SCIENCE CHINA
Chemistry

•REVIEWS•
SPECIAL ISSUE: Celebrating the 130th Anniversary of Wuhan University

November 2023 Vol.66 No.11: 3006–3023
https://doi.org/10.1007/s11426-023-1727-9

Circulating nucleic acids as liquid biopsies for disease prediction, screening and diagnosis

Yuqing Sun[†], Yongling Tang[†], Long Yu, Yuhao Wang, Shuang Peng[*] & Xiang Zhou[*]

College of Chemistry and Molecular Sciences, Key Laboratory of Biomedical Polymers-Ministry of Education, Department of Hematology of Zhongnan Hospital, Taikang Center for Life and Medical Sciences, Wuhan University, Wuhan 430072, China

Received June 3, 2023; accepted August 1, 2023; published online October 24, 2023

Liquid biopsy used molecular information in body liquid to perform early diagnosis, screening, monitoring, prognosis, and treatment of various diseases. Circulating free nucleic acids (cfNA) are important diagnostic biomarkers, providing a window to accurately and immediately observe the body's vital activity status. With the development of gene sequencing technology and bioinformatics technology, genetic, epigenetic, and fragtomics alterations that can be detected in cfDNA, as well as the expression level of miRNA and cf-mRNA can be quantified, this can reflect its tissue origin, gene regulation, genome evolution, and disease pathogenesis. This review focuses on the clinical utility of cfNA in different body liquids (blood, urine, bile), and discusses the diagnostic efficacy and accuracy of cfNA as diagnostic biomarkers in a variety of diseases. Blood is widely used to diagnose various tissue lesions for liquid biopsies as a body fluid circulating throughout the body, reflecting the state of the entire body. Bile and urine, as local circulating body fluids, can better reflect the changing state of tissues around the biliary tract and tissues around the bladder, respectively. In addition, normalized sample preservation, cfNA extraction, and detection procedures will help the practical application of cfNA in the clinic.

cliec acids, liquid biopsy, biomarkers

Citation: Sun Y, Tang Y, Yu L, Wang Y, Peng S, Zhou X. Circulating nucleic acids as liquid biopsies for disease prediction, screening and diagnosis. *Sci China Chem*, 2023, 66: 3006–3023, https://doi.org/10.1007/s11426-023-1727-9

1 Introduction

Minimally invasive liquid biopsy is attracting attention because precision diagnosis seeks to improve patient outcomes by using molecular information about the disease to perform prediction, screening diagnosis, and prognosis of disease [1]. Multidimensional molecules such as circulating tumor cells (CTCs), extracellular vesicles (EVs), circulating free nucleic acids (cfNA), proteins, and metabolites in biological fluids such as blood, urine, bile, and saliva are important targets molecules for liquid biopsy. Novel diagnostic biomarkers were developed by detecting the distribution and content levels of the above multidimensional molecules, screening differential molecules in disease and normal control groups, building diagnostic models through machine learning, and verifying the diagnostic performance of these biomarkers in large sample [2].

Compared with traditional pathological diagnosis methods such as needle biopsy, liquid biopsy technology is less invasive, biological fluids are easier to obtain, and a variety of new detection technologies can find disease-specific molecules at a low level, which is helpful to clinically monitoring the changes of patient's conditions and the effect of treatment, so it has broad research prospects in disease diagnosis [3]. Body fluids, including blood, urine, bile, cerebrospinal fluid, saliva, and other fluids, are the carrier of body metabolism, and material exchange, and carry the code of life

†These authors contributed equally to this work.
*Corresponding authors (email: pengshuang@whu.edu.cn; xzhou@whu.edu.cn)

© Science China Press 2023

activities. Blood is widely used for liquid biopsies as a body fluid circulating throughout the body, reflecting the state of the entire body, and with the advantage of being more accessible. Bile and urine, as local circulating body fluids, are more likely to contact some tissues and organs, and can better reflect the changing state of specific tissues and organs. Liquid biopsy facilitates the design of personalized treatments without the need for invasive tissue biopsy sampling, which is a core goal of disease treatment.

Circulating cell-free nucleic acids (cfNA) include circulating cell-free DNA (cfDNA) and circulating cell-free RNA (cfRNA). The release of nucleic acids into body fluids is thought to be associated with apoptosis and necrosis of cancer cells in the tumor microenvironment, as well as from active secretion by cells [4]. Tumor cells circulating in the blood, as well as micro-metastatic deposits in the bone marrow and liver, also contribute to the release of cfNA. Together, these cfNA have the potential to provide information about features of primary tumours or metastases. High-throughput sequencing yields a large amount of data with cfNA, and bioinformatics techniques are used to analyze this data, among which mathematical algorithmic deconvolution techniques can obtain the tissue origin of cfNA. Tissue deconvolution is the process of extracting cell type-specific information from heterogeneous samples. In liquid biopsies, this might involve resolving plasma DNA fragments into their constituent elements (for example, determining their tissue of origin on the basis of methylation markers). In another review [1], Speicher *et al*. showed the contributions of different tissues to plasma nucleic acids on the basis of data from four studies, which all confirmed that white blood cells are the main contributors to the plasma cell-free DNA pool, followed by placenta (in pregnant females) and liver. However, the relative contribution from other tissues differed substantially between studies.

On average, the size of cfDNA in blood ranges from small fragments of 70 to 200 base pairs and large fragments of about 21 kilobases [5]. Overall cfDNA concentrations vary widely, with cancer patients generally having higher levels of cfDNA than healthy control donors. The range of 0 to >1,000 ng/mL of blood has been measured in cancer patients, with an average of 180 ng/mL cfDNA. In contrast, healthy subjects had concentrations of cfDNA in their blood ranging from 0 to 100 ng/mL, with an average of 30 ng/mL cfDNA [4]. The half-life of naked RNA in the circulation is approximately 15 s [6]. In contrast, the half-life of endogenous cfNAs significantly increased to several minutes to hours [7] due to their association with proteins, lipoproteins and shielding by extracellular vesicles [8].

Circulating free nucleic acids (cfNA) are important diagnostic biomarkers, providing a window to observe the body's vital activity status, which can accurately and immediately feedback on the change of the body's state. Cancer-associated point mutations [9], gene expression variations [10,11], methylation [12], and fragmentation patterns [13] of cfDNA in blood have been identified as potential diagnostic or prognostic markers [14]. Historically, cfRNA research mainly focuses on microRNA in blood [15], cf-mRNA research is in its infancy, and more exploration is needed.

This review focuses on the clinical utility of cfNA in different body liquids (blood, urine, bile), including genetic and epigenetic alterations that can be detected in cfDNA, as well as the quantification of the expression level of miRNA and cf-mRNA, and discusses the diagnostic efficacy and accuracy of cfNA as diagnostic biomarkers in a variety of diseases (Figure 1).

2 Circulating nucleic acids as biomarkers in blood for liquid biopsy

2.1 cfDNA in plasma

Circulating cell-free DNA (cfDNA) mutations have been extensively researched and developed as a potential biomarker for various diseases. In the past decade, there has been reported a high degree of consistency between cfDNA mutations and tumor. This has drawn more attention from researchers to individual-specific and precise treatment. cfDNA contains fragments of DNA that are released into the bloodstream by tumor cells, known as circulating tumor DNA (ctDNA). These ctDNA fragments contain mutation information that is specific to certain tumors. Zhang and coworkers [16] used parallel sequencing between plasma and leukocytes in 10,000 Chinese pan-cancer patients and found that 14% of plasma cell-free DNA samples contained clonal hematopoietic (CH) variants whose detection rate increased with age. After the elimination of CH variants, ctDNA was detected in 73.5% of plasma samples, with the highest detection rates in small-cell lung cancer (91.1%) and prostate cancer (87.9%). Early detection of cancer through cfDNA testing can lead to earlier interventions and better patient outcomes. Chen and coworkers [17] showed that cancer could be detected noninvasively up to 4 years before the current standard of care. It tested plasma samples from 605 asymptomatic individuals, 191 of whom were diagnosed with gastric, esophageal, colorectal, lung, or liver cancer within 4 years after blood collection. The flowchart of ruitment, baseline survey, sample collection, and cohort follow-up for TZL was shown in Figure 2. cfDNA analysis can be used as an adjunct to therapy by identifying tumor-specific mutations. McDonald and coworkers [18] developed a method for simultaneously multiplexed analysis of patient-specific cancer mutations. Mutant allele fraction levels in cfDNA of breast cancer patients were significantly reduced after treatment. ctDNA assays molecular levels and residual lesions during adjuvant therapy. Bratman and coworkers

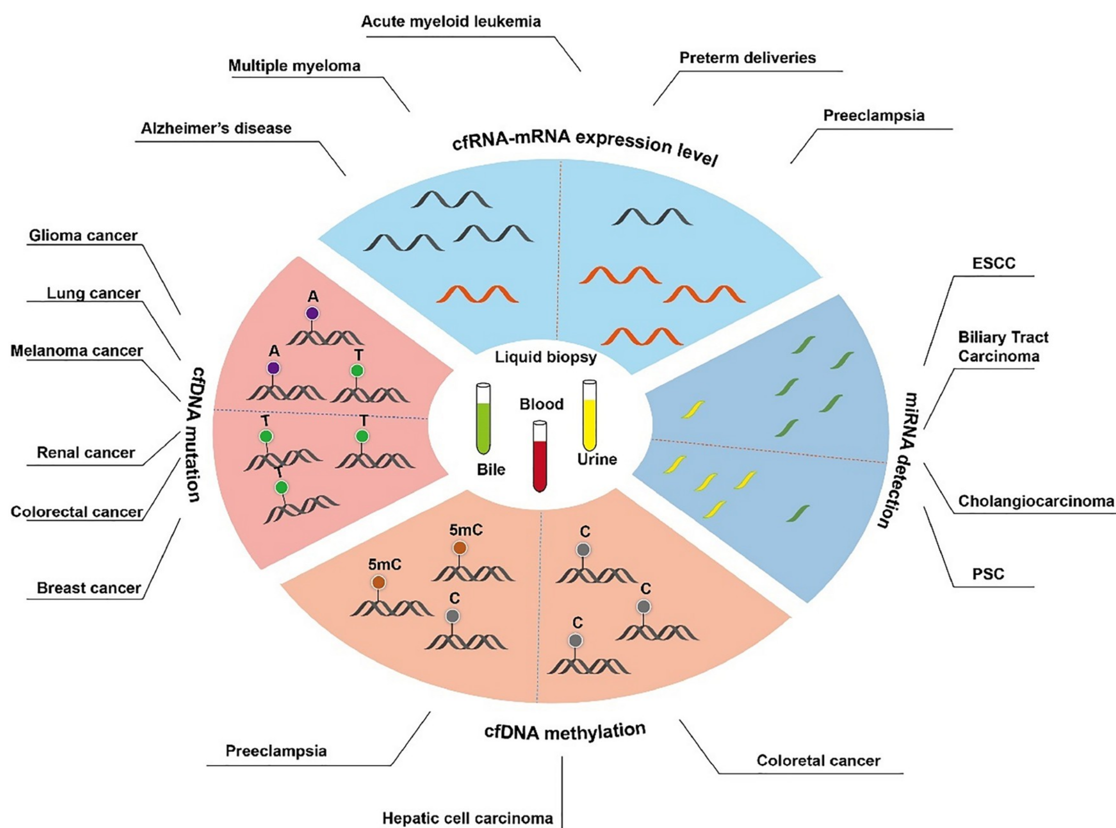


Figure 1 Clinical utility of cfNA including cfDNA and cfRNA in different body liquids (blood, urine, bile) (color online).

evaluated ctDNA in five different cohorts of patients with advanced solid tumors who were treated with monoclonal antibodies. The results showed that the baseline ctDNA concentration was associated with progression-free survival, overall survival, clinical response, and clinical benefit. This study demonstrates that ctDNA can be used clinically to monitor the course of disease in patients undergoing therapy [19]. Postoperative monitoring of small lesions can effectively predict the recurrence of the disease. Leal and coworkers [20] developed a phase 3 randomized controlled study of the perioperative treatment of patients with operable gastric cancer. After filtering for changes in matched leukocytes, the presence of ctDNA was predictive of relapse in analyses performed within 9 weeks of preoperative therapy and 9 weeks after surgery in patients eligible for multimodal therapy. Yeh and coworkers [21] study pathological variants in cfDNA from 60 patients with colorectal cancer. The concentration of cfDNA in patients with positive disease was significantly higher than that in patients with negative disease. The cfDNA selected to be added to the analysis flow at a specific concentration. This comprehensive framework of cfDNA analysis pipelines has satisfactory sensitivity and specificity for colorectal cancer postoperative surveillance. Wan and coworkers [22] detected whole-exome and whole-genome sequencing of cfDNA from patients with different types of cancer both early and advanced disease. They build patient-specific mutation lists, which offer a way to monitor for signs of relapse with greater sensitivity. Kingston and coworkers [23] used plasma ctDNA sequencing to analyze the genomic profile of 800 patients with advanced breast cancer. Their work demonstrated multiple subclonal resistance mutations, identified distinct mutational processes in advanced ER-positive breast cancer, and identified novel therapeutic opportunities.

To obtain more comprehensive biological information from cfDNA, many scientists are concerned about the development of cfDNA testing methods. In CNV profile from plasma cell-free DNA of cancer patients shows that nanopore sequencing has the same performance as SGS approaches in terms of throughput and sequencing costs [24]. Shen and coworkers [25] reported cell-free methylated DNA immunoprecipitation and high-throughput sequencing (cfMeDIP-seq) as a sensitive, low-input, cost-effective, and bisulfite-free method for the analysis of DNA methylomes of plasma cfDNA. Moreover, it is suitable for methylome analysis of low input DNA samples and capable of generating cfMeDIP-seq libraries from plasma cfDNA within ~3–4 days. To improve the utility of cfDNA in the diagnosis of minimal residual disease. Kurtz and coworkers [26] described phased variant enrichment and detection sequencing

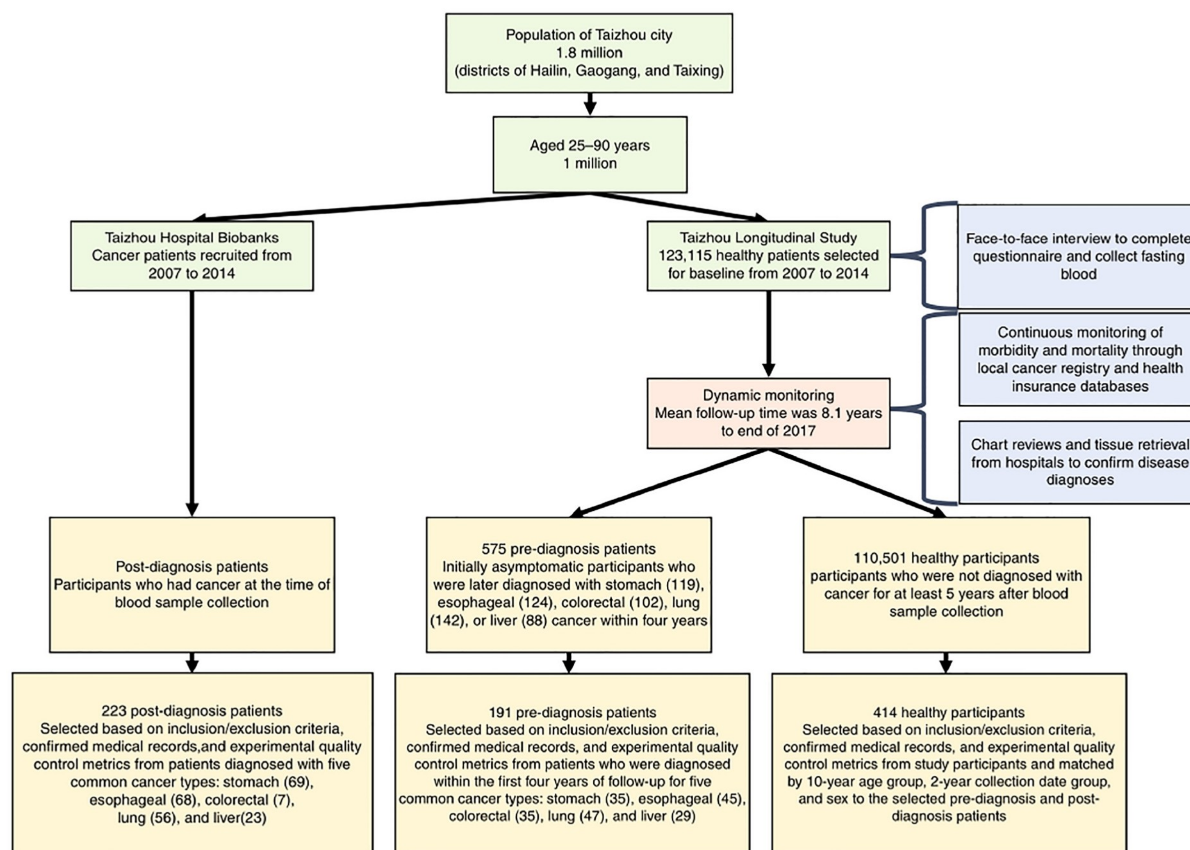


Figure 2 Summary of the Taizhou longitudinal study (TZL). The flowchart shows recruitment, baseline survey, sample collection, and cohort follow-up for TZL. Qualified pre-diagnosis patients and healthy participants were selected from the TZL cohort and qualified post-diagnosis patients were selected from local Taizhou hospital biobanks; 328 samples were processed but later excluded due to not meeting inclusion criteria or failing quality control metrics. Reproduced with permission from Ref. [17] (color online).

(PHased-seq), a method that uses multiple somatic mutations in a single DNA fragment to improve the sensitivity of cfDNA detection, with a better detection limit than previous methods.

Apart from cancer, cfDNA has shown great potential as a clinical biomarker for several diseases. Vlaminck and coworkers [27] extracted cfDNA from plasma and quantified the fraction of donor-derived cfDNA (cf-dDNA) through sequencing (Figure 3a). They used the results of a prospective cohort study to examine the utility of cfDNA in acute rejection after heart transplantation. They demonstrated that cfDNA can diagnose acute rejection after a heart transplant (AUC=0.83, Figure 3b). This non-invasive genomic transplantation dynamic method can be used for routine monitoring of allograft health. cfDNA has been explored as a potential biomarker for diagnosing and monitoring infectious diseases. Blauwkamp and coworkers [28] identified and quantified microbial cell-free DNA from plasma from 1,250 clinically relevant bacteria, DNA viruses, fungi, and eukaryotic parasites. The test showed a 93.7% agreement in blood cultures in 350 patients with sepsis alerts, and identified independently determined sepsis alert causes more frequently than all microbiological tests combined. Caggiano and coworkers proposed CelFiE, an algorithm suitable for low-coverage data, which can accurately estimate the relative abundance of cell types and tissues contributing to cfDNA from epigenetic cfDNA sequencing [29].

Epigenetics, fragmentomics, and the topology of cell-free DNA can also reflect its tissue origin, gene regulation, genome evolution, and disease pathogenesis. DNA methylation is the addition of methyl groups to cytosines, which usually leads to gene repression and is associated with disease. Fedyuk and coworkers [30] profiled the epigenetics of plasma-isolated nucleosomes, DNA methylation, and cancer-specific protein biomarkers with single-molecule multiparametric assay. The system detects histone modifications and their ratios and combinatorial patterns by single-molecule imaging. The technology detected cancer with high accuracy and sensitivity, even at early stages, while revealing the tissue of origin of colorectal, pancreatic, lung, and breast tumors. Wang and coworkers [31] developed a technology, which performed de novo screening of methylation markers on cfDNA samples. Mutations and methylation were tested in parallel in the training cohort. They applied the model to a

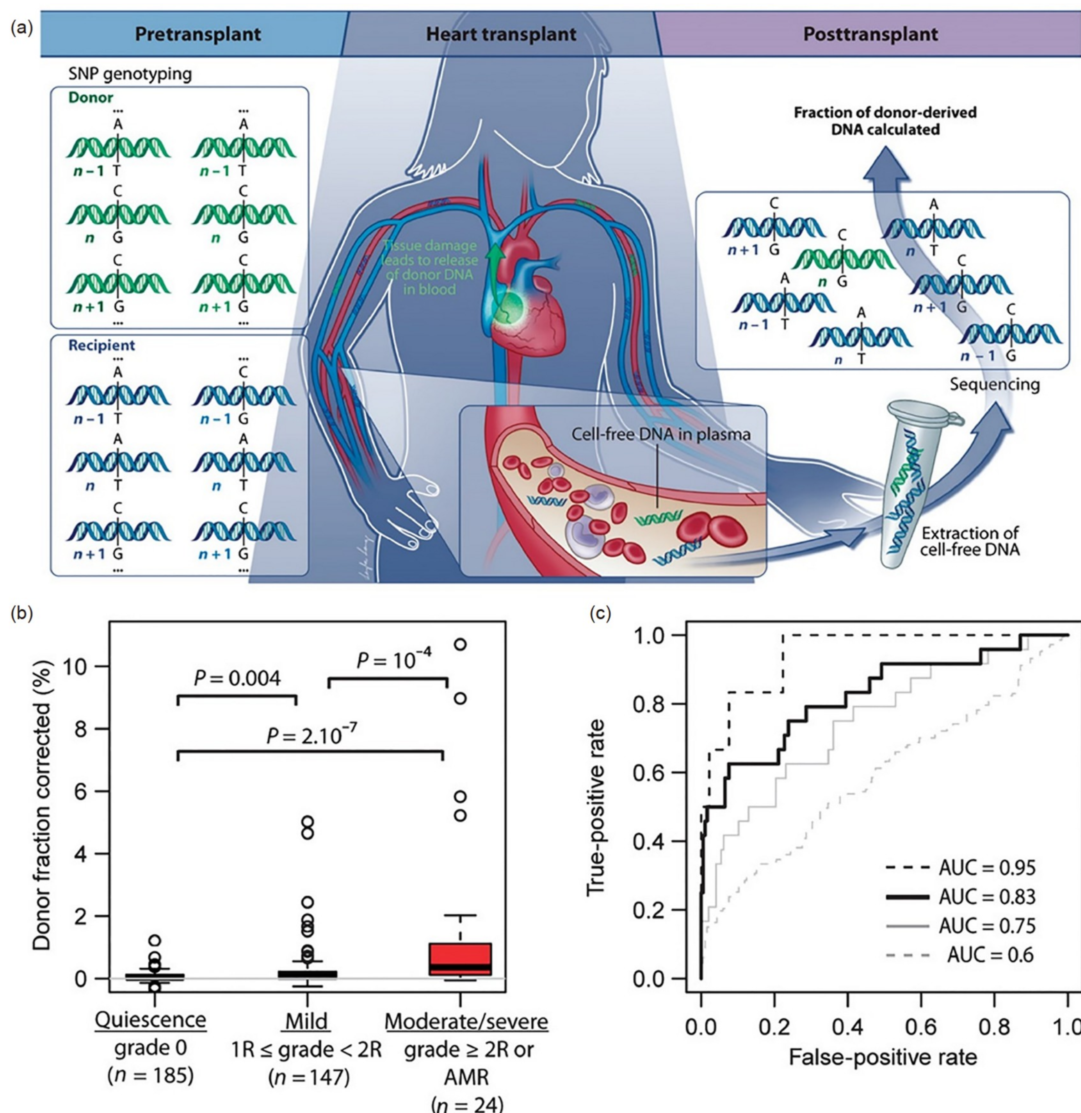


Figure 3 Single-nucleotide polymorphisms (SNPs) distribution of cfDNA enables noninvasive diagnosis of heart transplant rejection. (a) Working principle of the assay. Performance of cfDNA as a marker for heart transplant rejection. (b) Box plots of the fraction of cfDNA for stable heart transplant recipients, recipients diagnosed with mild rejection, and recipients diagnosed with moderate-to-severe rejection. (c) ROC analysis of the performance of cfDNA in classifying moderate-to-severe rejecting (AUC 0.83, black solid line) and non-rejecting recipients (grade 0). Reproduced with permission [27]. Copyright 2014, American Association for the Advancement of Science (color online).

prospective cohort (311 HBV carriers). The model detected four of the five HCC cases in the cohort, showing 80% sensitivity and 94% specificity. Guler and coworkers [32] noninvasively detected pancreatic ductal adenocarcinoma (PDAC) by 5-hydroxymethylcytosine (5hmC) changes in circulating cell-free DNA in PDAC patients. The most significant differential hydroxymethylation was found in genes related to pancreatic development or function and cancer pathogenesis. Genes that are commonly deregulated in PDAC tumors upon KRAS activation and TP53 inactivation are differentially enriched in the cfDNA hydroxymethyl group.

Fragment size and terminal motifs of plasma DNA molecules are related to nucleosome organization and nuclease content of the tissue of origin [33]. The cfDNA fragmentation pattern reflects the nucleosome pattern in white blood cells, and the fragmentation pattern is altered in patients with cancer. Cristiano and coworkers [13] have analyzed the fragmentation profiles of cancer patients with different cancers and healthy people using the assessment method of genome-wide cell-free DNA fragmentation patterns. Fragmentation, when combined with mutational cell-free DNA analysis, detects 91% of cancer patients. Yu and coworkers [34] analyzed a large number of long DNA molecules of fetal and maternal origin in maternal plasma. They observed that the longest fetal-derived plasma DNA exhibited a predominance of either A or G 5′ fragment ends. Pregnant women with preeclampsia have reduced plasma long DNA molecules with a decreased frequency of specific 5′ 4-mer

terminal motifs ending in G or A. They developed a method to determine the tissue origin of long DNA molecules by analyzing the methylation patterns of a series of CpG sites on them. Extrachromosomal circular DNAs (eccDNAs/ecDNAs) with a topological structure may play an important role in a variety of cellular processes. Sin and coworkers [35] isolated eccDNAs/ecDNAs molecules from the plasma of pregnant women, which is more stable than linear DNA, and generated from 5′-untranslated regions (5′-UTRs), exonic regions, and CpG island regions. The eccDNAs/ecDNAs molecules of fetal origin are usually shorter than those of maternal origin. After enriching the circular DNA, they linearized it using the restriction enzyme MspI and sequenced it. They found that numerous genomic annotations of ncRNAs overlapped highly with those of eccDNAs/ecDNAs. eccDNAs/ecDNAs is a potential driver of FGR through immune signaling pathways.

2.2 cfRNA in plasma

Cell-free RNA (cfRNA) offers the possibility to detect cancer, predict the origin of tumor tissue, and determine cancer subtypes. Larson and coworkers [36] performed transcriptomic profiling of cfRNA in cancer stage III breast cancer, lung cancer, and participants. Sixty-eight percent of annotated genes were not detected in the cfRNA of non-cancer individuals. Within these low-noise regions, they identified tissue- and cancer-specific genes that were repeatedly detected in cancer patients and whose levels correlated with shedding rates and RNA expression in paired tissues. Potential use of cf-mRNA for noninvasive monitoring of bone marrow lesions. Ibarra and coworkers [37] used next-generation sequencing-based analysis that cf-mRNA was enriched in transcripts derived from bone marrow compared with circulating cells. Based on longitudinal studies of multiple myeloma and acute myeloid leukemia patients undergoing hematopoietic stem cell transplantation after bone marrow ablation, cf-mRNA levels reflect the transcriptional activity of the hematopoietic lineage resident in the bone marrow during bone marrow reconstitution. During pregnancy, these cfRNAs can provide important biomarkers that reflect some important changes, such as placental health and fetal growth. Ngo and coworkers [38] found that measurement of nine cell-free RNA (cfRNA) transcripts in maternal blood predicted gestational age with comparable accuracy to ultrasonography but at a significant lower cost. They identified seven cfRNA transcripts that accurately classified women whose preterm delivery was 2 months early. Rasmussen and coworkers [39] analyzed pooled transcriptome data from eight independent prospective cohorts and found that the cfRNA signature from a single blood draw reliably predicted preeclampsia with a sensitivity of 75% and a positive prediction rate of 32.3%. cfRNA examines the molecular changes that occur in the brain, helping us understand the etiology and progression of Alzheimer's disease (AD) and identify effective treatment strategies. Toden *et al*. [40] performed a comprehensive analysis of cf-mRNA in AD patients. They found 2,591 dysregulated genes in cf-mRNA in AD patients, which are enriched in biological processes known to be associated with AD.

cf-miRNA refers to the circulating free microRNA in body fluids such as blood, which has certain biological and medical significance. Miyoshi and coworkers [41] evaluated markers based on circulating microRNA (miRNA) for early detection of esophageal squamous cell carcinoma (ESCC). They confirmed overexpression of eight miRNAs (miR-103, miR-106b, miR-151, miR-17, miR-181a, miR-21, miR-25, and miR-93) in serum samples. These eight miRNA markers are superior to current clinical serological markers in distinguishing patients with early ESCC from healthy controls (Table 1) [42–44].

3 Circulating nucleic acids as biomarkers in urine for liquid biopsy

3.1 The source and composition of urine

Urine is produced by filtering blood through the kidneys, transported to the bladder for storage through the ureter, and excreted through the urethra when the desire to urinate occurs. A normal person emits a large amount of urine daily, totaling 1,000–2,000 mL, with many samples available for collection. It is currently widely accepted that urine is a biological fluid composed of water (96%), salts, metabolites, cells, and other biological molecules (such as cfNA) [45]. Urine contains cells including white blood cells, kidney cells, urothelial cells, prostate cells, and exfoliated tumor cells. Urine biochemical components are influenced by diet, medications, and disease states, and when examined properly, these tests provide important insights into the mechanisms and treatments of various clinical diseases and small amounts of compounds found in urine can indicate the health of a human being [46–48]. The concentrations of different types of cfNA in urine vary widely, generally well below 0.01 μM [47]. Bryzgunova *et al*. [49,50] reported that the concentration of cfRNA in urine was 20–140 ng/mL, miRNA may be more resistant to nucleic acid hydrolases due to their small volume sequence, but the exact concentration remains unclear. cfDNA concentrations range from 1 to 200 ng/mL [51]. These cfNA are considered promising biomarkers for disease diagnosis. However, the development and utilization of cfNA have not been fully developed, most likely due to the small amount of cfDNA [52,53] and the complex composition of their systems.

There are two main sources of cfNA in the urine. One is

Table 1 Summary of characteristics of cfNA as biomarkers in blood for disease detection[a)]

Analytical characteristics	Result reported	Cohort characteristics	Cancer type	Ref.
ctDNA mutation	Sensitivity: 87.5% (stages I–III diseases for SCLC); 68.0% (HCC); 63.4% (NSCLC); 60.9% (cervical cancer)	NSCLC (*n*=5,548), colorectal cancer (*n* =1,195), breast cancer (*n*=1,178), upper gastrointestinal cancer (*n*=575), HCC (*n*=571)	Pan-cancer	[16]
ctDNA methylation	Specificity: 95%	Asymptomatic individuals four years before (*n*=605)	Stomach, esophageal, colorectal, lung, liver cancer	[17]
cfDNA mutation	AUC=0.83	Stage I-III breast cancer (*n*=33), healthy (*n*=47)	Breast cancer	[18]
ctDNA mutation	Odds ratio (OR): 28.74	Patients at the beginning of cycle 3 of pembrolizumab treatment (*n*=40)	SCCHN, TNBC, HGSOC, malignant melanoma, MST	[19]
cfDNA mutation	–	Patients with operable gastric cancer (*n*=50)	GC	[20]
cfDNA mutation and concentration	Accuracy=0.767	No clinical evidence (*n*=31), with clinical evidence (*n*=29)	CRC	[21]
cfDNA mutation (mutant allele)	Advanced cancers AUC=0.98 Early-stage AUC=0.80	Cancer (*n*=176), healthy (*n*=45)	Melanoma, lung, renal, glioma, breast cancer	[22]
ctDNA mutation	Sensitivity: 90.9%	ABC with measurable disease within plasma MATCH (*n*=682)	ABC	[23]
cfDNA mutation (CNV)	Sensitivity: 94% Specificity: 89%	Lung cancer (*n*=6), healthy (*n*=4)	Lung cancer	[24]
ctDNA mutation	Specificity: 97%	Large B cell lymphomas (*n*=107)	B cell lymphomas	[26]
cfDNA (microbial)	specificity: 62.7%	Patients that met the sepsis alert criteria (n=350)	Sepsis	[28]
DNA methylation and cancer- specific protein	AUC=0.96	CRC (*n*=63), pancreatic cancer (*n*=10), healthy (*n*=33)	CRC	[30]
cfDNA methylation	AUC=0.93–0.96	Training cohort HCC (*n*=60), non-HCC (*n*=60) Validation set HCC (*n*=58), non-HCC (*n*=198) Prospective cohort HBV carriers (*n*=311)	HCC	[31]
DNA hydroxymethylation (5hmc)	AUC=0.88	PDAC cohort (*n*=64) in comparison with a non-cancer cohort (*n*=243)	PDAC	[32]
cfDNA (fragmentation)	AUC=0.94	Healthy individuals (*n*=245) and patients with seven cancer types (*n*=236)	Breast, colorectal, lung, ovarian, pancreatic, gastric, bile duct cancer	[13]
cfDNA methylation	AUC=0.88	Early-onset preeclampsia (*n*=5), late-onset preeclampsia (*n*=5), normotensive pregnant women (*n*=10)	Preeclampsia	[34]
eccDNAs/ecDNAs	–	Third-trimester pregnancy (*n*=5)	Fetal growth restriction	[35]
cfRNA expression profiles	*P* value=7×10^{-13} (tissue-specific markers for lung DCBs), and *P* value =3×10^{-9} (for breast DCBs).	Sage III breast cancer (*n*=46), lung cancer (*n*=30) and non-cancer (*n*=89) participants from the Circulating Cell-free Genome Atlas	Stage III breast cancer, lung cancer	[36]
cfRNA transcripts	Discovery (AUC=0.86) Validation (AUC=0.81)	Full-term (*n*=25) Preterm deliveries (*n*=13)	Preterm deliveries	[38]
cf-mRNA molecular changes	AUC=0.83	Training cohort Ads (*n*=66), NCIs (*n*=24); Testing cohort, ADs (*n*=60), NCIs (*n*=92)	AD	[40]
miRNAs(overexpression)	AUC=0.92, 0.93	Training cohort (*n*=408) Validation cohort (*n*1=126, *n*2=165), Prospective cohort (n1=185, n2=188)	ESCC	[41]
cfDNA mutation	Sensitivity: 92%	681 patients who were enrolled in an IRB-approved research protocol	31 distinct tumor types	[42]
cfDNA (SVP)	AUC=0.83	Heart transplant recipient (*n*=21) Adult (*n*=44)	Heart transplant rejection	[27]

(*To be continued on the next page*)

(*Continued*)

Analytical characteristics	Result reported	Cohort characteristics	Cancer type	Ref.
ctDNA mutation	Mutation concordance: 83.4%	Metastatic urothelial carcinoma (mUC) patient (*n*=104)	Metastatic bladder cancer	[43]
Tumoural exRNA concentration	–	Mouse (*n*=30)	Xenografted with 11 different tumours	[44]
cf-mRNA transcripts	–	–	Multiple myeloma, acute myeloid leukemia	[37]
cfRNA transcripts	AUC=0.76–0.88	pre-eclampsia (*n*=72) Healthy (*n*=452)	Pre-eclampsia	[39]

a) NSCLC, non-small cell lung cancer; HCC, hepatocellular carcinoma; SCCHN, squamous cell carcinoma of the head and neck; TNBC, triple-negative breast cancer; HGSOC, high-grade serous ovarian cancer; GC, gastric carcinoma; MST, mixed solid tumors; CRC, colorectal cancer; ABC, advanced breast cancer; PDAC, pancreatic ductal adenocarcinoma; DCB, dark channel biomarker; ESCC, esophageal squamous cell carcinoma; NCIs, non-cognitive controls; AD, Alzheimer's disease.

released by dead cells of the non-urinary system or secreted by living cells, into the blood and the crude urine when the blood flows through the kidney. The other is secreted by organs within the urinary system, such as the kidney and bladder. cfNA of varying sizes are selectively filtered due to the limited permeability of the basement membrane and slit membrane between the glomerular podiatocytes' pedicles. For example, only complexes less than 6.4 nm in diameter and with molecular weight no greater than 70 kDa can enter the nephron lumen; it corresponds to DNA that is about 100 bp in size. From the physiology of urine formation, urine contains information from tissues throughout the body and is a microcosm of health and disease in the body. Studies have shown that the concentration of cfDNA in the blood is relatively low in healthy individuals but increases in the presence of tissue damage or change (such as cancer), and cfDNA may increase by a factor of 10 when compared with healthy controls [54]. The concentration of DNA in the urine of healthy volunteers was comparable to that in plasma [55]. In addition, tumor-specific RNA was detected in the plasma of cancer patients [56,57]. Judging from the "affinity" between urine and blood, cfNA in urine also correspondingly inherited cancer-related biomarkers transmitted from blood. Moreover, because urine is in direct contact with urothelial tumor cells, studies that detect biomarkers in the urine associated with cancers of the urinary system, such as bladder cancer, and prostate cancer, are attractive [58]. In recent years, with the development of nucleic acid enrichment and separation and gene sequencing technology, the enrichment, separation, and analysis of trace cfNA from body fluids have been realized. The prospect of urine cfNA as markers to provide disease diagnosis information from the level of genetic molecules is promising.

3.2 Techniques for the extraction and detection of cfNA in urine

Over the past two decades, there has been increasing attention to cfNA, including cfDNA, mRNA, and microRNA (miRNA), which are considered promising biomarkers in liquid biopsies [59–62]. Due to the characteristics mentioned above, such as extremely low concentration, short length, and easy degradation of cfNA in urine, which are not conducive to detection, to solve this problem and realize the non-invasive precision molecular diagnosis of cancer and other diseases, researchers have developed some techniques and methods for the extraction and detection of cfNA in urine. The extraction methods of intracellular nucleic acid and plasma cfDNA are relatively mature after great development. Unfortunately, the traditional boom method is not suitable for the extraction of urine cfNA, as it is not designed for short fragments and some new suitable methods for the extraction of urine cfNA have been developed.

For the extraction of cfDNA from urine, the Wizard/GuSCN method uses GuSCN with a high concentration (>3 mol/L) of high dissociation sequence to adsorb DNA to Wizard silica [55]. This method was originally used to demonstrate the presence of cfDNA in urine. The Q agarose gel method uses quaternary ammonium anion exchange resin to pre-concentrate DNA before desalting it on a rotating column of silica. It improves the recovery of short segments of urine cfDNA compared with Wizard/GuSCN in tumors for cancer diagnosis, monitoring, and prognosis [63]. And there are commercialized kits for the extraction of cfDNA in urine. Oreskovic and colleagues [51] developed a method for the capture of urine cfDNA using botanized sequentially specific probes and streptavidin-coated magnetic beads, and compared their method with the Wizard/GuSCN method, the Q agarose gel method, and three kits (QC, NU, MM). It was found that their hybridization capture and agarose gel methods performed best, with high recovery rates of short fragments (as low as 25 and 40 nt, respectively), sensitive detection of diluted fragments, good tolerance to different urine conditions, and resistance to PCR inhibition, which is recommended for the extraction of urine cfDNA from clinical samples.

Compared with cfDNA, the methods for the extraction of cfRNA from urine were developed relatively late. There are three main types of biomarkers: microRNA (miRNA), messenger RNA (mRNA), and long-chain non-coding RNA (lncRNA). Several methods [64–69] have been reported for the isolation of urinary cfRNA with commercial kits. After the cfNA is isolated and extracted, precise sequencing is required, and new analytical techniques such as quantitative PCR (qPCR), microdrop digital PCR (ddPCR), or next-generation sequencing (NGS) are advancing the field by allowing more accurate detection of smaller amounts and fragments of cfDNA [70–73].

3.3 cfNA in urine as a marker to indicate disease

Although blood has been the primary source of cfNA and other biomolecules as biomarkers, urinary cfNA is also an important biomarker associated with diseases such as cancer [67,74,75]. Bryzgunova *et al*. [76] isolated and compared different cfNA in human urine, and their work demonstrated the applicability of urinary cfDNA in cancer diagnosis. Urine is the preferred sample for the discovery of biomarkers of urinary diseases due to its non-invasive collection, relative stability, and exposure to damaged microenvironments. However, the diagnostic applications of urinary cfNA as biomarkers are far beyond urinary diseases. They are also useful in the diagnosis of lung cancer breast cancer, gynecological cancer, gastrointestinal cancer [45], hypertensive diseases [77], and neurodegenerative diseases of Parkinson's disease [62].

3.3.1 cfDNA in urine

cfDNA can be released directly into the urine *via* necrotic and apoptotic cells in the genitourinary system. The association of urinary cfDNA with cancers of the urinary system has been extensively reported. Sidransky *et al*. [78] first reported the presence of the p53 mutated gene in urinary sediment in patients with invasive bladder cancer in 1991. Given the quality and accuracy of urinary prostate cancer gene 3 (PCA3), Hessels *et al*. [79] developed an RT-PCR quantitative assay for PCA as a molecular urine analysis tool to provide a basis for molecular diagnosis in clinical urological practice. Urinary PCA3 detection has been approved by Food and Drug Administration (FDA) as a diagnostic tool for prostate cancer [80]. In addition to this, in hepatocellular carcinoma, Lin *et al*. [81] have demonstrated elevated levels of urinary cfDNA p53 mutations and may be explored for screening. Su *et al*. [82] reported a higher incidence of K-ras gene mutations detected in the urine of patients with colorectal cancer or adenomatous polyps than those detected in serum and plasma. Urinary cfDNA has also been reported by Chen *et al*. [83] for EGFR mutation studies in non-small cell lung cancer, and they found no significant differences in the sensitivity and specificity of urinary cfDNA compared with plasma. In addition to detecting mutations of cfDNA in urine, Zhang *et al*. [84] compared serum and urine ctDNA levels in 200 breast cancer patients and healthy volunteers using a ddPCR technique. They found that urine circulating tumor DNA (ctDNA) and wild-type PIK3CA genotype levels were 3.5 times higher in patients with early-stage breast cancer than in healthy volunteers. Studies are also being conducted to examine specific modifications of cfDNA in urine that are associated with cancer. A test designed by Nuzzo *et al*. [85] called methylated cfDNA immunoprecipitation and high-throughput sequencing (cfMeDIP-seq) can detect early renal cell carcinoma from plasma and urine samples using small amounts of DNA (⩽10 ng). Bach and colleagues [86] measured cfDNA methylation levels for six CRC-related markers in 40 milliliters of urine from patients with colorectal cancer (CRC) and healthy volunteers. The SEPT9 methylation analysis provided the first evidence for CRC tests in urine at levels close to those found in plasma SEPTIN9 methylation CRC tests (75% to 81%), which have been approved for testing by the FDA. In urine, the nucleic acid is affected by higher levels of DNase activity, which breaks down cfDNA [45]. However, the amount of nuclease present in the urine may provide us with another useful biomarker. Zhou and colleagues [87] found that urine cfDNA concentrations (double-stranded DNA with single-strand protruding) with jagged ends were higher than plasma cfDNA, possibly due to different DNase activity levels (Figure 4). In urine samples from bladder cancer patients, they found that urine cfDNA levels with jagged ends were lower than those of healthy volunteers, possibly due to cancer-induced reduced nuclease activity.

3.3.2 cfRNA in urine

Urinary cfRNAs, such as mRNA, miRNA, lncRNA, and cirRNA, have been reported to be useful as biomarkers for urinary cancers. For cfRNA from urine, researchers mainly focus on detecting the difference in the expression level of cfRNA in urine to indicate the occurrence of diseases. Kim *et al*. [88] identified urinary UBE2C cf-mRNA levels as a possible diagnostic marker for bladder cancer, and their work is the first study to identify urinary UBE2C cf-mRNA as a diagnostic marker. Urquidi *et al*. [65] used a multivariate model to identify an optimal 18-gene diagnostic signature set by monitoring 44 urine-like mRNA transcripts from 196 (89 patients with bladder cancer). This mRNA biomarker diagnostic team was able to accurately detect bladder cancer using non-invasive urine assays. miRNA is a class of small non-coding short single-stranded RNA (22–24 nt in length) that is involved in cell proliferation, differentiation, stress response, inflammation, and cell death, and plays a role in RNA silencing and post-transcriptional regulation of gene

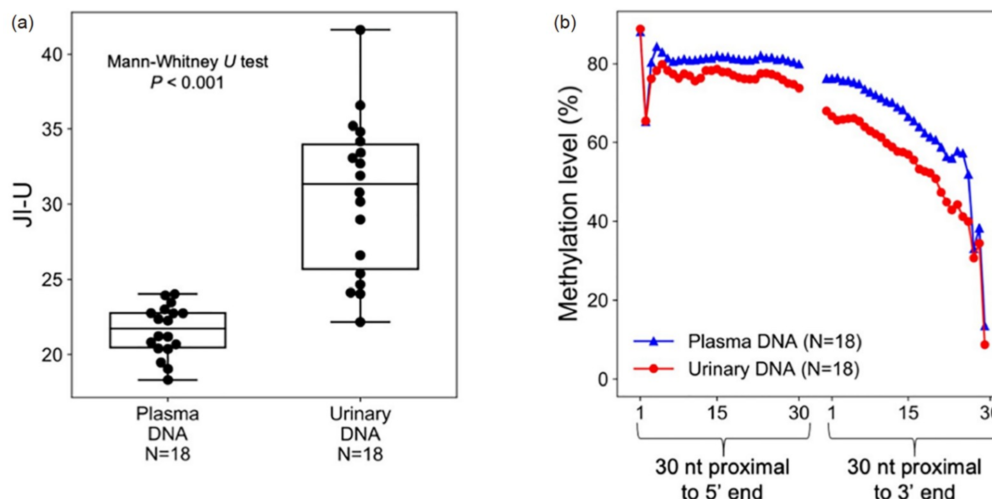


Figure 4 Comparison of jaggedness between plasma and urinary DNA. (a) Jagged index-unmethylated (JI-U) values of plasma DNA and urinary DNA of control patients with hematuria. (b) Methylation levels of plasma DNA (blue) and urinary DNA (red) across different loci at the first 30 nucleotides (nt) and the last 30 nt of a cfDNA fragment. Reproduced with permission from Ref. [87] (color online).

expression [89]. Alterations in miRNA have been implicated in the pathogenesis of various types of human cancer, and these show potential as novel biomarkers due to the stability of tumor-derived free miRNA [15]. Mall *et al*. [90] investigated the stability of miRNA in human urine and found that miRNA was more resistant to nucleases than mRNA that was easily degraded by RNA hydrolases, and that miRNA was relatively stable in urine under a variety of storage conditions, supporting their usefulness as urine biomarkers. Piao *et al*. [66] explored a new method for distinguishing bladder cancer from hematuria in non-malignant diseases, by measuring differences in urine free miR-6124 and miR-4511 expression between BC patients and hematuria patients, allowing non-invasive diagnosis and thus reducing unnecessary cystoscopy in hematuria patients being evaluated for BC, with minimal loss of sensitivity to detect cancer. Giri *et al*. [62] suggested that miRNAs, which are overexpressed in patients with Parkinson's disease (PD), Alzheimer's disease, and dementia, could be used as biomarkers in urine for early clinical diagnosis and prognostic detection of these neurodegenerative diseases. Hung *et al*. [91] reported that miR-376c promotes the proliferation and migration of gastric cancer cells and is increased the urine and plasma levels of patients with gastric cancer, indicating that miR-376c in urine has the potential to be used as a marker of gastric cancer cells. Long non-coding RNAs (lncRNA) are transcripts larger than 200 nucleotides in length and do not encode proteins. They are gene regulators involved in many biological functions and are disorders in various cancers [92]. Iyer *et al*. [93] applied *de novo* assembly to RNA sequencing (RNA-seq) libraries from multiple tumors, revealing thousands of pedigrees and cancer-associated lncRNAs, highlighting the importance of incorporating lncRNA into biomarker and therapeutic target discovery algorithms. Bussemakers *et al*. [94] found that prostate cancer antigen 3 (PCA3) is overexpressed in more than 95% of prostate cancers in 1999. PCA3 was the first lncRNA localized to chromosome 9q21-22. Human uroepithelial carcinoma-associated 1 (UCA1), a 2,314-bp lncRNA located on human chromosome 19, is upregulated in many cancers, such as hepatocellular carcinoma, colorectal cancer, gastric cancer, esophageal squamous cell carcinoma, and epithelial ovarian [45]. circRNA is a class of single-stranded covalent closed RNA molecules produced by pre-mRNA through a reverse splicing process. Studies have suggested that circRNA may be involved in miRNA inhibition [95] and tumorigenesis [96]. Studies have shown that circRNA is an evolving class of promising cancer biomarkers [97]. Cieslik *et al*. [60] developed a method to detect prostate cancer tissue-associated circRNA, such as circ-CPNE4 and circ-ACPP, in non-invasive urine assays in prostate cancer patients starting from a small amount of RNA (50 ng) (Table 2).

3.4 Summarization

Over the past decade, urine as a biomarker source for liquid biopsies has gradually moved toward the clinical prediction of multiple diseases, and there has been increasing attention to cfNA in urine, including cfDNA, mRNA, and miRNA, which are considered promising biomarkers in liquid biopsies. The current research has focused on the feasibility of finding disease-indicating cfNA in urine, but further clinical trials and technical optimization are needed before they can be applied to the clinic. Free nucleic acid molecules in urine may be more sensitive to disease than MRI and traditional methods and can detect both urinary and non-urinary diseases. Despite the benefits of non-invasive, simple equip-

Table 2 Summary of characteristics of cfNA as biomarkers in urine for disease detection[a)]

Analytical characteristics	Result reported	Cohort characteristics	Cancer type	Ref.
mRNA transcripts (44 candidate diagnostic biomarkers)	AUC=0.935	Bladder disease status (*n*=96), BC (*n*=89)	BC	[65]
RNA expression (miR-6124 to miR-4511)	Sensitivity >90% (AUC=0.888, *p*<0.001)	BC (*n*=326), hematuria (*n*=174) and pyuria without cancer (*n*=43)	BC, hematuria	[68]
cfDNA quantification ($DD3^{PCA3}$ transcripts)	AUC=0.72	Patients admitted for prostatic biopsies with serum PSA level>3 ng/mL (*n*=108)	Prostate cancer	[79]
DNA mutation (TP53)	Sensitivity: 0.1% (the mutant/ wild-type ratio)	HCC (*n*=17)	HCC	[81]
cfDNA mutation (K-ras)	Detected K-ras DNA from urine was significantly higher than that from plasma (*P*-value=0.00043)	CRC or Adenomatous Polyps (*n*=20)	CRC, adenomatous polyps	[82]
cfDNA mutation (EGFR)	AUC=0.976	Activating EGFR mutation and received EGFR-TKIs (*n*=150)	NSCLC	[83]
Quantities of ctDNA (PIK3CA)	AUC=0.965 (plasma), AUC=0.972 (urine)	Breast cancer (*n*=200), Healthy (*n*=50)	Breast cancer	[84]
cfDNA methylomes	AUC=0.99 (plasma), AUC=0.86 (urine)	Healthy (*n*= 28); stage I–IV RCC cases (*n* = 99); Stage IV UBC samples (*n*= 21)	RCC	[85]
ctDNA methylated fragments SEPT9, TMEFF2, SDC2, NDRG4, VIM and ALX4)	Specificity: 86%	CRC (*n*=92) and healthy (*n*=63)	CRC	[86]
cfDNA (jagged ends)	AUC=0.83	BC (*n*=46) and without BC (n=39)	BC	[87]
cfRNA levels (UBE2C)	AUC=0.839	BC (*n*=212); healthy (*n*=64); hematuria (*n*=42)	BC, hematuria	[88]
miRNA expression level (miR-376c)	AUC=0.70	GC (*n*=65), control samples (*n*=108)	GC	[91]

a) BC, bladder cancer; UBC, urothelial bladder cancer; RCC, renal cell carcinoma.

ment, and patient-friendly access to urine samples, this urine-based molecular diagnostic approach to cfNA requires scaling up study sizes and sample diversity for prospective validation in large cohorts. Because urine samples are easily affected by non-diseases such as food, drugs, and the environment. Due to the previously mentioned unfavorable characteristics of cfNA in urine, such as extremely low concentration, short length, and easy degradation, it is still necessary to develop extraction methods that can extract and capture cfNA in sufficient quantity and purity from urine systems containing complex components while minimizing damage to it, at the lowest possible cost for clinical application. It is also important to develop techniques for accurately detecting cfNA sequences with high sensitivity and to analyze the fragmented cfNA to extract useful information. As a source of multiple omics information including metabolomics, proteomics, genomics, and transcriptomics, it is also a major trend of future development to build multiple omics platforms for interaction, information interchange, and mutual verification. There is no doubt that with more research and technological advances, cfNA in urine as non-invasive molecular diagnostic tool will one day be useful for patients in the diagnosis of a variety of diseases, including cancer.

4 Circulating nucleic acids as biomarkers in bile for liquid biopsy

4.1 The source and composition of bile

Bile is secreted by the liver and flows into the bile duct tree through the hepatic bile duct, which connects the liver, pancreas, gallbladder, and other digestive organs to the duodenum [98]. Bile is drained by biliary ducts from the liver into the gallbladder, where it is stored and concentrated approximately tenfold [99,100]. It mainly consists of water, in which there are organic and inorganic substances in suspension, dissolved, or in equilibrium between both states. The most important components of bile are organic molecules, including bile salts, fatty acids, cholesterol, proteins, phospholipids, bile pigments, etc., and contain a certain level of inorganic salt ions [101]. Bile secretion, flow, and storage properties allow more opportunities for bile to come into direct contact with biliary pancreatic malignancy. Bile not only bears digestive functions but also contains metabolites. A series of changes occurring in the processes of cancer cell genesis, proliferation, and metastasis may be detected in bile. Bile is more sensitive and specific for biliary tract diseases [102]. More mutations are detected in bile than in plasma, and the concordance between bile and tumors is higher

[103,104]. The sensitivity and positive predictive value of bile are higher compared with cytology with ERCP [105]. And the AUG value is significantly higher than that of urine and plasma. The unique advantages of bile in the detection of biliopancreatic malignancies make it more suitable as a biopsy fluid for biomarker development.

4.2 Characterization and extraction of cfNA from bile

The main components of bile were different from those of serum. Phenol chloroform extraction is a classical method for extracting nucleic acids from liquid, but it is not efficient. Most commercial kits on the market use spin columns with silica membranes to bind cfNA but with preference. Some kits capture free nucleic acids by selective adsorption of magnetic beads [106]. By analyzing cfDNA in bile, the researchers found that cfDNA in the bile with higher levels [103] and length distributions. cfDNA fragments >6,000 are widely present in bile [106–108]. Moreover, the particularity of bile components may lead to difficulties in downstream detection, such as, biliverdin may affect fluorescence signals [109]. There is still room to improve the extraction efficiency of cfNA in bile.

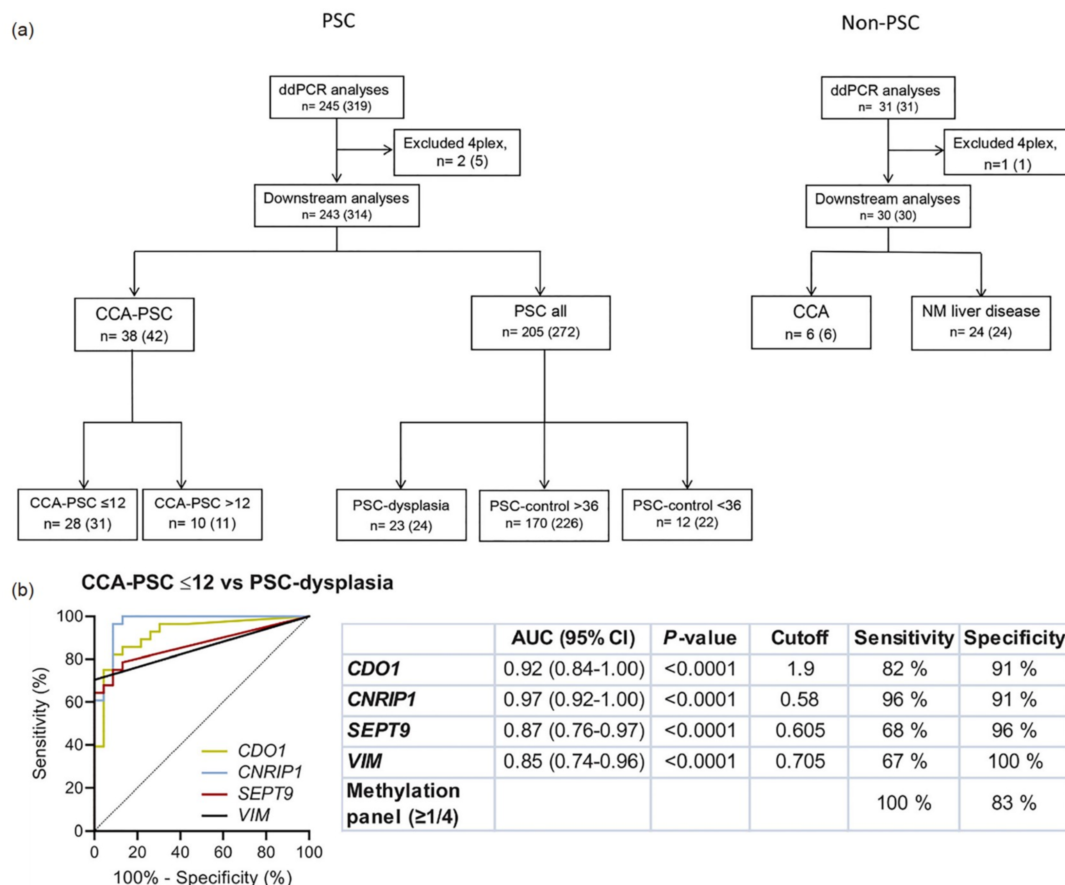
4.3 cfNA in bile as a marker to indicate disease

Bile is used for the detection and analysis of biliary duct-related diseases, among which biliary tract carcinoma (BTC) is a heterogeneous malignant tumor with biliary duct differentiation characteristics, including cholangiocarcinoma and gallbladder carcinoma. Cholangiocarcinoma is divided into intrahepatic cholangiocarcinoma (iCCA), perihilar cholangiocarcinoma (pCCA), and distal cholangiocarcinoma (dCCA) according to its anatomical site [110]. In recent years, most cfNA tumor markers of bile have been developed for CCA. In addition to other benign liver diseases, the control cohort also focused on primary sclerosing cholangitis (PSC), a group of people at potential risk for CCA. CCAs are the second most common primary liver tumor and their incidence is increasing worldwide. It is characterized by strong invasiveness, late diagnosis, and high mortality [111].

4.3.1 cfDNA in bile

Detection of cfDNA mutations at the early diagnostic stage of biliary stenosis can significantly improve the detection of malignancies, reduce delays in patient clinical management, and help select patients for targeted therapy [108]. The consistency of ctDNA mutations between bile and tissue samples was high, while that between plasma and tissue samples was less than 50% [102,104]. The expression levels of KRAS-related signaling oncogenes in bile and tissue samples showed a strong positive correlation. Liquid biopsies of bile reliably detect mutated variants in the ctDNA of BTC patients [104]. These results indicate that bile is an effective biopsy fluid for ctDNA analysis. At present, the carcinoembryonic antigen (CEA) and the carbohydrate antigens 19-9 (CA 19-9) are used clinically as serum biomarkers in the monitoring and diagnosis of biliary tract tumors, but the sensitivity and specificity are not adequate for early detection [112]. Detection of cfDNA mutations at the initial diagnostic stage of biliary stenosis can significantly improve the detection of malignancies, reduce delays in patient clinical management, and help select patients for targeted therapy [103]. The frequency of TP53, KRAS, CDKN2A, PIK3CA, and ERBB2 oncogene mutations increased significantly in the bile of patients with biliary diseases. The mutation rate of alleles (MAF) was much higher than that of plasma and tissue [102,103]. The detection of cfDNA single nucleotide variation (SNV)/insertion and deletion (Indel) in bile showed high sensitivity and specificity [107,113]. The copy number variant also has a low recurrence amplified gene consistent with the tumor [107].

In addition to cfDNA changes such as gene mutations and single nucleotide variation, epigenetic modifications such as methylation in biliary tract diseases also play an important role in the occurrence and development of cancer. In 2003, Klump and coworkers [114] adopted the difference in DNA promoter methylation in cancer and non-cancer patients as a screening condition for tumor markers. They found that cyclin dependent kinase inhibitor 2A (CDKN2A) was methylated in 52% (p16) and 48% (p14) of bile samples from patients with CCA and only 6% of healthy individuals. Shin and coworkers [115] analyzed 17 aberrantly methylated biomarkers in dCCA in training and validation set of bile specimens. In an independent test set of bile fluid samples, a five-gene panel detected CCA at a sensitivity of 83%. Vedeld and coworkers [116] analyzed DNA methylation of cysteine dioxygenase type I (CDO1), cannabinoid receptor interacting protein type I (CNRIP), SEPT9, and vimentin in CCA patients and benign PSC patients with another liver disease by ddPCR. Positive methylation markers were found in all bile samples obtained 12 months before the definitive CCA diagnosis. They used PSC and few false positives detected were all among the PSC controls; and if only including other NM liver diseases excluding PSC (Figure 5a), they obtained a specificity of 100%. The ability to distinguish the two groups remained high with AUCs (Figure 5b). Shun and coworkers [117] established a bile screen model through many samples, which combined mutation and methylation, at least one mutant KRAS, TP53, and other genes, and gradually selected five known methylated genes such as SEPT9 as the prediction model through punitive logistic regression method. Either positive is defined as positive, with a sensitivity and specificity of 93% and 98%, respectively, and a predictive value of 95% for potential tumor patients through



	AUC (95% CI)	*P*-value	Cutoff	Sensitivity	Specificity
CDO1	0.92 (0.84-1.00)	<0.0001	1.9	82 %	91 %
CNRIP1	0.97 (0.92-1.00)	<0.0001	0.58	96 %	91 %
SEPT9	0.87 (0.76-0.97)	<0.0001	0.605	68 %	96 %
VIM	0.85 (0.74-0.96)	<0.0001	0.705	67 %	100 %
Methylation panel (≥1/4)				100 %	83 %

Figure 5 Early and accurate detection of cholangiocarcinoma (CCA) in patients with primary sclerosing cholangitis (PSC) by methylation markers in bile. (a) Flowchart of bile samples included in the study. (b) Receiver operator curves (ROC), calculated AUCs, and sensitivity and specificity values for the four individual DNA methylation biomarkers in bile, samples from patients with CCA-PSC⩽12 (n=28) *versus* PSC-dysplasia (n=23). Reproduced with permission from Ref. [116] (color online).

patient follow-up. Biliary cfDNA can accurately reflect the physiological state of the tumor tissue and can detect the genomic changes determined by the tumor tissue. Bile biomarkers are of great significance for screening, risk assessment, early diagnosis, and prognosis of tumors.

4.3.2 cfRNA in bile

Although the constituents of bile are complex, endogenous miRNAs can exist stably in bile, which provides a prerequisite for their development as biomarkers [118]. Due to the stability of miRNA, the studies of bile cfRNA mostly focused on miRNA, and there are still a few studies on mRNA. miR-9 and miR-145* are candidate biomarkers for the diagnosis of cholangiocarcinoma, potential candidate genes with high sensitivity and specificity, and can be used as good diagnostic markers of BTC by ROC curve analysis [118,119]. miRNA concentrations in bile were significantly higher in patients with PSC than in patients with CCA, and four miRNAs were significantly different between patients with PSC and PSC/CCA. Among them, miR-412 was up-regulated in cholangiocarcinoma, while miR-640, miR-3189, and miR-1537 were down-regulated, and the area under the ROC curve was between 0.78 and 0.81 [120]. Han and coworkers [121] screened 1209 miRNAs on the miRNA microarray platform and found that miR-30d-5p and miR-92a-3p were significantly up-regulated in bile in the CCA group compared with the BBD group. miR-30d-5p had the best diagnostic effect, with a sensitivity of 81.1% and a specificity of 60.5%. Meanwhile, the miRNA wrapped in exosome of bile is also concerned, such as miR-451a and miR-3619-3p, which were identified as reproducible upregulated markers [122]. Uchida and coworkers [123] used endoscopic transpapillary cholecystectomy (ETCG) to collect and detect human telomerase reverse transcriptase mRNA in gallbladder bile to investigate its value in the diagnosis of gallbladder cancer. Sensitivity and specificity were 83.3% and 100% in 12 cases of gallbladder cancer and 8 cases of cholecystitis (Table 3).

5 Circulating nucleic acids as biomarkers in other body fluids for liquid biopsy

Genetic analysis and diagnosis of pancreatic bile duct dis-

Table 3 Summary of characteristics of cfNA as biomarkers in bile for disease detection[a)]

Analytical characteristics	Result reported	Cohort characteristics	Cancer type	Ref.
Gene mutation KRAS, TP53	–	PDAC/CCA (*n*=21), non-malignant biliary obstructions (*n*=21)	CCA PDAC	[102]
Gene mutation	MAF 3.84%	BTC (*n*=13)	BTC	[103]
Gene mutation (CNV, SNV)	–	BTC (*n*=10)	BTC	[107]
Gene mutation KRAS, TP53	Sensitivity 96.4% Specificity 69.2%	Benign (*n*=26), indeterminate (*n*=9), malignant (*n*=33)	CCA, PDAC	[108]
DNA methylation (*CDO1*, *CNRIP*, *SEPT9*, VIM)	AUC =0.77–0.87	CCA-PSC (*n*=38), PSC (*n*=205), CCA (*n*=6), NM liver disease (*n*=24)	CCA	[116]
DNA methylation CDKN2A (p16, p14)	–	CDL (*n*=5), PSC (*n*=11), BTC (*n*=23), GBC (*n*=5), NAD (*n*=6)	CCA	[114]
DNA methylation (CDH13, GRIN2B, RUNX3, and TWIST1)	Specificity100.0 % Sensitivity 83.3%	Training cohort malignant (*n*=116), benign (*n*=93) Validation cohort (*n*=45) Test cohort (*n*=40)	CCA	[115]
DNA methylation and mutation	AUC=0.85–0.96	Training cohort (*n*=104), validation cohort (*n*=105), test cohort (*n*=50)	Pancreatobiliary tract cancer	[117]
miR-9 and miR-145*	AUC=0.975	Malignant (*n*=9), benign (*n*=9)	BTC	[118]
miR-412, miR-640, miR-1537 and miR-3189	AUC=0.78-0.81	PSC (*n*=52), CCA (*n*=19), PSC-CCA (*n*=12), Healthy individuals (*n*=12)	PSC/CCA	[120]
miR-30d-5p and miR-92a-3p	AUC=0.730 AUC=0.652	Training cohort BBD (*n*=10), CCA (*n*=11) Validation cohort BBD (*n*=48), CCA (*n*=37)	CCA	[121]
Exosomal microRNAs miR-451a and miR-3619-3p	AUC=0.819	BTC (*n*=34), non-cancer (*n*=43)	BTC	[122]
cf-mRNA (hTERT)	Sensitivity 83.3% Specificity 100%	Gallbladder cancer (*n*=12), cholecystitis (*n*=8)	Gallbladder cancer	[123]

a) PSC, primary sclerosing cholangitis; CDL, choledocholithiasis; NAD, nothing abnormal detected; NM, non-malignant; BD benign biliary disease.

eases are not only being explored in bile, researchers are also detecting mutations of cfDNA in pancreatic fluid to identify the development and progression of tumors. Kosuke Nagai and coworkers [124] obtained pancreatic fluid from a study of 50 patients with pancreatic bile duct stricture undergoing endoscopic retrograde cholangiopancreatography. The concentration of cfDNA in the pancreatic fluid of the intraductal papillary myxoma group was higher than that of other groups. The sensitivity, specificity, and predictive values of cfDNA analysis were better than those of cytological analysis in the pancreatic fluid. The results indicate that pancreatic fluid-derived cfDNA can be used to diagnose pancreatic bile duct stenosis. Currently, other body fluids used in liquid biopsies include cerebrospinal and pancreatic fluids, which do not circulate throughout the body and may be better suited for disease-specific tests. Detection of tumor-derived cell-free DNA in the blood of patients with brain tumors is challenging due to the blood-brain barrier. De Mattos-Arruda and coworkers [125] confirmed that ctDNA derived from central nervous system tumors was more abundant in cerebrospinal fluid (CSF) than in plasma. Massive parallel sequencing of CSF ctDNA more comprehensively characterized the genomic changes of brain tumors, and CSF ctDNA levels varied with changes in brain tumor load, providing biomarkers for monitoring brain malignant tumors. Miller and coworkers [126] evaluated the cerebrospinal fluid of 85 patients with glioma and found that ctDNA was detected in the CSF in 42 patients (49.4%) and was associated with disease burden and adverse outcomes. Mutations that occur early in tumorigenesis, such as codeletion of chromosome arms (1p/19q deletion) and IDH1 and IDH2 mutations, are existed among all matched ctDNA-positive cerebrospinal fluid tumor pairs. Escudero and coworkers [127] reported that ctDNA analysis of medulloblastoma patients reproduces the genomic alterations of the tumor. ctDNA contributes to subgroup and risk stratification and provides valuable information about diagnosis and prognosis. The detection of glioma genome mutations by liquid biopsy techniques can facilitate the clinical development and use of genotype-targeted therapies for glioma, one of the most aggressive human cancers, and the monitoring of brain tumor metastasis.

In conclusion, liquid biopsy is a non-invasive means of detection, which can be used as an effective method of early diagnosis, screening, monitoring, prognosis, and treatment of malignant tumors, early eclampsia, and organ transplantation diseases. With the development of gene sequencing technology and bioinformatics technology, the analysis and

identification of free nucleic acids in body fluids can accurately discover the etiology and therapeutic targets of diseases. From the current research process, free nucleic acid in body fluids is closely related to the process and state of disease development, which provides strong support for the development and improvement of precision medicine.

6 Limitations, challenges, and future development trends for circulating nucleic acids

The main difficulties that limit its technical translation and clinical application are the complexity of the cfNA pool and the technical limitations of purification [8]. Firstly, cancer-related changes are subtle and submerged in a flood of noninformative cfNA without diagnostic value. It is challenging to parse disease-related information from the complex cfNA pool from different tissues, which depends on the development of bioinformatic technology. The second point is the limitation of extraction technology, the protective layer of the RNA is stripped off, and the cfRNA becomes fragile during the isolation process, so the extraction process should be fast and maintain the integrity of the RNA [128]. The quantitative detection results of miRNAs in body fluids are highly variable in different reports, possibly due to the small size of miRNAs and their attachment to lipids and proteins, so efficient and reproducible extraction methods are required. Therefore, there is an urgent need for a rapid, bias-free extraction method that preserves the integrity of cfRNA. More importantly, standardizing the cfNA detection process, specifying sample collection and storage conditions, cfNA extraction methods, and library construction methods, and developing data analysis processes will greatly promote the clinical application of cfNA. Third, tumors are heterogeneous, and normal cells are also heterogeneous to some extent, but the degree of mutation is relatively small. These functional mutations that are difficult to recognize can lead to false negatives and false positives [129]. False negatives and false positive test results in the study will have a greater error rate in people with a large base [130]. In the face of such challenges, well-established analytical methods are the key to solving problems. Focusing on intra-gene mutations alone will not completely solve the problem. Changes in gene copy number and epigenetic modifications are strongly associated with the occurrence of diseases. The analysis of multi-omics is the trend of the future, such as many methods in the article have found breakthroughs through omics. With the continuous development of information technology, it is possible to comprehensively monitor the occurrence, progression, treatment and recurrence of diseases from all dimensions using genomics, transcriptomics, proteomics, metabolomics.

At present, the dynamics of cancer cfDNA turnover is yet largely unknown [131]. Through the analysis of cfNA, finding changes in upstream genes related to disease is the first step for researchers. Whether it is genetic mutations, epigenetic modifications, or omics studies, the aim is to find more correlations between disease cells and normal cells. There are still prospects and potential for exploration in the discovery of new mutations as detection criteria and therapeutic targets, exploration of the distribution of the cell-specific degree of epigenetic modification, and exploration of tumor microenvironment.

Acknowledgements This work was supported by the National Natural Science Foundation of China (92153303, 21721005, 91940000 to X. Z), and the National Science Foundation for Young Scientists of China (21907078 to S.P.).

Conflict of interest The authors declare no conflict of interest.

1 Heitzer E, Haque IS, Roberts CES, Speicher MR. *Nat Rev Genet*, 2019, 20: 71–88
2 Gardner L, Kostarelos K, Mallick P, Dive C, Hadjidemetriou M. *Nat Rev Clin Oncol*, 2022, 19: 551–561
3 Underwood JJ, Quadri RS, Kalva SP, Shah H, Sanjeeviah AR, Beg MS, Sutphin PD. *Radiology*, 2020, 294: 5–17
4 Schwarzenbach H, Hoon DSB, Pantel K. *Nat Rev Cancer*, 2011, 11: 426–437
5 Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, Knippers R. *Cancer Res*, 2001, 61: 1659–1665
6 Tsui NB, Ng EK, Lo YD. *Clin Chem*, 2002, 48: 1647–1653
7 Fleischhacker M, Schmidt B. *Biochim Biophys Acta Rev Cancer*, 2007, 1775: 181–232
8 Zaporozhchenko IA, Ponomaryova AA, Rykova EY, Laktionov PP. *Expert Rev Mol Diagnostics*, 2018, 18: 133–145
9 Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R, Le Quesne J, Moore DA, Veeriah S, Rosenthal R, Marafioti T, Kirkizlar E, Watkins TBK, McGranahan N, Ward S, Martinson L, Riley J, Fraioli F, Al Bakir M, Grönroos E, Zambrana F, Endozo R, Bi WL, Fennessy FM, Sponer N, Johnson D, Laycock J, Shafi S, Czyzewska-Khan J, Rowan A, Chambers T, Matthews N, Turajlic S, Hiley C, Lee SM, Forster MD, Ahmad T, Falzon M, Borg E, Lawrence D, Hayward M, Kolvekar S, Panagiotopoulos N, Janes SM, Thakrar R, Ahmed A, Blackhall F, Summers Y, Hafez D, Naik A, Ganguly A, Kareht S, Shah R, Joseph L, Marie Quinn A, Crosbie PA, Naidu B, Middleton G, Langman G, Trotter S, Nicolson M, Remmen H, Kerr K, Chetty M, Gomersall L, Fennell DA, Nakas A, Rathinam S, Anand G, Khan S, Russell P, Ezhil V, Ismail B, Irvin-Sellers M, Prakash V, Lester JF, Kornaszewska M, Attanoos R, Adams H, Davies H, Oukrif D, Akarca AU, Hartley JA, Lowe HL, Lock S, Iles N, Bell H, Ngai Y, Elgar G, Szallasi Z, Schwarz RF, Herrero J, Stewart A, Quezada SA, Peggs KS, Van Loo P, Dive C, Lin CJ, Rabinowitz M, Aerts HJWL, Hackshaw A, Shaw JA, Zimmermann BG, Swanton C. *Nature*, 2017, 545: 446–451
10 Sadeh R, Sharkia I, Fialkoff G, Rahat A, Gutin J, Chappleboim A, Nitzan M, Fox-Fisher I, Neiman D, Meler G, Kamari Z, Yaish D, Peretz T, Hubert A, Cohen JE, Salah A, Temper M, Grinshpun A, Maoz M, Abu-Gazala S, Ben Ya'acov A, Shteyer E, Safadi R, Kaplan T, Shemer R, Planer D, Galun E, Glaser B, Zick A, Dor Y, Friedman N. *Nat Biotechnol*, 2021, 39: 586–598
11 Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, Jahn SW, Abete L, Pristauz G, Petru E, Geigl JB, Heitzer E, Speicher MR. *Nat Genet*, 2016, 48: 1273–1278
12 Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Liu MC, Oxnard GR, Klein EA, Smith D, Richards D, Yeatman TJ, Cohn AL,

Lapham R, Clement J, Parker AS, Tummala MK, McIntyre K, Sekeres MA, Bryce AH, Siegel R, Wang X, Cosgrove DP, Abu-Rustum NR, Trent J, Thiel DD, Becerra C, Agrawal M, Garbo LE, Giguere JK, Michels RM, Harris RP, Richey SL, McCarthy TA, Waterhouse DM, Couch FJ, Wilks ST, Krie AK, Balaraman R, Restrepo A, Meshad MW, Rieger-Christ K, Sullivan T, Lee CM, Greenwald DR, Oh W, Tsao CK, Fleshner N, Kennecke HF, Khalil MF, Spigel DR, Manhas AP, Ulrich BK, Kovoor PA, Stokoe C, Courtright JG, Yimer HA, Larson TG, Swanton C, Seiden MV, Cummings SR, Absalan F, Alexander G, Allen B, Amini H, Aravanis AM, Bagaria S, Bazargan L, Beausang JF, Berman J, Betts C, Blocker A, Bredno J, Calef R, Cann G, Carter J, Chang C, Chawla H, Chen X, Chien TC, Civello D, Davydov K, Demas V, Desai M, Dong Z, Fayzullina S, Fields AP, Filippova D, Freese P, Fung ET, Gnerre S, Gross S, Halks-Miller M, Hall MP, Hartman AR, Hou C, Hubbell E, Hunkapiller N, Jagadeesh K, Jamshidi A, Jiang R, Jung B, Kim TH, Klausner RD, Kurtzman KN, Lee M, Lin W, Lipson J, Liu H, Liu Q, Lopatin M, Maddala T, Maher MC, Melton C, Mich A, Nautiyal S, Newman J, Newman J, Nicula V, Nicolaou C, Nikolic O, Pan W, Patel S, Prins SA, Rava R, Ronaghi N, Sakarya O, Satya RV, Schellenberger J, Scott E, Sehnert AJ, Shaknovich R, Shanmugam A, Shashidhar KC, Shen L, Shenoy A, Shojaee S, Singh P, Steffen KK, Tang S, Toung JM, Valouev A, Venn O, Williams RT, Wu T, Xu HH, Yakym C, Yang X, Yecies J, Yip AS, Youngren J, Yue J, Zhang J, Zhang L, Zhang LQ, Zhang N, Curtis C, Berry DA. *Ann Oncol*, 2020, 31: 745–759

13 Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, Jensen SØ, Medina JE, Hruban C, White JR, Palsgrove DN, Niknafs N, Anagnostou V, Forde P, Naidoo J, Marrone K, Brahmer J, Woodward BD, Husain H, van Rooijen KL, Ørntoft MBW, Madsen AH, van de Velde CJH, Verheij M, Cats A, Punt CJA, Vink GR, van Grieken NCT, Koopman M, Fijneman RJA, Johansen JS, Nielsen HJ, Meijer GA, Andersen CL, Scharpf RB, Velculescu VE. *Nature*, 2019, 570: 385–389

14 Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, Pacey S, Baird R, Rosenfeld N. *Nat Rev Cancer*, 2017, 17: 223–238

15 Schwarzenbach H, Nishida N, Calin GA, Pantel K. *Nat Rev Clin Oncol*, 2014, 11: 145–156

16 Zhang Y, Yao Y, Xu Y, Li L, Gong Y, Zhang K, Zhang M, Guan Y, Chang L, Xia X, Li L, Jia S, Zeng Q. *Nat Commun*, 2021, 12: 1048

17 Chen X, Gole J, Gore A, He Q, Lu M, Min J, Yuan Z, Yang X, Jiang Y, Zhang T, Suo C, Li X, Cheng L, Zhang Z, Niu H, Li Z, Xie Z, Shi H, Zhang X, Fan M, Wang X, Yang Y, Dang J, McConnell C, Zhang J, Wang J, Yu S, Ye W, Gao Y, Zhang K, Liu R, Jin L. *Nat Commun*, 2020, 11: 3475

18 McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, Chin SF, Farooq M, Mejia R, Cronin PA, Anderson KS, Kosiorek HE, Northfelt DW, McCullough AE, Patel BK, Weitzel JN, Slavin TP, Caldas C, Pockaj BA, Murtaza M. *Sci Transl Med*, 2019, 11: eaax7392

19 Bratman SV, Yang SYC, Iafolla MAJ, Liu Z, Hansen AR, Bedard PL, Lheureux S, Spreafico A, Razak AA, Shchegrova S, Louie M, Billings P, Zimmermann B, Sethi H, Aleshin A, Torti D, Marsh K, Eagles J, Cirlan I, Hanna Y, Clouthier DL, Lien SC, Ohashi PS, Xu W, Siu LL, Pugh TJ. *Nat Cancer*, 2020, 1: 873–881

20 Leal A, van Grieken NCT, Palsgrove DN, Phallen J, Medina JE, Hruban C, Broeckaert MAM, Anagnostou V, Adleff V, Bruhm DC, Canzoniero JV, Fiksel J, Nordsmark M, Warmerdam FARM, Verheul HMW, van Spronsen DJ, Beerepoot LV, Geenen MM, Portielje JEA, Jansen EPM, van Sandick J, Meershoek-Klein Kranenbarg E, van Laarhoven HWM, van der Peet DL, van de Velde CJH, Verheij M, Fijneman R, Scharpf RB, Meijer GA, Cats A, Velculescu VE. *Nat Commun*, 2020, 11: 525

21 Yeh YM, Lin PC, Lee CT, Chen SH, Lin BW, Lin SC, Chen PC, Chan RH, Shen MR. *Mol Cancer*, 2020, 19: 150

22 Wan JCM, Heider K, Gale D, Murphy S, Fisher E, Mouliere F, Ruiz-Valdepenas A, Santonja A, Morris J, Chandrananda D, Marshall A, Gill AB, Chan PY, Barker E, Young G, Cooper WN, Hudecova I, Marass F, Mair R, Brindle KM, Stewart GD, Abraham JE, Caldas C, Rassl DM, Rintoul RC, Alifrangis C, Middleton MR, Gallagher FA, Parkinson C, Durrani A, McDermott U, Smith CG, Massie C, Corrie PG, Rosenfeld N. *Sci Transl Med*, 2020, 12: eaaz8084

23 Kingston B, Cutts RJ, Bye H, Beaney M, Walsh-Crestani G, Hrebien S, Swift C, Kilburn LS, Kernaghan S, Moretti L, Wilkinson K, Wardley AM, Macpherson IR, Baird RD, Roylance R, Reis-Filho JS, Hubank M, Faull I, Banks KC, Lanman RB, Garcia-Murillas I, Bliss JM, Ring A, Turner NC. *Nat Commun*, 2021, 12: 4479

24 Martignano F, Munagala U, Crucitta S, Mingrino A, Semeraro R, Del Re M, Petrini I, Magi A, Conticello SG. *Mol Cancer*, 2021, 20: 32

25 Shen SY, Burgener JM, Bratman SV, De Carvalho DD. *Nat Protoc*, 2019, 14: 2749–2780

26 Kurtz DM, Soo J, Co Ting Keh L, Alig S, Chabon JJ, Sworder BJ, Schultz A, Jin MC, Scherer F, Garofalo A, Macaulay CW, Hamilton EG, Chen B, Olsen M, Schroers-Martin JG, Craig AFM, Moding EJ, Esfahani MS, Liu CL, Dührsen U, Hüttmann A, Casasnovas RO, Westin JR, Roschewski M, Wilson WH, Gaidano G, Rossi D, Diehn M, Alizadeh AA. *Nat Biotechnol*, 2021, 39: 1537–1547

27 De Vlaminck I, Valantine HA, Snyder TM, Strehl C, Cohen G, Luikart H, Neff NF, Okamoto J, Bernstein D, Weisshaar D, Quake SR, Khush KK. *Sci Transl Med*, 2014, 6: 241ra277

28 Blauwkamp TA, Thair S, Rosen MJ, Blair L, Lindner MS, Vilfan ID, Kawli T, Christians FC, Venkatasubrahmanyam S, Wall GD, Cheung A, Rogers ZN, Meshulam-Simon G, Huijse L, Balakrishnan S, Quinn JV, Hollemon D, Hong DK, Vaughn ML, Kertesz M, Bercovici S, Wilber JC, Yang S. *Nat Microbiol*, 2019, 4: 663–674

29 Caggiano C, Celona B, Garton F, Mefford J, Black BL, Henderson R, Lomen-Hoerth C, Dahl A, Zaitlen N. *Nat Commun*, 2021, 12: 2717

30 Fedyuk V, Erez N, Furth N, Beresh O, Andreishcheva E, Shinde A, Jones D, Zakai BB, Mavor Y, Peretz T, Hubert A, Cohen JE, Salah A, Temper M, Grinshpun A, Maoz M, Zick A, Ron G, Shema E. *Nat Biotechnol*, 2023, 41: 212–221

31 Wang P, Song Q, Ren J, Zhang W, Wang Y, Zhou L, Wang D, Chen K, Jiang L, Zhang B, Chen W, Qu C, Zhao H, Jiao Y. *Sci Transl Med*, 2022, 14: abp8704

32 Guler GD, Ning Y, Ku CJ, Phillips T, McCarthy E, Ellison CK, Bergamaschi A, Collin F, Lloyd P, Scott A, Antoine M, Wang W, Chau K, Ashworth A, Quake SR, Levy S. *Nat Commun*, 2020, 11: 5270

33 Lo YMD, Han DSC, Jiang P, Chiu RWK. *Science*, 2021, 372: eaaw3616

34 Yu SCY, Jiang P, Peng W, Cheng SH, Cheung YTT, Tse OYO, Shang H, Poon LC, Leung TY, Chan KCA, Chiu RWK, Lo YMD. *Proc Natl Acad Sci USA*, 2021, 118: e2114937118

35 Sin STK, Jiang P, Deng J, Ji L, Cheng SH, Dutta A, Leung TY, Chan KCA, Chiu RWK, Lo YMD. *Proc Natl Acad Sci USA*, 2020, 117: 1658–1665

36 Larson MH, Pan W, Kim HJ, Mauntz RE, Stuart SM, Pimentel M, Zhou Y, Knudsgaard P, Demas V, Aravanis AM, Jamshidi A. *Nat Commun*, 2021, 12: 2357

37 Ibarra A, Zhuang J, Zhao Y, Salathia NS, Huang V, Acosta AD, Aballi J, Toden S, Karns AP, Purnajo I, Parks JR, Guo L, Mason J, Sigal D, Nova TS, Quake SR, Nerenberg M. *Nat Commun*, 2020, 11: 400

38 Ngo TTM, Moufarrej MN, Rasmussen MLH, Camunas-Soler J, Pan W, Okamoto J, Neff NF, Liu K, Wong RJ, Downes K, Tibshirani R, Shaw GM, Skotte L, Stevenson DK, Biggio JR, Elovitz MA, Melbye M, Quake SR. *Science*, 2018, 360: 1133–1136

39 Rasmussen M, Reddy M, Nolan R, Camunas-Soler J, Khodursky A, Scheller NM, Cantonwine DE, Engelbrechtsen L, Mi JD, Dutta A, Brundage T, Siddiqui F, Thao M, Gee EPS, La J, Baruch-Gravett C, Santillan MK, Deb S, Ame SM, Ali SM, Adkins M, DePristo MA, Lee M, Namsaraev E, Gybel-Brask DJ, Skibsted L, Litch JA, Santillan DA, Sazawal S, Tribe RM, Roberts JM, Jain M, Høgdall E, Holzman C, Quake SR, Elovitz MA, McElrath TF. *Nature*, 2022,

601: 422–427
40 Toden S, Zhuang J, Acosta AD, Karns AP, Salathia NS, Brewer JB, Wilcock DM, Aballi J, Nerenberg M, Quake SR, Ibarra A. *Sci Adv*, 2020, 6: eabb1654
41 Miyoshi J, Zhu Z, Luo A, Toden S, Zhou X, Izumi D, Kanda M, Takayama T, Parker IM, Wang M, Gao F, Zaidi AH, Baba H, Kodera Y, Cui Y, Wang X, Liu Z, Goel A. *Mol Cancer*, 2022, 21: 44
42 Rose Brannon A, Jayakumaran G, Diosdado M, Patel J, Razumova A, Hu Y, Meng F, Haque M, Sadowska J, Murphy BJ, Baldi T, Johnson I, Ptashkin R, Hasan M, Srinivasan P, Rema AB, Rijo I, Agarunov A, Won H, Perera D, Brown DN, Samoila A, Jing X, Gedvilaite E, Yang JL, Stephens DP, Dix JM, DeGroat N, Nafa K, Syed A, Li A, Lebow ES, Bowman AS, Ferguson DC, Liu Y, Mata DA, Sharma R, Yang SR, Bale T, Benhamida JK, Chang JC, Dogan S, Hameed MR, Hechtman JF, Moung C, Ross DS, Vakiani E, Vanderbilt CM, Yao JJ, Razavi P, Smyth LM, Chandarlapaty S, Iyer G, Abida W, Harding JJ, Krantz B, O'Reilly E, Yu HA, Li BT, Rudin CM, Diaz L, Solit DB, Arcila ME, Ladanyi M, Loomis B, Tsui D, Berger MF, Zehir A, Benayed R. *Nat Commun*, 2021, 12: 3770
43 Vandekerkhove G, Lavoie JM, Annala M, Murtha AJ, Sundahl N, Walz S, Sano T, Taavitsainen S, Ritch E, Fazli L, Hurtado-Coll A, Wang G, Nykter M, Black PC, Todenhöfer T, Ost P, Gibb EA, Chi KN, Eigl BJ, Wyatt AW. *Nat Commun*, 2021, 12: 184
44 Vermeirssen V, Deleu J, Morlion A, Everaert C, De Wilde J, Anckaert J, Durinck K, Nuytens J, Rishfi M, Speleman F, Van Droogenbroeck H, Verniers K, Baietti MF, Albersen M, Leucci E, Post E, Best MG, Van Maerken T, De Wilde B, Vandesompele J, Decock A. *NAR Cancer*, 2022, 4: zcac037
45 Oshi M, Murthy V, Takahashi H, Huyser M, Okano M, Tokumaru Y, Rashid OM, Matsuyama R, Endo I, Takabe K. *Cancers*, 2021, 13: 2652
46 Bulterys MA, Wagner B, Redard-Jacot M, Suresh A, Pollock NR, Moreau E, Denkinger CM, Drain PK, Broger T. *J Clin Med*, 2019, 9: 111
47 Yeasmin S, Ammanath G, Onder A, Yan E, Yildiz UH, Palaniappan A, Liedberg B. *TrAC Trends Anal Chem*, 2022, 157: 116786
48 Palmer BF, Clegg DJ. *Clin J Am Soc Nephrol*, 2019, 14: 306–316
49 Salfer B, Li F, Wong DTW, Zhang L. *Clin Chem*, 2022, 68: 1493–1501
50 Bryzgunova OE, Laktionov PP. *Acta Naturae*, 2015, 7: 48–54
51 Oreskovic A, Brault ND, Panpradist N, Lai JJ, Lutz BR. *J Mol Diagnostics*, 2019, 21: 1067–1078
52 Greytak SR, Engel KB, Parpart-Li S, Murtaza M, Bronkhorst AJ, Pertile MD, Moore HM. *Clin Cancer Res*, 2020, 26: 3104–3109
53 Johansson G, Andersson D, Filges S, Li J, Muth A, Godfrey TE, Ståhlberg A. *Biomol Detection Quantif*, 2019, 17: 100078
54 Werner B, Warton K, Ford CE. *Cancers*, 2022, 14: 1309
55 Botezatu I, Serdyuk O', Potapova G, Shelepov V, Alechina R, Molyaka Y, Anan'ev V, Bazin I, Garin A, Narimanov M, Knysh V, Melkonyan H, Umansky S, Lichtenstein A. *Clin Chem*, 2000, 46: 1078–1084
56 Kopreski MS, Benko FA, Kwak LW, Gocke CD. *Clin Cancer Res*, 1999, 5: 1961–1965
57 Chen XQ, Bonnefoi H, Pelte MF, Lyautey J, Lederrey C, Movarekhi S, Schaeffer P, Mulcahy HE, Meyer P, Stroun M, Anker P. *Clin Cancer Res*, 2000, 6: 3823–3826
58 Takahashi H, Yasui T, Baba Y. *Chem Lett*, 2021, 50: 1244–1253
59 van der Vaart M, Pretorius PJ. *Ann New York Acad Sci*, 2008, 1137: 18–26
60 Vo JN, Cieslik M, Zhang Y, Shukla S, Xiao L, Zhang Y, Wu YM, Dhanasekaran SM, Engelke CG, Cao X, Robinson DR, Nesvizhskii AI, Chinnaiyan AM. *Cell*, 2019, 176: 869–881.e13
61 Fleischhacker M. *Int J Mol Sci*, 2016, 1–4
62 Giri B, Seamon M, Banerjee A, Chauhan S, Purohit S, Morgan J, Baban B, Wakade C. *Metab Brain Dis*, 2022, 37: 1687–1696
63 Shekhtman EM, Anne K, Melkonyan HS, Robbins DJ, Warsof SL, Umansky SR. *Clin Chem*, 2009, 55: 723–729
64 Kim WT, Kim YH, Jeong P, Seo SP, Kang HW, Kim YJ, Yun SJ, Lee SC, Moon SK, Choi YH, Lee GT, Kim IY, Kim WJ. *Oncotarget*, 2018, 9: 14354–14365
65 Urquidi V, Netherton M, Gomes-Giacoia E, Serie D, Eckel-Passow J, Rosser CJ, Goodison S. *Oncotarget*, 2016, 7: 38731–38740
66 Piao XM, Jeong P, Kim YH, Byun YJ, Xu Y, Kang HW, Ha YS, Kim WT, Lee JY, Woo SH, Kwon TG, Kim IY, Moon SK, Choi YH, Cha EJ, Yun SJ, Kim WJ. *Int J Cancer*, 2019, 144: 380–388
67 Guo J, Yang J, Zhang X, Feng X, Zhang H, Chen L, Johnson H, Persson JL, Xiao K. *Anticancer Res*, 2018, 38: 1471–1477
68 Woo HK, Park J, Ku JY, Lee CH, Sunkara V, Ha HK, Cho YK. *Lab Chip*, 2019, 19: 87–97
69 Ngamdee T, Chalermwatanachai T, Siriwan C, Warachit O, Rijiravanich P, Surareungchai W. *Anal Bioanal Chem*, 2022, 414: 5695–5707
70 Saha S, Araf Y, Promon SK. *J Egypt Natl Canc Inst*, 2022, 34: 8
71 Zmrzljak UP, Košir R, Krivokapić Z, Radojković D, Nikolić A. *Genes*, 2021, 12: 289
72 Ståhlberg A, Krzyzanowski PM, Jackson JB, Egyud M, Stein L, Godfrey TE. *Nucleic Acids Res*, 2016, 44: e105
73 Ståhlberg A, Krzyzanowski PM, Egyud M, Filges S, Stein L, Godfrey TE. *Nat Protoc*, 2017, 12: 664–682
74 Lei R, Huo R, Mohan C. *Expert Rev Mol Diagnostics*, 2020, 20: 69–84
75 Jordaens S, Zwaenepoel K, Tjalma W, Deben C, Beyers K, Vankerckhoven V, Pauwels P, Vorsters A. *Intl J Cancer*, 2023, 152: 2186–2205
76 Bryzgunova OE, Skvortsova TE, Kolesnikova EV, Starikov AV, Rykova EY, Vlassov VV, Laktionov PP. *Ann New York Acad Sci*, 2006, 1075: 334–340
77 Chen C, Lu C, Qian Y, Li H, Tan Y, Cai L, Weng H. *Sci Rep*, 2017, 7: 17737
78 Sidransky D, Von Eschenbach A, Tsai Y, Jones P, Summerhayes I, Marshall F, Paul M, Green P, Hamilton P, Frost P, et A. *Science*, 1991, 252: 706–709
79 Hessels D, Klein Gunnewiek JMT, van Oort I, Karthaus HFM, van Leenders GJL, van Balken B, Kiemeney LA, Witjes JA, Schalken JA. *Eur Urol*, 2003, 44: 8–16
80 Groskopf J, Aubin SM, Deras IL, Blase A, Bodrug S, Clark C, Brentano S, Mathis J, Pham J, Meyer T, Cass M, Hodge P, Macairan ML, Marks LS, Rittenhouse H. *Clin Chem*, 2006, 52: 1089–1095
81 Lin SY, Dhillon V, Jain S, Chang TT, Hu CT, Lin YJ, Chen SH, Chang KC, Song W, Yu L, Block TM, Su YH. *J Mol Diagn*, 2011, 13: 474–484
82 Su YH, Wang M, Brenner DE, Norton PA, Block TM. *Ann New York Acad Sci*, 2008, 1137: 197–206
83 Chen S, Zhao J, Cui L, Liu Y. *Clin Transl Oncol*, 2017, 19: 332–340
84 Zhang J, Zhang X, Shen S. *Future Oncol*, 2020, 16: 849–858
85 Nuzzo PV, Berchuck JE, Korthauer K, Spisak S, Nassar AH, Abou Alaiwi S, Chakravarthy A, Shen SY, Bakouny Z, Boccardo F, Steinharter J, Bouchard G, Curran CR, Pan W, Baca SC, Seo JH, Lee GSM, Michaelson MD, Chang SL, Waikar SS, Sonpavde G, Irizarry RA, Pomerantz M, De Carvalho DD, Choueiri TK, Freedman ML. *Nat Med*, 2020, 26: 1041–1043
86 Bach S, Paulis I, Sluiter NR, Tibbesma M, Martin I, van de Wiel MA, Tuynman JB, Bahce I, Kazemier G, Steenbergen RDM. *Sci Rep*, 2021, 11: 2363
87 Zhou Z, Cheng SH, Ding SC, Heung MMS, Xie T, Cheng THT, Lam WKJ, Peng W, Teoh JYC, Chiu PKF, Ng CF, Jiang P, Chan KCA, Chiu RWK, Lo YMD. *Clin Chem*, 2021, 67: 621–630
88 Kim WT, Jeong P, Yan C, Kim YH, Lee IS, Kang HW, Kim YJ, Lee SC, Kim SJ, Kim YT, Moon SK, Choi YH, Kim IY, Yun SJ, Kim WJ. *Oncotarget*, 2016, 7: 58193–58202
89 Bartel DP. *Cell*, 2004, 116: 281–297
90 Mall C, Rocke DM, Durbin-Johnson B, Weiss RH. *Biomarkers Med*, 2013, 7: 623–631
91 Hung PS, Chen CY, Chen WT, Kuo CY, Lo SS. *PLoS One*, 2017, 12:

e0177346
92 Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermuller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H, Gingeras TR. *Science*, 2007, 316: 1484–1488
93 Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, Poliakov A, Cao X, Dhanasekaran SM, Wu YM, Robinson DR, Beer DG, Feng FY, Iyer HK, Chinnaiyan AM. *Nat Genet*, 2015, 47: 199–208
94 Bussemakers MJG, van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HF, Schalken JA, Debruyne FM, Ru N, Isaacs WB. *Cancer Res*, 1999, 59: 5975–5979
95 Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. *Nature*, 2013, 495: 384–388
96 Guarnerio J, Bezzi M, Jeong JC, Paffenholz SV, Berry K, Naldini MM, Lo-Coco F, Tay Y, Beck AH, Pandolfi PP. *Cell*, 2016, 165: 289–302
97 Zhang S, Zeng X, Ding T, Guo L, Li Y, Ou S, Yuan H. *Sci Rep*, 2018, 8: 2878
98 Kristiansen TZ, Maitra A, Pandey A. Proteomics of human bile. In: *Proteomics of Human Body Fluids*. Totowa: Humana Press, 2007. 399–414
99 Housset C, Chrétien Y, Debray D, Chignard N. *Compr Physiol*, 2016, 6: 1549–1577
100 Farina A, Dumonceau JM, Lescuyer P. *Expert Rev Proteomics*, 2009, 6: 285–301
101 Esteller A. *World J Gastroenterol*, 2008, 14: 5641–5649
102 Driescher C, Fuchs K, Haeberle L, Goering W, Frohn L, Opitz FV, Haeussinger D, Knoefel WT, Keitel V, Esposito I. *Cancers*, 2020, 13: 39
103 Li Z, Liu Y, Fu J, Mugaanyi J, Yan J, Lu C, Huang J. *Front Oncol*, 2022, 12: 961939
104 Han JY, Ahn KS, Kim TS, Kim YH, Cho KB, Shin DW, Baek WK, Suh SI, Jang BC, Kang KJ. *Cancers*, 2021, 13: 4581
105 Kinugasa H, Nouso K, Ako S, Dohi C, Matsushita H, Matsumoto K, Kato H, Okada H. *Cancer Biol Ther*, 2018, 19: 934–938
106 Shen N, Zhu B, Zhang W, Nian B, Xu X, Yu L, Ruan X, Chen S, Liu Y, Cao X, Shi X, Li Z, Huang X, Wang X, Chen C, Xiong L, Zhang D, Fu X, Zhang Y. *Front Oncol*, 2022, 12: 891917
107 Shen N, Zhang D, Yin L, Qiu Y, Liu J, Yu W, Fu X, Zhu B, Xu X, Duan A, Chen Z, Wang X, Cao X, Zhao T, Zhou Z, Yu L, Qin H, Fang Z, Li J, Liu Y, Xiong L, Yuan B, Li F, Zhang Y. *Oncol Rep*, 2019, doi: 10.3892/or.2019.7177
108 Arechederra M, Rullán M, Amat I, Oyon D, Zabalza L, Elizalde M, Latasa MU, Mercado MR, Ruiz-Clavijo D, Saldaña C, Fernández-Urién I, Carrascosa J, Jusué V, Guerrero-Setas D, Zazpe C, González-Borja I, Sangro B, Herranz JM, Purroy A, Gil I, Nelson LJ, Vila JJ, Krawczyk M, Zieniewicz K, Patkowski W, Milkiewicz P, Cubero FJ, Alkorta-Aranburu G, G Fernandez-Barrena M, Urman JM, Berasain C, Avila MA. *Gut*, 2022, 71: 1141–1151
109 Roest HP, IJzermans JNM, van der Laan LJW. *BMC Biotechnol*, 2021, 21: 48
110 Valle JW, Kelley RK, Nervi B, Oh DY, Zhu AX. *Lancet*, 2021, 397: 428–444
111 Banales JM, Cardinale V, Carpino G, Marzioni M, Andersen JB, Invernizzi P, Lind GE, Folseraas T, Forbes SJ, Fouassier L, Geier A, Calvisi DF, Mertens JC, Trauner M, Benedetti A, Maroni L, Vaquero J, Macias RIR, Raggi C, Perugorria MJ, Gaudio E, Boberg KM, Marin JJG, Alvaro D. *Nat Rev Gastroenterol Hepatol*, 2016, 13: 261–280
112 Wang YF, Feng FL, Zhao XH, Ye ZX, Zeng HP, Li Z, Jiang XQ, Peng ZH. *World J Gastroenterol*, 2014, 20: 4085–4092
113 Gou Q, Zhang CZ, Sun ZH, Wu LG, Chen Y, Mo ZQ, Mai QC, He J, Zhou ZX, Shi F, Cui W, Zou W, Lv L, Zhuang WH, Xu RD, Li WK, Zhang J, Du HW, Xiang JX, Wang HZ, Hou T, Li ST, Li Y, Chen XM, Zhou ZJ. *ESMO Open*, 2021, 6: 100275
114 Klump B, Hsieh CJ, Dette S, Holzmann K, Kieβlich R, Jung M, Sinn U, Ortner M, Porschen R, Gregor M. *Clin Cancer Res*, 2003, 9: 1773–1778
115 Shin SH, Lee K, Kim BH, Cho NY, Jang JY, Kim YT, Kim D, Jang JJ, Kang GH. *J Mol Diagn*, 2012, 14: 256–263
116 Vedeld HM, Grimsrud MM, Andresen K, Pharo HD, von Seth E, Karlsen TH, Honne H, Paulsen V, Färkkilä MA, Bergquist A, Jeanmougin M, Aabakken L, Boberg KM, Folseraas T, Lind GE. *Hepatology*, 2022, 75: 59–73
117 He S, Zeng F, Yin H, Wang P, Bai Y, Song Q, Chu J, Huang Z, Liu Y, Liu H, Chen Q, Liu L, Zhou J, Hu H, Li X, Li T, Wang G, Cai J, Jiao Y, Zhao H. *eClinicalMedicine*, 2023, 55: 101736
118 Shigehara K, Yokomuro S, Ishibashi O, Mizuguchi Y, Arima Y, Kawahigashi Y, Kanda T, Akagi I, Tajiri T, Yoshida H. *PLoS One*, 2011, 6: e23584
119 Macias RIR, Banales JM, Sangro B, Muntané J, Avila MA, Lozano E, Perugorria MJ, Padillo FJ, Bujanda L, Marin JJG. *Biochim Biophys Acta Mol Basis Dis*, 2018, 1864: 1468–1477
120 Voigtlander T, Gupta SK, Thum S, Fendrich J, Manns MP, Lankisch TO, Thum T. *PLoS One*, 2015, 10: e0139305
121 Han HS, Kim MJ, Han JH, Yun J, Kim HK, Yang Y, Kim KB, Park SM. *Hepatob Pancreat Dis Int*, 2020, 19: 41–50
122 Yoshida M, Yukawa H, Hayashi K, Naitoh I, Miyabe K, Hori Y, Natsume M, Jinno N, Kato A, Kachi K, Asano G, Sahashi H, Toyohara T, Kuno K, Kito Y, Kondo H, Hirano A, Okumura F, Anbe K, Baba Y, Kataoka H, Tanaka Y. *Cancer Sci*, 2023, 114: 295–305
123 Uchida N, Tsutsui K, Kobara H, Ezaki T, Fukuma H, Yachida M, Masaki T, Watanabe S, Nakatsu T, Nakano H, Kuriyama S. *Endoscopy*, 2003, 35: 185–188
124 Nagai K, Kuwatani M, Hirata K, Suda G, Hirata H, Takishin Y, Furukawa R, Kishi K, Yonemura H, Nozawa S, Sugiura R, Kawakubo K, Sakamoto N. *Diagnostics*, 2022, 12: 2704
125 De Mattos-Arruda L, Mayor R, Ng CKY, Weigelt B, Martínez-Ricarte F, Torrejon D, Oliveira M, Arias A, Raventos C, Tang J, Guerini-Rocco E, Martínez-Sáez E, Lois S, Marín O, de la Cruz X, Piscuoglio S, Towers R, Vivancos A, Peg V, Ramon y Cajal S, Carles J, Rodon J, González-Cao M, Tabernero J, Felip E, Sahuquillo J, Berger MF, Cortes J, Reis-Filho JS, Seoane J. *Nat Commun*, 2015, 6: 8839
126 Miller AM, Shah RH, Pentsova EI, Pourmaleki M, Briggs S, Distefano N, Zheng Y, Skakodub A, Mehta SA, Campos C, Hsieh WY, Selcuklu SD, Ling L, Meng F, Jing X, Samoila A, Bale TA, Tsui DWY, Grommes C, Viale A, Souweidane MM, Tabar V, Brennan CW, Reiner AS, Rosenblum M, Panageas KS, DeAngelis LM, Young RJ, Berger MF, Mellinghoff IK. *Nature*, 2019, 565: 654–658
127 Escudero L, Llort A, Arias A, Diaz-Navarro A, Martínez-Ricarte F, Rubio-Perez C, Mayor R, Caratù G, Martínez-Sáez E, Vázquez-Méndez É, Lesende-Rodríguez I, Hladun R, Gros L, Ramón y Cajal S, Poca MA, Puente XS, Sahuquillo J, Gallego S, Seoane J. *Nat Commun*, 2020, 11: 5376
128 Brunet-Vega A, Pericay C, Quílez ME, Ramírez-Lázaro MJ, Calvet X, Lario S. *Anal Biochem*, 2015, 488: 28–35
129 Reiter JG, Baretti M, Gerold JM, Makohon-Moore AP, Daud A, Iacobuzio-Donahue CA, Azad NS, Kinzler KW, Nowak MA, Vogelstein B. *Nat Rev Cancer*, 2019, 19: 639–650
130 Ignatiadis M, Sledge GW, Jeffrey SS. *Nat Rev Clin Oncol*, 2021, 18: 297–312
131 Li W, Zhang X, Lu X, You L, Song Y, Luo Z, Zhang J, Nie J, Zheng W, Xu D, Wang Y, Dong Y, Yu S, Hong J, Shi J, Hao H, Luo F, Hua L, Wang P, Qian X, Yuan F, Wei L, Cui M, Zhang T, Liao Q, Dai M, Liu Z, Chen G, Meckel K, Adhikari S, Jia G, Bissonnette MB, Zhang X, Zhao Y, Zhang W, He C, Liu J. *Cell Res*, 2017, 27: 1243–1257