfite sequencing, and protein quantification in plasma [83], which employed a machine learning-based classifier to integrate different tumour- and immune-derived signals from epigenetic, cfDNA, and protein biomarkers. Ma et al. accomplished early detection of advanced CRC through integrative analysis of the ratio and distribution fragment sizes, end motif, breakpoint motif, as well as CNVs inferred from plasma cfDNA

WGS data with an ensemble stacked model [84]. They implemented different ML methods (e.g., RF, generalised linear model, XGBoost, gradient boosting machine, and deep learning) in their ensemble strategy and performed 10-fold cross-validation for criteria optimisation. Therefore, these studies showed that a diversity of tumour-associated signals can be combined to enhance the assay performance for different types of cancers.

Notably, because of the very low incidence of individual cancers in the general population, a large number of patients are needed for developing single cancer detection approaches. Moreover, currently reported single-cancer screening tests often values sensitivity over specificity to detect as many cancers as possible, which may lead to overdiagnosis. If different cancer types can be screened together, a higher overall positive predictive value (PPV) could be achieved due to the increase in aggregate prevalence [85, 86].

Compared with SCED, detecting multiple cancers simultaneously (MCED) may generate more value and greater impacts on public health. Furthermore, if multiple single-cancer screening tests were performed to determine the TOO, it may lead to a high cumulative false-positive rate. By contrast, MCED can theoretically minimise such influence by employing a low fixed false-positive rate and benefit the overall cancer detection of investigated population. Additionally, TOO dissection in MCED tests is valuable for reducing unnecessary diagnostic workups and the anxiety of patients, which is very useful for streamlining the diagnostic

Table 1. Multi-omics studies regarding early cancer detection.

Cancer type	Cancer stage	Cohort design	Detection methods	Target features	Algorithms	Major conclusions	Ref.
Hepatocellular carcinoma	Early stage	2250 patients with liver cirrhosis, 508 with hepatocellular carcinoma, and 476 healthy controls	5-hmc sequencing and low-pass WGS	5-hydroxymethylcytosine, nucleosome footprint, 5′ end motif, and fragmentation profiles of cfDNAs	Support vector machine and logistic regression	Their HIFI strategy achieved >95% sensitivity and >95% specificity in differentiating HCC from liver cirrhosis	[18]
Hepatocellular carcinoma	Early stage	Training cohort includes 65 patients and 70 healthy controls; independent test cohort includes 331 individuals with normal liver ultrasonography and serum AFP levels	Targeted sequencing and protein quantification	cfDNA mutation, structural variant, cfDNA concentration, protein markers (AFP and DCP), age and gender	Logistic regression	The combination of cfDNA alterations and protein markers could effectively identify early-stage HCC from asymptomatic community populations with unknown HCC status	[70]
Hepatocellular carcinoma	Early and late stage	135 patients and 302 controls	Multiplex bisulfite PCR and protein quantification	Methylated DNA markers (HOXA1, EMX1, TSPYL5, B3GALT6, etc.) and protein markers (AFP, AFP-L3, and DCP)	Logistic regression	Their multi-target panel based on the markers of methylated DNA and proteins could detect early-stage HCC with high sensitivity	[68]
Non-small cell lung cancer	Stage I–III	104 patients and 56 risk-matched controls	CAPP-Seq	cfDNA mutation, background fragment frequencies, cfDNA fragment size, likelihood of clonal haematopoiesis, copy-number alterations, etc.	Lung cancer likelihood in plasma' multi-tiered machine learning	Their integrated genomic strategy could effectively discriminate early-stage lung cancer patients from risk-matched controls using blood plasma	[82]
Lung cancer	Early and late stage	Training cohort contains 365 individuals; independent validation cohort includes 46 patients and 385 healthy controls	Low coverage WGS and protein quantification	cfDNA fragments, clinical risk factors, CEA levels and CT imaging	Logistic regression with a LASSO penalty	Their approach based on multi-omics and multi-modal data successfully detected 94% of lung cancer patients across different stages and subtypes	[38]
Colorectal cancer	Early stage, late stage, and unknown	43 patients and 548 colonoscopy-confirmed colorectal cancer negative controls	WGS, WGBS, and protein quantification	cfDNA, epigenetic, and protein biomarkers	Machine learning	The combination of tumour- and immune-derived signals from cfDNA, epigenetic, and protein biomarkers could detect early-stage colorectal cancer with high sensitivity and specificity	[83]

Table 1. continued

Cancer type	Cancer stage	Cohort design	Detection methods	Target features	Algorithms	Major conclusions	Ref.
Colorectal cancer	Early and advanced stage	Training cohort: 149 early-stage CRC patients, 46 advCRA patients and 115 healthy individuals; independent cohort: 149 early-stage CRC, 46 advCRA and 116 healthy controls	WGS	Fragment size ratio, fragment size distribution, end motif, breakpoint motif and copy number variation	Multi-dimensional ensembled stacked model	Integrative analysis of different signals inferred from WGS data held great potential for accurate noninvasive screening of colorectal cancer prior to colonoscopy	[84]
Breast cancer	Early and late stage	Training dataset includes 87 patients and 80 healthy controls; validation dataset includes 58 patients and 53 healthy controls	Droplet digital methylation-specific PCR	Methylation markers and cfDNA concentration	Support vector machine	Their model can discriminate breast cancer from normal with AUC values of 0.916 and 0.876 in the training and validation datasets, respectively	[69]
Breast, colorectal, lung, ovarian, pancreatic, gastric, or bile duct cancer	Stage I–III	236 patients with cancers and 245 healthy individuals	Low coverage WGS	cfDNA fragments and gene mutations	Gradient tree boosting machine learning	Combining fragmentation profiles with mutation-based cfDNA analyses enabled the detection of 91% of cancer patients	[32]
Bladder, breast, cholangiocarcinoma, colorectal, glioblastoma, lung, melanoma, ovarian, pancreatic, prostate, renal, etc.	Early and late stage	239 patients and 65 healthy controls	Low-pass WGS	5 fragmentation features (proportion of fragments of 160–180, 180–220, 250–320 bp, 10-bp amplitude) and t-MAD score (copy number aberration)	Logistic regression and random forest	Integration of cfDNA fragment size analysis and somatic CAN detection improved the discrimination between cancer patients and healthy individuals	[34]
Ovary, liver, stomach, pancreas, oesophagus, colorectum, lung, or breast cancers (8 cancers)	Early and late stage	1005 patients with non-metastatic clinically detected studied cancers, 812 healthy controls	Multiplex-PCR and protein quantification	16 genes mutation and 8 protein level (CA-125, CEA, CA19-9, PRL, HGF, OPN, MPO, TIMP-1)	Logistic regression and random forest	CancerSEEK detected eight cancer types with sensitivities ranging from 69 to 98 and 99% specificity based on multi-omics data	[71]
Women with unknown cancer status	Unknown cancer status	10,006 women with no personal history of cancer	Massively parallel sequencing, protein quantification, and PET-CT scan	Gene mutations, protein markers, and PET-CT scan results	Statistical methods	Multi-cancer blood testing combined with PET-CT could facilitate cancer screen and guide intervention	[16]

Note: Early-stage often refers to stage 0–1 and late-stage usually represents stages II–IV, while advanced or metastatic cancers generally mean stage IV.

HIFI 5-hydroxymethylcytosine/motif/fragmentation/nucleosome footprint, *HCC* hepatocellular carcinoma, *CAPP-Seq* cancer personalised profiling by deep sequencing, *t-MAD* trimmed median absolute deviation from copy number neutrality, *PET-CT* positron emission tomography–computed tomography.

evaluation of cancer [85, 86]. Some studies have investigated the feasibility of MCED and the localisation of TOO based on the methylation profile of ctDNA alone [17, 87]. To further improve the performance of liquid biopsy, multi-omics approaches are increasingly been applied in MCED through integrative analysis of different types of omics data (Table 1). For example, Cristiano et al. revealed that combing fragmentation profiles (e.g., fragmentation size and coverage) with mutation detection in cfDNA can increase the sensitivity of multicancer detection using low-coverage WGS (9× coverage) [32]. They applied gradient tree boosting machine learning to incorporate genome-wide fragmentation features and mutations to distinguish cancer patients from healthy individuals. Using a much lower WGS coverage (0.4×), Mouliere et al. also showed that joint analysis of fragment length and CNVs of cfDNA can boost the performance of MCED, which employed LR and random forest algorithms to identify the best predictor variables with recursive feature selection [34]. Both these two studies demonstrated that joint analysis of the variants and fragment information inferred from low-pass WGS of cfDNA hold great potential for MCED. Furthermore, Cohen et al. developed an effective blood test (called CancerSEEK) by taking into account both cfDNA mutations and the abundance of protein biomarkers (e.g., CA-125, CA19-9, CEA, HGF, Myeloperoxidase, OPN, Prolactin, and TIMP-1) for MCED and TOO localisation based on 1005 patients with nonmetastatic cancers and 812 healthy controls [71]. They performed cancer type prediction using LR and RF models, achieving a sensitivity of 69 to 98 and 99% specificity on eight different cancers. Subsequently, Lennon et al. further showed that the performance of CancerSEEK for MCED can be further enhanced by integrating three data types of genetic mutations, protein biomarkers (e.g., CEA, CA15–3, CA19–9, and CA125), and PET-CT imaging based on 10,006 participants [16]. Accordingly, multi-omics strategies are appealing to boost the performance of MCED as well as the determination of TOO.

However, systematic evaluation of MCED in terms of performance (e.g., sensitivity and specificity), analytical and clinical validity, benefit-risk, safety, and clinical utility is crucial for applying it in the healthcare system [2, 71, 81]. Moreover, substantial efforts are still needed to overcome the technical limitations, the complexity of tumour biology, and clinic features to develop highly efficient screening approaches. Specifically, the ctDNA level in early-stage disease is considerably low and the somatic mutations (e.g., CHIP-associated mutations) of noncancerous cells can also hinder the identification of real cancer-specific mutations [88]. Age is one of the most important risk factors for early cancer detection, where older individuals have a remarkably higher CHIP mutation rate and higher cancer risk than younger individuals [80, 89–91]. Furthermore, the amount of ctDNA shed into the blood varies among different cancer types and tumour stages, which could also influence the sensitivity of MCED [92]. Additionally, it is important to systematically examine the assay performance using a large-scale cohort without cancer (including the individuals with benign neoplasms and inflammatory conditions) to ensure the feasibility in the real-world asymptotic population [85]. On the other hand, since some existing standard-of-care (SOC) screening approaches are already effective for certain cancers, MCED is complementary to conventional SOC methods for increasing cancer detection rates rather than replacing them [16]. Overall, multi-omics-based strategies are very promising to improve the performance of early cancer detection and MCED may transform the landscape of cancer screening.

MRD DETECTION OF LIQUID BIOPSY
Molecular/minimal residual disease (MRD) detection is another important application in cancer management. MRD usually denotes the small number of residual tumour cells left behind after cancer treatment (e.g., surgery) that are not able to be detected with current medical imaging modalities and may eventually result in relapse or metastasis. Traditional methods for surveilling disease relapse of patients often undergo imaging scans of CT and/or PET and clinical assessment. Moreover, adjuvant treatments (e.g., chemotherapy and/or radiotherapy) are often provided to high-risk patients after surgery in many cancer types to reach a complete cure. Although such decision-making of adjuvant therapy is traditionally made based on several clinical and pathological factors (e.g., node and metastasis (TNM) staging system), none of these factors can effectively assess the MRD.

If MRD can be early detected during a disease-free follow-up period, it will open a valuable time window for identifying the patients who will ultimately relapse to improve their outcomes through tailored therapy [20]. Previous studies have revealed that blood-based MRD detection could identify the disease at a low level and provide multiple months of lead time prior to the radiographic clinical relapse [19–21, 93, 94]. Moreover, ctDNA was shown as the strongest prognostic marker compared with conventionally used risk markers based on multivariable analysis [20, 95]. Specifically, postoperative patients with detectable ctDNA can be classified as high-risk, providing valuable guidance for the following treatment decisions [96]. Therefore, MRD monitoring is appealing to monitoring disease relapse and stratifying cancer patients after treatment, which may revolutionise cancer management. It could transform approaches for patient stratification and recurrence prediction, as well as benefit precision oncology in terms of guiding de-escalation or escalation of adjuvant therapy and increasing the curable possibility while minimising unnecessary toxicities of treatment [97].

TUMOUR-INFORMED AND TUMOUR-NATIVE MRD MONITORING BASED ON MULTI-OMICS DATA
Existing approaches developed for MRD detection can be mainly grouped into tumour-informed and tumour-native/uninformed methods according to whether the prior information is used. Currently, most of the reported MRD detection assays are tumour-informed. Those methods usually identified the tumour-derived mutations specific to patients first based on tumour tissues and related controls, then track these precise genomic alterations in cfDNA [19–21]. The performance of MRD detection using single omics is still limited, especially for cancer patients with a very low tumour burden or ctDNA concentration. Same to the aforementioned early cancer detection, multi-omics strategies can also improve the performance of MRD monitoring (Table 2). For instance, Cai et al. revealed that integrative analysis of ctDNA mutations and Des-Gamma-Carboxy Prothrombin (DCP) could increase the sensitivity in MRD detection of resected patients with hepatocellular carcinoma (HCC) [98]. It could generate better prognostic value for both overall survival (OS) and relapse-free survival (RFS) than using ctDNA or DCP alone. Furthermore, Radovich et al. showed that the combination of ctDNA and CTCs is also able to improve the performance of MRD detection for early-stage Triple-Negative Breast Cancer patients after neoadjuvant chemotherapy [99]. Their method could enhance the prediction of disease recurrence and clinical outcomes of patients. Additionally, Przybyl et al. found that the MRD detection confidence in leiomyosarcoma patients could be substantially improved using a combinatorial approach of CAPP-Seq for SNV and indel identification, and genome representation profiling for CNA detection [100]. Their strategy benefited from the detection of different types of genomic alterations in ctDNA to increase the number of molecular markers for plasma tracking. Therefore, these tumour-informed studies demonstrated that multi-omics-based methods outperform the approaches using only one type of tumour signal for MRD monitoring.

G. Chen et al.

Table 2. Multi-omics-based MRD monitoring and treatment response surveillance.

MRD monitoring based on multi-omics data

Cancer type	Cohort design	Follow-up	Blood collection	Detection methods	Control	Target features	Main outcome and measures	Setting (ctDNA monitoring)	Major conclusions	Ref.
Hepatocellular carcinoma	34 patients	Long term	At preoperative, postoperative, and multiple follow-up time points	Targeted sequencing, low-coverage whole-genome sequencing, and protein quantification	Tumour-informed	ctDNA somatic mutation and Des-Gamma CarboxyProthrombin (DCP)	RFS, OS	Adjuvant	Integrative analysis of ctDNA mutations and DCP enabled a better evaluation of patients' prognostic risk and tumour occurrence detection prior to traditional strategies	[98]
Triple-negative breast cancer	196 patients	Median of 17.2 months	At the time of treatment assignment	Targeted sequencing (62–70 cancer-related genes + 7–6 genes frequently rearranged in cancer) and CTC enumeration	Tumour-informed	ctDNA somatic mutation and CTC	DDFS, DFS, and OS	Neoadjuvant	Detection of ctDNA and CTCs in patients with early-stage triple-negative breast cancer after neoadjuvant chemotherapy could be an important stratification factor	[99]
Leiomyosarcoma	7 patients and 24 healthy controls	Long term	At follow-up time points of post-treatment	CAPP-seq and GRP	Tumour-informed	Single-nucleotide variants, small indels, and copy-number alterations in ctDNA	OS	Surgery or chemotherapy	Joint analysis of SNVs, indels, and genome-wide CNAs could allow the comprehensive monitoring of tumour-specific markers in plasma for leiomyosarcoma	[100]
Colorectal cancer	103 patients	Up to 1 year	1 month after definitive therapy	Targeted sequencing (500 kb panel)	Tumour-naive	ctDNA somatic mutation and methylation	RFS	Surgery only or surgery with neoaduvant/ adjuvant therapy	Plasma-only MRD detection based on integrative analysis of epigenomic and genomic alterations could achieve comparable performance to tumour-informed approaches	[101]
Head and neck squamous cell carcinoma	30 patients and 20 risk-matched healthy controls	Minimum 2 years	At diagnosis and at least one timepoint post-treatment	CAPP-seq and cfMeDIP-seq	Tumour-naive	ctDNA somatic mutation and methylation	OS	Surgery only or surgery with adjuvant therapy	Tumour-navie detection of ctDNA through multimodal profiling of mutation and methylation could facilitate corresponding biomarker discovery and clinical use	[102]

Table 2. continued

Treatment response surveillance based on multi-omics data

Cancer type	Treatment methods	Cohort design	Follow-up	Blood collection	Detection methods	Target features	Major conclusions	Ref
Non-small cell lung cancer	Treated with immune checkpoint inhibitors (ICIs)	99 patients	Monitoring at least 6 months after treatment with PD-(L)1 blockade-based ICI	Before treatment, and after initiation of therapy with median of 18.5 days	RNA-seq, CAPP-Seq, WES and flow cytometry	ctDNA, PD-L1 expression, circulating CD8 T cell and bTMB	Integrative analysis of ctDNA and circulating immune cells in plasma could enable accurate and early forecasting of ultimate outcomes for NSCLC patients receiving immune checkpoint inhibitors	[120]
Non-small cell lung cancer	Treated with immune checkpoint blockade	24 metastatic patients and 14 early stage patients	Median duration of follow-up was 12.7 months and 16 months for metastatic and early stage patients, respectively.	At least 2 serial blood samples (range 2–8) over the course of treatment	TEC-Seq and TCR clonal expansion analysis	ctDNA and TCR dynamics	The combination of ctDNA and TCR repertoire was useful for the rapid determination of treatment outcomes for NSCLC patients treated with immune checkpoint inhibitors	[121]
Metastatic breast cancers	Treated with endocrine therapy	45 patients	Monitored at least 6 months	−28 to −1 days prior to treatment initiation, on the day of treatment initiation, 4 weeks for the first six cycles and 6 weeks thereafter	ddPCR and CTC enumeration	ctDNA ESR1$_{LBD}$m and CTC	Joint analysis of ctDNA and CTC could improve outcome prediction and mechanism identification of therapy resistance than using a single biomarker for metastatic breast cancers	[122]
Locally advanced or metastatic pancreatic cancer	With first-line treatment	100 patients	Median of 7.7 months	Before initiation of chemotherapy and after the first cycle of chemotherapy	DNA chip and cfDNA fragment size analysis	cfDNA level and cfDNA fragment	They demonstrated that cfDNA fragment size and cfDNA levels were useful for predicting the clinical outcome in patients with advanced pancreatic cancer	[119]
Metastatic colorectal cancer	Treated in first- or second-line chemotherapy	82 patients	Long term	Before the first cycle, before the second and/or third cycle of chemotherapy	Picodroplet digital PCR	ctDNA gene mutations (KRAS, BRAF, TP53) or hypermethylation (WIF1, NPY) based ctDNA concentration	Early change in ctDNA concentration inferred from tumour-specific genetic or epigenetic alterations could predict the therapeutic efficacy in patients with metastatic colorectal cancer	[117]

CAPP-seq cancer personalised profiling by deep sequencing, GRP genome representation profiling, cfMeDIP-seq cell-free methylated DNA immuno-precipitation and high-throughput sequencing, OS overall survival, RFS recurrence-free survival, DDFS distant disease-free survival, DFS disease-free survival, ICI immune checkpoint inhibitor, CTC circulating tumour cell, TEC-Seq targeted error-correction sequencing, TCR T cell receptor, ddPCR droplet digital PCR, NSCLC non-small cell lung cancer.

In contrast, tumour-naive assays aim to detect MRD using plasma only and do not rely on any prior knowledge of tumour tissues. For patients with insufficient or unavailable tumour tissues, tumour-naive strategies are the preferred choice since it is impracticable to employ a tumour-informed approach. Most of the currently reported tumour-naive assays for MRD detection were measured with a static panel based on preselected actionable/hotspot mutations or DNA methylation of specific genes in corresponding cancers. However, due to the tumour's inherent heterogeneity and a limited number of variants or DNA methylation markers for tracking, the inherent limitations of those tumour-naive methods may restrict the performance of MRD detection. Thus, joint analysis of multi-omics data provides great opportunities to improve tumour-naive MRD detection (Fig. 2, Table 2). For example, Parikh et al. demonstrated that the combination of genomic and epigenomic tumour signatures can allow effective tumour-naive MRD detection in CRC patients [101]. Their multi-omics method could increase sensitivity by 25–36% versus genomic signatures alone, producing similar sensitivity and specificity to previously reported tumour-informed approaches. Using a similar strategy, Burgener et al. showed that tumour-uninformed detection based on joint analysis of genetic mutations and methylation changes could facilitate MRD detection for head and neck squamous cell carcinoma [102]. They showed that their multi-omics approach was suitable for the scenario of low ctDNA abundance. Therefore, integrative analysis of different types of signals in the blood for tumour-naive MRD detection could also improve the performance compared to using a single modality alone.

Generally, tumour-informed and tumour-naive strategies have distinct advantages and disadvantages. The performance of MRD detection for tumour-informed methods could be enhanced with the known information on tumour-derived mutations [90, 91]. However, tumour-informed approaches may face several challenges and limitations. First, if the cancer patients are not eligible for surgery or the surgical specimen is insufficient for sequencing because of low tumour cellularity, quality, or DNA yield (e.g., neoadjuvant therapy influence), it would preclude the tumour-informed approaches [103]. Second, tumour-informed approaches may suffer from high costs and long turnaround times because of the dissection of both tumour and plasma samples. By contrast, tumour-naive methods for MRD detection may possess several advantages, such as less turnaround time due to no tumour tissue analysis, greater flexibility, and potentially lower cost. It can be conducted just based on the plasma draw of the patients no matter whether the tumour tissue is sufficient or inadequate, or non-available. Thus, tumour-naive assays provide an alternative approach for MRD detection; however, they still lack enough performance validation and breadth of supporting data [104]. Additionally, the specificity of MRD detection is largely impacted by the biological noise resulting from germline and CHIP mutations, tumour-naive approaches could be susceptible to the background noise owing to the lack of guidance information from primary tumours. Therefore, tumour-informed and tumour-naive strategies have distinct suitable scenarios for MRD detection.

Notably, the sensitivity of reported tumour-informed and tumour-naive MRD approaches remains modest, great efforts are needed to improve the assay performance. Increased sensitivity for MRD detection can further avoid inappropriate selection of high-risk patients to receive less intensive therapy or low-risk patients to take unnecessary treatment. The low ctDNA concentration is one main factor hindering MRD detection, which requires the methods should be very sensitive. Strategies for controlling sequencing artifacts could be useful, such as applying UMIs to mitigate the PCR error effect and other approaches for reducing background noise [5, 105–107]. For instance, Dai et al. recently presented a convenient and versatile strategy for

accurate mutation quantitation below 0.01% VAF by integrating variant enrichment into UMI quantitation [108].

On the other hand, serial blood testing should be conducted for patients (including both relapsing and non-relapsing patients) to establish the PPVs and negative predictive values (NPVs). Because an effective assay needs to demonstrate that the results of MRD positivity and negativity should remain reliable over time for recurrence or response monitoring. Moreover, serial blood sample analysis is an effective way to improve relapse prediction compared to the single time point analysis. For example, recent studies showed that serial ctDNA assessment for the patients with surveillance draws achieved significantly higher sensitivity of MRD monitoring compared to single ctDNA analysis in colorectal cancer [20, 95]. Serial blood testing is also crucial for proving the lead time from first MRD detection to radiographic recurrence (e.g., CT). If an MRD test shows higher sensitivity than other assays, it is expected to exhibit a longer lead time for detecting the recurrence when the tumour burden is lower and more curative. Additionally, the combination of ctDNA analysis and clinical risk assessment after treatment could further benefit patients' survival prediction and decision-making guidance [95]. Besides, DNA methylation and other epigenomic modalities, as well as DNA fragmentation patterns and motifs are also appealing for enhancing MRD detection [27, 32, 109] despite they are still not well established. Consequently, we envision that the innovation of experimental protocols and the development of multi-omics methods will improve the performance of MRD monitoring for both tumour-informed and tumour-naive tests.

TREATMENT RESPONSE AND RESISTANCE SURVEILLANCE BASED ON MULTI-OMICS DATA

In addition, the liquid biopsy also holds great promise to monitor treatment response and resistance. Specifically, treatment response surveillance is valuable to prevent ineffective therapies, avoid unnecessary side effects, and evaluate the efficacy of novel therapeutics [22, 110]. Traditional approaches for treatment response assessment often rely on serial imaging; however, radiographic measurements are unable to accurately detect the changes in tumour burden. It has been shown that serial changes of ctDNA in plasma were superior to traditional CT imaging for treatment response monitoring, having the potential to reflect treatment response before imaging [22]. For example, Dawson et al. observed that the increase of ctDNA level was closely associated with the disease progression of breast cancer, which could lead an average of 5 months to its discovery on radiographic imaging [46]. Previous studies also revealed that liquid biopsy could enable effective tracking of lethal clones during CRC treatment with EGFR-targeted therapies [111, 112]. Moreover, blood-based NGS methods have been successfully applied to track tumour evolution of metastatic cancers in response to treatment with serial plasma samples in different cancers (e.g., advanced lung, breast, and ovarian cancers) [113]. Besides, serial ctDNA analyses during and after adjuvant chemotherapy (ACT) could allow the assessment of ACT efficacy [95, 114]. Liquid biopsy was able to monitor the response of diverse chemotherapeutic drugs, including palbociclib, bevacizumab, and fulvestrant, among others [115–117]. Moreover, Qiu et al. found that postoperative ctDNA status could guide ACT, where postsurgical ctDNA positive patients benefited from ACT rather than the ctDNA negative patients [118]. Accordingly, liquid biopsy is valuable for promoting the surveillance of treatment response.

Joint analysis of multi-omics data is also an appealing strategy to improve assay performance for treatment response surveillance compared to the methods based on single modality data (Fig. 2, Table 2). For example, Lapin et al. revealed that the combination of cfDNA levels and cfDNA fragment size could enhance the

prediction of disease outcomes for advanced pancreatic cancer patients than using one type of data alone [119]. For NSCLC, Nabet et al. showed that integrative analysis of ctDNA and circulating immune cells could discriminate the patients with a durable response to immunotherapy from those without a durable response [120], while Anagnostou et al. found that both ctDNA and T cell receptor dynamics were reflective for the pathologic response to immune checkpoint blockade [121]. Using a different strategy, Paoletti suggested that integrating CTC and ctDNA biomarkers was superior to a single biomarker for patients' outcome prediction and resistance mechanism exploration for metastatic breast cancers with endocrine therapy [122]. Additionally, Garlan et al. revealed that early change in ctDNA concentration measured by quantitative analysis of gene mutation (TP53, KRAS, and BRAF) or hypermethylation (WIF1 and NPY) could effectively predict the therapeutic efficacy in metastatic colorectal cancer patients [117]. Thus, these studies showed that multi-omics-based blood biopsy can effectively improve the surveillance of treatment response and resistance to cancers. Notably, more prospective studies are essential to further systematically evaluate whether blood-based assays can accurately predict the efficacy of therapy in real time.

CHALLENGES AND FUTURE DIRECTIONS

With the innovation of approaches for precise cancer detection and monitoring, liquid biopsies will likely transform cancer management and become a routine part of clinical practice in a non-invasive and real-time way. However, many challenges remain to be resolved to further improve the performance of blood-based assays. First, the ctDNA levels of most solid tumour types in blood plasma are generally low due to limited ctDNA shedding, especially under the conditions of early-stage cancers and after definitive treatment (e.g., tumour resection) [5, 19, 47, 50]. Moreover, the half-life of ctDNA is short and the contamination, loss, or degradation of ctDNA during sample processing may lead to biased and inaccurate results and conclusions [123, 124]. Therefore, it is crucial to standardise the procedures of sample selection, collection, processing, enrichment as well as analysis. Additionally, different somatic mosaicisms, especially those cancer-unrelated CHIP mutations, could further confound the performance of blood-based assays [91, 125]. It is worth noting that CHIP mutations are age-related and often show higher frequency in older people than younger individuals without clinically apparent haematologic disease [126–129]. Accordingly, appropriate controls are essential to mitigate such effects for identifying and tracking real cancer-specific mutations.

On the other hand, most of the published studies regarding early cancer detection were mainly based on the case–control cohort rather than the population without symptoms. Although they achieved high performance, their results were not fully representative of a real-world screening population. Before clinical application, further large-scale prospective validation in an asymptomatic population will be important and necessary [85]. Moreover, preanalytical considerations that could affect liquid biopsy are worthy to be further explored, such as the type and quantity of specimen, storage conditions, sampling time points, clinical variables, as well as biological factors of patients [130]. Additionally, novel experimental strategies for reducing sequencing artifacts and noise as well as more advanced bioinformatics methods for improving joint analysis of different types of data are also in urgent need.

Besides those commonly interrogated tumour-associated signals in liquid biopsy, other cancer-related features are also worth to be systematically evaluated in the future, including TEPs, tumour exosomes, ctRNAs, metabolites, and tumour microorganisms [6]. The combination of different omics data with radiological imaging and diverse clinical risk factors could further improve the performance [6, 24]. Furthermore, the analytical validity (including sensitivity, specificity, limit of detection, PPV, NPV, and robustness) of distinct liquid biopsy approaches also need to be further evaluated and optimised. Remarkably, liquid biopsy can complement conventional approaches for better improvement of cancer management rather than fully supplant those traditional methods with proved clinical efficacy [23].

CONCLUSION

Collectively, multi-omics-based liquid biopsy through integrative analysis of different types of data is an appealing strategy to improve the performance of early cancer identification and TOO prediction, MRD monitoring, and surveillance of treatment response and resistance. The innovation and advances of liquid biopsy in both experimental protocols and bioinformatics methods will gradually transform clinical management and reduce the mortality and morbidity of various cancers.

DATA AVAILABILITY
Not applicable.

REFERENCES

1. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nat Rev Cancer. 2017;17:223–38.
2. Cescon DW, Bratman SV, Chan SM, Siu LL. Circulating tumor DNA and liquid biopsy in oncology. Nat Cancer. 2020;1:276–90.
3. Babayan A, Pantel K. Advances in liquid biopsy approaches for early detection and monitoring of cancer. Genome Med. 2018;10:4–6.
4. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. Nature. 2013;501:355–64.
5. Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature. 2017;545:446–51.
6. Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. Nat Rev Genet. 2019;20:71–88.
7. Alix-Panabières C, Pantel K. Liquid biopsy: from discovery to clinical application. Cancer Discov. 2021;11:858–73.
8. Im YR, Tsui DWY, Diaz LA, Wan JCM. Next-generation liquid biopsies: embracing data science in oncology. Trends Cancer. 2021;7:283–92.
9. Diaz LA, Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol. 2014;32:579–86.
10. Thierry AR, El Messaoudi S, Gahan PB, Anker P, Stroun M. Origins, structures, and functions of circulating DNA in oncology. Cancer Metastasis Rev. 2016;35:347–76.
11. Han X, Wang J, Sun Y. Circulating tumor DNA as biomarkers for cancer detection. Genomics Proteomics Bioinformatics. 2017;15:59–72.
12. Fu Y, Yang Z, Hu Z, Yang Z, Pan Y, Chen J, et al. Preoperative serum ctDNA predicts early hepatocellular carcinoma recurrence and response to systemic therapies. Hepatol Int. 2022;16:868–78.
13. Baselga J, Im SA, Iwata H, Cortés J, De Laurentiis M, Jiang Z, et al. Buparlisib plus fulvestrant versus placebo plus fulvestrant in postmenopausal, hormone receptor-positive, HER2-negative, advanced breast cancer (BELLE-2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol. 2017;18:904–16.
14. Jia J, Morse MA, Nagy RJ, Lanman RB, Strickler JH. Cell-free DNA profiling to discover mechanisms of exceptional response to cabozantinib plus panitumumab in a patient with treatment refractory metastatic colorectal cancer. Front Oncol. 2018;8:305.
15. 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.
16. Lennon AM, Buchanan AH, Kinde I, Warren A, Honushefsky A, Cohain AT, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science. 2020;369:eabb9601.
17. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Liu MC, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol. 2020;31:745–59.

18. Chen L, Abou-Alfa GK, Zheng B, Liu JF, Bai J, Du LT, et al. Genome-scale profiling of circulating cell-free DNA signatures for early detection of hepatocellular carcinoma in cirrhotic patients. Cell Res. 2021;31:589–92.

19. Chaudhuri AA, Chabon JJ, Lovejoy AF, Newman AM, Stehr H, Azad TD, et al. Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling. Cancer Discov. 2017;7:1394–403.

20. Reinert T, Henriksen TV, Christensen E, Sharma S, Salari R, Sethi H, et al. Analysis of plasma cell-free DNA by ultradeep sequencing in patients with stages I to III colorectal cancer. JAMA Oncol. 2019;5:1124–31.

21. Chen G, Peng J, Xiao Q, Wu HX, Wu X, Wang F, et al. Postoperative circulating tumor DNA as markers of recurrence risk in stages II to III colorectal cancer. J Hematol Oncol. 2021;14:80.

22. Kilgour E, Rothwell DG, Brady G, Dive C. Liquid biopsy-based biomarkers of treatment response and resistance. Cancer Cell. 2020;37:485–95.

23. De Rubis G, Rajeev Krishnan S, Bebawy M. Liquid biopsies in cancer diagnosis, monitoring, and prognosis. Trends Pharmacol Sci. 2019;40:172–86.

24. Ignatiadis M, Sledge GW, Jeffrey, SS. Liquid biopsy enters the clinic—implementation issues and future challenges. Nat Rev Clin Oncol. 2021;18:297–31.

25. Snyder MW, Kircher M, Hill AJ, Daza RM, Correspondence JS. Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. Cell. 2016;164:57–68.

26. Costello JF, Frühwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, et al. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nat Genet. 2000;24:132–8.

27. Shen SY, Singhania R, Fehringer G, Chakravarthy A, Roehrl MHA, Chadwick D, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. Nature. 2018;563:579–83.

28. Garrigou S, Perkins G, Garlan F, Normand C, Didelot A, Le Corre D, et al. A study of hypermethylated circulating tumor DNA as a universal colorectal cancer biomarker. Clin Chem. 2016;62:1129–39.

29. Pietrasz D, Wang-Renault S, Taieb J, Dahan L, Postel M, Durand-Labrunie J, et al. Prognostic value of circulating tumour DNA in metastatic pancreatic cancer patients: post-hoc analyses of two clinical trials. Br J Cancer. 2021;126:440–8.

30. Beinse G, Borghese B, Métairie M, Just P-A, Poulet G, Garinet S, et al. Highly specific droplet-digital PCR detection of universally methylated circulating tumor DNA in endometrial carcinoma. Clin Chem. 2022. https://doi.org/10.1093/CLINCHEM/HVAC020.

31. Tan HT, Low J, Lim SG, Chung MCM. Serum autoantibodies as biomarkers for early cancer detection. FEBS J. 2009;276:6880–904.

32. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. Nature. 2019;570:385–9.

33. Heider K, Wan JCM, Hall J, Belic J, Boyle S, Hudecova I, et al. Detection of ctDNA from dried blood spots after DNA size selection. Clin Chem. 2020;66:697–705.

34. Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Barlebo Ahlborn L, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci Transl Med. 2018;10:4921.

35. Jiang P, Sun K, Peng W, Cheng SH, Ni M, Yeung PC, et al. Plasma DNA end-motif profiling as a fragmentomic marker in cancer, pregnancy, and transplantation. Cancer Discov. 2020;10:664–73.

36. Larson MH, Pan W, Kim HJ, Mauntz RE, Stuart SM, Pimentel M, et al. A comprehensive characterization of the cell-free transcriptome reveals tissue- and subtype-specific biomarkers for cancer detection. Nat Commun. 2021;12:2357.

37. Poore GD, Kopylova E, Zhu Q, Carpenter C, Fraraccio S, Wandro S, et al. Microbiome analyses of blood and tissues suggest cancer diagnostic approach. Nature. 2020;579:567–74.

38. Mathios D, Johansen JS, Cristiano S, Medina JE, Phallen J, Larsen KR, et al. Detection and characterization of lung cancer using cell-free DNA fragmentomes. Nat Commun. 2021;12:5060.

39. Li Y, Ma L, Wu D, Chen G. Advances in bulk and single-cell multi-omics approaches for systems biology and precision medicine. Brief Bioinform. 2021;22:1–18.

40. Amelio I, Bertolo R, Bove P, Buonomo OC, Candi E, Chiocchi M, et al. Liquid biopsies and cancer omics. Cell Death Discov. 2020;6:131.

41. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol. 2019;20:590–607.

42. Zaporozhchenko IA, Ponomaryova AA, Rykova EY, Laktionov PP. The potential of circulating cell-free RNA as a cancer biomarker: challenges and opportunities. Expert Rev Mol Diagn. 2018;18:133–45.

43. Geary B, Walker MJ, Snow JT, Lee DCH, Pernemalm M, Maleki-Dizaji S, et al. Identification of a biomarker panel for early detection of lung cancer patients. J Proteome Res. 2019;18:3369–82.

44. Peng L, Cantor DI, Huang C, Wang K, Baker MS, Nice EC. Tissue and plasma proteomics for early stage cancer detection. Mol Omics. 2018;14:405–23.

45. Chen F, Dai X, Zhou C-C, Li K, Zhang Y, Lou X-Y, et al. Integrated analysis of the faecal metagenome and serum metabolome reveals the role of gut microbiome-associated metabolites in the detection of colorectal cancer and adenoma. Gut. 2021. https://doi.org/10.1136/gutjnl-2020-323476.

46. Dawson S-J, Tsui DWY, Murtaza M, Biggs H, Rueda OM, Chin S-F, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med. 2013;368:1199–209.

47. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. Sci Transl Med. 2015;7:1–12.

48. Diehl F, Li M, Dressman D, He Y, Shen D, Szabo S, et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. Proc Natl Acad Sci USA. 2005;102:16368–73.

49. Newman AM, Bratman SV, To J, Wynne JF, Eclov NCW, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med. 2014;20:548–54.

50. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, et al. Direct detection of early-stage cancers using circulating tumor DNA. Sci Transl Med. 2017;9:eaan2415.

51. Razavi P, Li BT, Brown DN, Jung B, Hubbell E, Shen R, et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. Nat Med. 2019;25:1928–37.

52. Lianidou E. Detection and relevance of epigenetic markers on ctDNA: recent advances and future outlook. Mol Oncol. 2021;15:1683–1700.

53. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. Lancet. 2018;392:777–86.

54. Venegas V, Halberg MC. Quantification of mtDNA mutation heteroplasmy (ARMS qPCR). Methods Mol Biol. 2012;837:313–26.

55. Thierry AR. A targeted Q-PCR-based method for point mutation testing by analyzing circulating DNA for cancer management care. Methods Mol Biol. 2016;1392:1–16.

56. Milbury CA, Li J, Liu P, Makrigiorgos GM. COLD-PCR: improving the sensitivity of molecular diagnostics assays. Expert Rev Mol Diagn. 2011;11:159–69.

57. Perkins G, Lu H, Garlan F, Taly V. Droplet-based digital PCR: application in cancer research. Adv Clin Chem. 2017;79:43–91.

58. Diehl F, Li M, He Y, Kinzler KW, Vogelstein B, Dressman D. BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions. Nat Methods. 2006;3:551–9.

59. Tran NH, Kisiel J, Roberts LR. Using cell-free DNA for HCC surveillance and prognosis. JHEP Rep. 2021;3:100304.

60. Masfarré L, Vidal J, Fernández-Rodríguez C, Montagut C. ctDNA to guide adjuvant therapy in localized colorectal cancer (CRC). Cancers. 2021;13:2869.

61. Takemasa I, Hamabe A, Ishii M. Perspectives for circulating tumor DNA in clinical management of colorectal cancer. Int J Clin Oncol. 2021;26:1420–30.

62. Chen M, Zhao H. Next-generation sequencing in liquid biopsy: cancer screening and early detection. Hum Genomics. 2019;13:34.

63. Oxnard GR, Klein EA, Seiden MV, Hubbell E, Venn O, Jamshidi A, et al. Simultaneous multi-cancer detection and tissue of origin (TOO) localization using targeted bisulfite sequencing of plasma cell-free DNA (cfDNA). Ann Oncol. 2019;30:v912.

64. Liu MC, Klein E, Hubbell E, Maddala T, Aravanis AM, Beausang JF, et al. Plasma cell-free DNA (cfDNA) assays for early multi-cancer detection: the circulating cell-free genome atlas (CCGA) study. Ann Oncol. 2018;29:viii14–viii57.

65. Liu L, Chen X, Petinrin OO, Zhang W, Rahaman S, Tang ZR, et al. Machine learning protocols in early cancer detection based on liquid biopsy: a survey. Life. 2021;11:1–39.

66. Nalepa J, Kawulok M. Selecting training sets for support vector machines: a review. Artif Intell Rev. 2019;52:857–900.

67. Leo B. Random forests. Mach Learn. 2001;45:5–32.

68. Chalasani NP, Ramasubramanian TS, Bhattacharya A, Olson MC, Edwards V DK, Roberts LR, et al. A novel blood-based panel of methylated DNA and protein markers for detection of early-stage hepatocellular carcinoma. Clin Gastroenterol Hepatol. 2021;19:2597.e4–605.e4.

69. Uehiro N, Sato F, Pu F, Tanaka S, Kawashima M, Kawaguchi K, et al. Circulating cell-free DNA-based epigenetic assay can detect early breast cancer. Breast Cancer Res. 2016;18:129.

70. Qu C, Wang Y, Wang P, Chen K, Wang M, Zeng H, et al. Detection of early-stage hepatocellular carcinoma in asymptomatic HBsAg-seropositive individuals by liquid biopsy. Proc Natl Acad Sci USA 2019;116:6308–12.

71. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. 2018;17:20.

72. Zhong G, Wang LN, Ling X, Dong J. An overview on data representation learning: from traditional feature learning to recent deep learning. J Financ Data Sci. 2016;2:265–78.

73. Dong X, Yu Z, Cao W, Shi Y, Ma Q. A survey on ensemble learning. Front Comput Sci. 2020;14:241–58.

74. Bello M, Nápoles G, Sánchez R, Bello R, Vanhoof K. Deep neural network to extract high-level features and labels in multi-label classification problems. Neurocomputing. 2020;413:259–70.

75. Le QV. Building high-level features using large scale unsupervised learning. ICASSP, IEEE Int Conf Acoust Speech Signal Process Proc. 2013. https://doi.org/10.1109/ICASSP.2013.6639343.

76. Lones MA. How to avoid machine learning pitfalls: a guide for academic researchers. arXiv:2108.02497v2 [Preprint]. 2021 [cited 2021 Aug 5]: [19 p.]. Available from https://arxiv.org/abs/2108.02497

77. Kann BH, Hosny A, Aerts HJWL. Artificial intelligence for clinical oncology. Cancer Cell. 2021;39:916–27.

78. Aravanis AM, Lee M, Klausner RD. Next-generation sequencing of circulating tumor DNA for early cancer detection. Cell. 2017;168:571–4.

79. Campos-Carrillo A, Weitzel JN, Sahoo P, Rockne R, Mokhnatkin JV, Murtaza M, et al. Circulating tumor DNA as an early cancer detection tool. Pharmacol Ther. 2020;207:107458.

80. Clarke CA, Hubbell E, Ofman JJ. Multi-cancer early detection: a new paradigm for reducing cancer-specific and all-cause mortality. Cancer Cell. 2021;39:447–8.

81. Ahlquist DA. Universal cancer screening: revolutionary, rational, and realizable. npj Precis Oncol. 2018;2:1–5.

82. Chabon JJ, Hamilton EG, Kurtz DM, Esfahani MS, Moding EJ, Stehr H, et al. Integrating genomic features for non-invasive early lung cancer detection. Nature. 2020;580:245–51.

83. Putcha G, Liu T-Y, Ariazi E, Bertin M, Drake A, Dzamba M, et al. Blood-based detection of early-stage colorectal cancer using multiomics and machine learning. J Clin Oncol. 2020;38:66.

84. Hematol J, Ma X, Chen Y, Tang W, Bao H, Mo S, et al. Multi-dimensional fragmentomic assay for ultrasensitive early detection of colorectal advanced adenoma and adenocarcinoma. J Hematol Oncol. 2021. https://doi.org/10.1186/s13045-021-01189-w.

85. Braunstein GD, Ofman JJ. Criteria for evaluating multi-cancer early detection tests. touchREVIEWS Oncol Haematol. 2021;17:3–6.

86. Putcha G, Gutierrez A, Skates, S. Multicancer screening: one size does not fit all. JCO Precis Oncol. 2021;5:574–6.

87. Liu L, Toung JM, Jassowicz AF, Vijayaraghavan R, Kang H, Zhang R, et al. Targeted methylation sequencing of plasma cell-free DNA for cancer detection and classification. Ann Oncol. 2018;29:1445–53.

88. Adashek JJ, Janku F, Kurzrock R. Signed in blood: circulating tumor DNA in cancer diagnosis, treatment and screening. Cancers. 2021;13:3600.

89. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. Ca Cancer J Clin. 2020;70:7–30.

90. Abbosh C, Swanton C, Birkbak NJ. Clonal haematopoiesis: a source of biological noise in cell-free DNA analyses. Ann Oncol. 2019;30:358–9.

91. Hu Y, Ulrich BC, Supplee J, Kuang Y, Lizotte PH, Feeney NB, et al. False-positive plasma genotyping due to clonal hematopoiesis. Clin Cancer Res. 2018;24:4437–43.

92. Bettegowda 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.

93. Chen K, Zhao H, Shi Y, Yang F, Wang LT, Kang G, et al. Perioperative dynamic changes in circulating tumor DNA in patients with lung cancer (Dynamic). Clin Cancer Res. 2019;25:7058–67.

94. Xia L, Mei J, Kang R, Deng S, Chen Y, Yang Y, et al. Perioperative ctDNA-based molecular residual disease detection for non-small cell lung cancer: a prospective multicenter cohort study (LUNGCA-1). Clin Cancer Res. 2021. https://doi.org/10.1158/1078-0432.CCR-21-3044.

95. Henriksen TV, Tarazona N, Frydendahl A, Reinert T, Gimeno-Valiente F, Carbonell-Asins JA, et al. Circulating tumor DNA in stage III colorectal cancer, beyond minimal residual disease detection, towards assessment of adjuvant therapy efficacy and clinical behavior of recurrences. Clin Cancer Res. 2021. https://doi.org/10.1158/1078-0432.ccr-21-2404.

96. Chin RI, Chen K, Usmani A, Chua C, Harris PK, Binkley MS, et al. Detection of solid tumor molecular residual disease (MRD) using circulating tumor DNA (ctDNA). Mol Diagn Ther. 2019. https://doi.org/10.1007/s40291-019-00390-5.

97. Pantel K, Alix-Panabières C. Liquid biopsy and minimal residual disease — latest advances and implications for cure. Nat Rev Clin Oncol. 2019;16:409–24.

98. Cai Z, Chen G, Zeng Y, Dong X, Li Z, Huang Y, et al. Comprehensive liquid profiling of circulating tumor DNA and protein biomarkers in long-term follow-up patients with hepatocellular carcinoma. Clin Cancer Res. 2019;25:5284–94.

99. Radovich M, Jiang G, Hancock, BA, Chitambar C, Nanda R, Falkson C, et al. Association of circulating tumor DNA and circulating tumor cells after neoadjuvant chemotherapy with disease recurrence in patients with triple-negative breast cancer: preplanned secondary analysis of the BRE12-158 randomized clinical Trial. JAMA Oncol. 2020;6:1410–5.

100. Przybyl J, Chabon JJ, Spans L, Ganjoo KN, Vennam S, Varma S, et al. Combination approach for detecting different types of alterations in circulating tumor DNA in leiomyosarcoma. Clin Cancer Res. 2018;24:2688–99.

101. Parikh AR, Van Seventer EE, Siravegna G, Hartwig AV, Jaimovich A, He Y, et al. Minimal residual disease detection using a plasma-only circulating tumor DNA assay in patients with colorectal cancer. Clin Cancer Res. 2021. https://doi.org/10.1158/1078-0432.ccr-21-0410.

102. Burgener JM, Zou J, Zhao Z, Zheng Y, Shen SY, Huang SH, et al. Tumor-naïve multimodal profiling of circulating tumor DNA in head and neck squamous cell carcinoma. Clin Cancer Res. 2021;27:4230–44.

103. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med. 2017;23:703–13.

104. Signatera T. A comparison of tumor-informed and tumor-naive approaches for early-stage molecular residual disease (MRD) detection two different approaches: to personalize or not to personalize. 2021. https://www.natera.com/wp-content/uploads/2021/05/SGN_WP_Solar_20210503_NAT-9000052_FINAL_DWNLD.pdf.

105. Abbosh C, Birkbak NJ, Swanton C. Early stage NSCLC — challenges to implementing ctDNA-based screening and MRD detection. Nat Rev Clin Oncol. 2018;15:577–86.

106. Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, Scherer F, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. Nat Biotechnol. 2016;34:547–55.

107. Pécuchet N, Rozenholc Y, Zonta E, Pietraz D, Didelot A, Combe P, et al. Analysis of base-position error rate of next-generation sequencing to detect tumor mutations in circulating DNA. Clin Chem. 2016;62:1492–503.

108. Dai P, Wu LR, Chen SX, Wang MX, Cheng LY, Zhang JX, et al. Calibration-free NGS quantitation of mutations below 0.01% VAF. Nat Commun. 2021;12:1–9.

109. Xie H, Mahoney DW, Foote PH, Burger K, Doering KA, Taylor WR, et al. Novel methylated DNA markers in plasma detect distant recurrence of colorectal cancer. J Clin Oncol. 2020;38:4088.

110. Shoukry M, Broccard S, Kaplan J, Gabriel E. The emerging role of circulating tumor DNA in the management of breast cancer. Cancers. 2021;13:3813.

111. Parseghian CM, Loree JM, Morris VK, Liu X, Clifton KK, Napolitano S, et al. Anti-EGFR-resistant clones decay exponentially after progression: implications for anti-EGFR re-challenge. Ann Oncol. 2019;30:243–9.

112. Siravegna G, Mussolin B, Buscarino M, Corti G, Cassingena A, Crisafulli G, et al. Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. Nat Med. 2015;21:795–801.

113. Murtaza M, Dawson SJ, Tsui DWY, Gale D, Forshew T, Piskorz AM, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature. 2013;497:108–12.

114. Benhaim L, Bouché O, Normand C, Didelot A, Mulot C, Le Corre D, et al. Circulating tumor DNA is a prognostic marker of tumor recurrence in stage II and III colorectal cancer: multicentric, prospective cohort study (ALGECOLS). Eur J Cancer. 2021;159:24–33.

115. Darrigues L, Pierga JY, Bernard-Tessier A, Bièche I, Silveira AB, Michel M, et al. Circulating tumor DNA as a dynamic biomarker of response to palbociclib and fulvestrant in metastatic breast cancer patients. Breast Cancer Res. 2021;23:1–10.

116. Nakagomi H, Hirotsu Y, Amemiya K, Nakada H, Inoue M, Mochizuki H, et al. Rapid changes in circulating tumor DNA in serially sampled plasma during treatment of breast cancer: a case report. Am J Case Rep. 2017;18:26–32.

117. Garlan F, Laurent-Puig P, Sefrioui D, Siauve N, Didelot A, Sarafan-Vasseur N, et al. Early evaluation of circulating tumor DNA as marker of therapeutic efficacy in metastatic colorectal cancer patients (PLACOL Study). Clin Cancer Res. 2017;23:5416–25.

118. Qiu B, Guo W, Zhang F, Lv F, Ji Y, Peng Y, et al. Dynamic recurrence risk and adjuvant chemotherapy benefit prediction by ctDNA in resected NSCLC. Nat Commun. 2021;12:6770.

119. Lapin M, Oltedal S, Tjensvoll K, Buhl T, Smaaland R, Garresori H, et al. Fragment size and level of cell-free DNA provide prognostic information in patients with advanced pancreatic cancer. J Transl Med. 2018;16:300.

120. Nabet BY, Esfahani MS, Moding EJ, Hamilton EG, Chabon JJ, Rizvi H, et al. Noninvasive early identification of therapeutic benefit from immune checkpoint inhibition. Cell. 2020;183:363.e13–76.e13.

121. Anagnostou V, Forde PM, White JR, Niknafs N, Hruban C, Naidoo J, et al. Dynamics of tumor and immune responses during immune checkpoint blockade in Non-small cell lung cancer. Cancer Res. 2019;79:1214–25.

122. Paoletti C, Schiavon G, Dolce EM, Darga EP, Hedley Carr T, Geradts J, et al. Circulating biomarkers and resistance to endocrine therapy in metastatic breast cancers: correlative results from AZD9496 Oral SERD phase I trial. Clin Cancer Res. 2018;24:5860–72.

123. Pessoa LS, Heringer M, Ferrer VP. ctDNA as a cancer biomarker: a broad overview. Crit Rev Oncol Hematol. 2020;155:103109.

124. Zhu G, Guo YA, Ho D, Poon P, Poh ZW, Wong PM, et al. Tissue-specific cell-free DNA degradation quantifies circulating tumor DNA burden. Nat Commun. 2021;12:2229.
125. Swanton C, Frs F, Venn O, Aravanis A, Hubbell E, Maddala T, et al. Prevalence of clonal hematopoiesis of indeterminate potential (CHIP) measured by an ultra-sensitive sequencing assay: exploratory analysis of the Circulating Cell-free Genome Atlas (CCGA) study. J Clin Oncol. 2018;36:12003.
126. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. Nat Commun. 2016;7:12484.
127. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N. Engl J Med. 2014;371:2488–98.
128. Zink F, Stacey SN, Norddahl GL, Frigge ML, Magnusson OT, Jonsdottir I, et al. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. Blood. 2017;130:742–52.
129. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood. 2015;126:9–16.
130. Merker JD, Oxnard GR, Compton C, Diehn M, Hurley P, Lazar AJ, et al. Circulating tumor DNA analysis in patients with cancer: American Society of Clinical Oncology and College of American Pathologists joint review. Arch Pathol Lab Med. 2018;142:1242–53.

AUTHOR CONTRIBUTIONS
GC, JZ, QF, VT, and FT wrote the manuscript. GC and FT revised and finalised the manuscript. All authors read and approved the final manuscript.

FUNDING

COMPETING INTERESTS
The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE
Not applicable.

CONSENT FOR PUBLICATION
Not applicable.

ADDITIONAL INFORMATION
Correspondence and requests for materials should be addressed to Geng Chen, Valerie Taly or Fei Tan.

Reprints and permission information is available at http://www.nature.com/reprints

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