ANNUAL REVIEWS

Annual Review of Pathology: Mechanisms of Disease

Detection and Diagnostic Utilization of Cellular and Cell-Free Tumor DNA

Jonathan C. Dudley[1] and Maximilian Diehn[2]

[1]Ludwig Center, Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21287, USA

[2]Department of Radiation Oncology, Stanford Cancer Institute, and Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, California 94305, USA; email: diehn@stanford.edu

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Annu. Rev. Pathol. Mech. Dis. 2021. 16:199–222

First published as a Review in Advance on November 23, 2020

The *Annual Review of Pathology: Mechanisms of Disease* is online at pathol.annualreviews.org

https://doi.org/10.1146/annurev-pathmechdis-012419-032604

Copyright © 2021 by Annual Reviews.
All rights reserved

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

cell-free DNA, cfDNA, circulating tumor DNA, ctDNA, molecular genetic pathology, oncology, next-generation sequencing, benign clonal expansions

Abstract

Because cancer is caused by an accumulation of genetic mutations, mutant DNA released by tumors can be used as a highly specific biomarker for cancer. Although this principle was described decades ago, the advent and falling costs of next-generation sequencing have made the use of tumor DNA as a biomarker increasingly practical. This review surveys the use of cellular and cell-free DNA for the detection of cancer, with a focus on recent technological developments and applications to solid tumors. It covers (*a*) key principles and technology enabling the highly sensitive detection of tumor DNA; (*b*) assessment of tumor DNA in plasma, including for genotyping, minimal residual disease detection, and early detection of localized cancer; (*c*) detection of tumor DNA in body cavity fluids, such as urine or cerebrospinal fluid; and (*d*) challenges posed to the use of tumor DNA as a biomarker by the phenomenon of benign clonal expansions.

Personalis EX2166

INTRODUCTION

DNA is released from tumors into surrounding fluids through apoptosis, necrosis, or active secretion. DNA nucleases digest DNA between nucleosomal positions, resulting in fragmentation to a length of approximately 170 base pairs (1). The concept of using tumor-specific mutations as biomarkers for cancer was first described in 1991 through the detection of *TP53* mutations in the urinary cells of patients with bladder cancer, with follow-up studies in 1992 describing *KRAS* mutations in the stool of patients with colorectal cancer (CRC) and, in 1994, mutations in both genes in the sputum of patients who developed lung cancer (2–4).

In 1997, fetal-specific DNA sequences were identified in maternal plasma (5). Follow-up studies demonstrated the utility of cell-free fetal DNA for the detection of chromosomal abnormalities, and it is now routinely used for the noninvasive prenatal diagnosis of trisomies and determination of gender (6, 7). While some studies have demonstrated the possibility of using cell-free fetal DNA to screen for single-gene disorders, the background noise and cost of whole-exome or genome sequencing have made this application more challenging (8–10). Targeted testing to assess for predefined single-gene mutations may be more technically feasible (11). A final major application of cell-free DNA (cfDNA) is the detection of transplant rejection from the presence of increased, graft-specific DNA sequences in the recipient bloodstream (12, 13).

This review examines the use of cfDNA as a biomarker for solid tumors, with a focus on cfDNA in the plasma and both cellular DNA and cfDNA in body cavity fluids. Interest in cell-free tumor DNA has exploded in recent years, and several excellent reviews have covered the topic in depth (1, 14–16). This review therefore focuses on the most significant and recent developments in this rapidly growing field, rather than attempting to exhaustively survey historic contributions. Particular attention is paid to technological and conceptual milestones in detecting cell-free tumor DNA, the use of circulating tumor DNA (ctDNA) to detect minimal residual disease and screen for localized cancers, analysis of tumor DNA in body cavity fluids, and challenges posed to the field by recently described benign clonal expansions.

CONCEPTS AND TECHNOLOGY

Single-Locus Assays

Initial examinations of the diagnostic potential of tumor DNA focused on single loci and employed PCR-based target enrichment strategies (**Table 1**). Mutant amplicons were detected through oligomer-specific hybridization, mutant allele–specific PCR, or other PCR-based strategies (2, 3, 17, 18). More recently, the Roche Cobas *EGFR* real-time quantitative PCR (qPCR) assay became the first US Food and Drug Administration–approved test for tumor-specific mutations in cfDNA. This assay detects mutations that predict the response to tyrosine-kinase inhibitors, as well as resistance mutations (19). An advantage of PCR-based tests is their quick turnaround time and low cost. With the exception of mutant allele–specific PCR, however, they typically have a limit of detection above 0.5%, limiting their applicability to the assessment of advanced cancers.

A conceptual advance in the analysis of cfDNA came with the development of digital PCR in 1999 (20). A challenge in employing regular PCR-based methods for the detection of rare variants is that mutations occurring due to polymerase errors during PCR can be difficult to distinguish from rare variants present in the original sample. Digital PCR overcame this limitation by isolating the original DNA molecules prior to amplification. This was initially achieved by the dilution of DNA across a PCR plate at a concentration of 0.5 copies per well, ensuring that most wells only had one or zero starting DNA molecules (20). In current applications of digital PCR, DNA is typically diluted across emulsions, increasing throughput. The locus of interest is then amplified by

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

Table 1 Technology and applications for cell-free and cellular DNA analysis

Test type	Genomic space	Techniques	Variants detected	Example applications in the literature[a]	References
Single-locus	1–100 bp	Allele-specific qPCR Digital PCR UMI PCR (e.g., Safe-SeqS)	Point mutations Indels	Actionable point mutation detection in advanced cancer	19–29
Multilocus	≥100–500 kbp	Multiplexed PCR (e.g., TAM-Seq) Hybrid capture (e.g., CAPP-Seq)	Point mutations Indels Copy number variants Translocations	Tumor genotyping Minimal residual disease detection	29, 32–35
Genome-wide	Whole genome	Shallow whole-genome sequencing (± bisulfite conversion) Genome amplification (e.g., REALSeq)	Aneuploidy Fragment end profiles Methylation profiles	Early cancer detection	13, 43–45, 47–52

Abbreviations: bp, base pair; CAPP-Seq, cancer personalized profiling by deep sequencing; kbp, kilobase pair; qPCR, quantitative PCR; REALSeq, repetitive element aneuploidy sequencing; TAM-Seq, tagged-amplicon deep sequencing; UMI, unique molecular identifier.
[a]Only single-locus applications currently have strong evidence of clinical utility.

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

PCR, and the presence of mutant or wild-type DNA is detected by the use of fluorescently tagged DNA oligomers. In a variation of digital PCR called beads, emulsion, amplification, magnetics (BEAMing), the DNA is amplified on oligomer-coated magnetic beads, and fluorescently tagged mutant or wild-type oligomers are detected by flow cytometry (21).

Like qPCR, digital PCR can be performed quickly and at low cost. It is therefore routinely employed in some clinical settings for the analysis of actionable genetic variants, such as *EGFR* mutations in lung cancer or *KRAS* mutations in CRC (22, 23). Unlike qPCR, however, digital PCR has a very low limit of detection (0.01–0.1%) and has therefore also been used in some studies for the early detection of resistance mutations and the detection of minimal residual disease in solid tumors, with applications in lung and pancreatic cancers (24, 25). A disadvantage of both qPCR and digital PCR is that the potential mutation of interest has to be known ahead of time and specifically accounted for in the assay design.

Molecular Barcoding

Next-generation sequencing offers a way to screen for mutations in a broader target space. Loci are first amplified by PCR, and then the entire region within the amplicon can be sequenced, enabling the unbiased detection of any variants within the targeted space. For high-sensitivity applications, a unique molecular identifier (UMI) is incorporated into one of the primers in an initial round of PCR (26). Like digital PCR, this strategy, called Safe-SeqS in its initial description, enables one to distinguish variants arising due to polymerase errors during amplification from variants present in the original DNA molecule. Variants present in the original molecule should be present in all copies with the same UMI, whereas variants arising due to polymerase errors will usually only be present in a subset of molecules with the same UMI. An exception is if a polymerase error occurs in the first cycle of PCR, in which case it can be passed on to all copies with the same UMI, resulting in a false positive mutation.

Personalis EX2166

This challenge can be surmounted by attaching adapters to the double-stranded DNA that uniquely tag each molecule in the duplex, as well as the duplex itself. After adapter ligation, the strands are amplified separately, and target enrichment is typically accomplished through hybridization to oligonucleotide probes complementary to the regions of interest. Two types of error suppression can be performed using this approach. The first compares copies of each single strand of the original DNA duplex (i.e., it analyzes the strands individually). The second additionally requires a mutation to be present on both single strands of the original DNA duplex. False positive mutations are extremely rare with the latter approach because they would have to occur separately at the same spot on each strand of the duplex (27). This strategy therefore offers the lowest limit of detection of any available method, but it is limited by a low rate of recovery of both duplex strands (28, 29). Clinical studies employing this approach, therefore, have either begun with a massive quantity of DNA, beyond what is available from a normal blood draw, or integrated it with a strategy relying primarily on single-strand UMIs (29, 30).

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Multilocus Assays

PCR-based assays have limited multiplexability. With qPCR and digital PCR, it is difficult to assay for more than one mutation in a single reaction pool. When assaying for multiple mutations, therefore, the original sample is typically split across multiple separate reaction pools, a requirement that may pose challenges when cfDNA is limited. Single-locus next-generation sequencing strategies allow multiple targets to be assessed in the same reaction pool. Without technically complex strategies that require massive DNA input, however, they typically cannot be used to assess more than 1,000–2,000 base pairs of genomic space in a single reaction pool (31, 32).

Hybrid capture–based target enrichment strategies offer a way to overcome this limitation, enabling the enrichment and sequencing of tens of thousands to millions of base pairs in a single assay. These strategies begin with the ligation of adapters to the original DNA molecule and have the advantage of preserving unique fragment ends. In combination with UMIs in the adapters, these unique fragment ends can be used to identify PCR copies originating from the same molecule and may themselves provide additional diagnostic information, as discussed below (29, 33, 34). In some applications, the UMIs may be coupled with identifiers that allow the original DNA duplex to be uniquely tagged, enabling ultrasensitive duplex sequencing when both strands are recovered (29).

Hybrid capture assays that cover a large genomic space are often used in advanced cancer patients to genotype tumors from the DNA they shed into the plasma. In some instances, assays may be designed to detect copy number variations within the genomic space as well as translocations, indels, and single-nucleotide variants, providing a comprehensive genetic portrait of a patient's tumor from a noninvasive blood draw (29, 33–35).

An additional advantage of hybrid capture assays is that they may help surmount the challenge posed by stochastic sampling error. When tracking a single locus at a low mutant allele fraction, the probability that a fragment of mutant DNA will be absent from a single blood draw by chance may be high due to the low concentrations of ctDNA present in most cancer patients. However, the more loci one tracks, the higher the probability that at least a few mutant DNA molecules will be present in a given blood draw. This strategy, called cancer personalized profiling by deep sequencing (CAPP-Seq) in its initial implementation, enables the detection of ctDNA at levels from about 0.1% mutant allele fraction to as low as approximately 0.0001% (1, 29, 33). This may be especially helpful in the setting of minimal residual disease detection, as discussed below, when mutations in a patient's tumor are already known. In this context, the genomic space actually analyzed from a hybrid capture panel can be limited to the specific bases that were mutated in a

Personalis EX2166

patient's tumor, lowering the background noise of the panel. The identification of any one of those mutations may be sufficient to diagnose residual disease (36–38).

A challenge with hybrid capture approaches is that a larger genomic space provides more opportunities for artifactual mutations to arise. One strategy for reducing this problem is to first assess the pattern of stereotypic artifacts in a group of normal samples and informatically account for that background when assessing patient samples for the presence of oncogenic mutations (29, 39, 40). An additional limitation is that hybrid capture–based workflows typically have longer turnaround times, since the target enrichment process often takes days (instead of hours with PCR), and are more expensive, since they involve sequencing a larger genomic space. Finally, despite the advantages of overcoming stochastic sampling error, the larger genomic space provided by hybrid capture strategies may provide diminishing additional value in some contexts (41).

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Genome-Wide Assays

In the context of cancer detection, hybrid capture–based strategies typically rely on identifying the presence of cancer from the occurrence of known oncogenic changes within the panel space. However, a substantial fraction of tumors of any given cancer type may not harbor known oncogenic mutations, or they may harbor a small enough number of such mutations that they may be missed by stochastic sampling error, even with a large panel.

Genome-wide assays can potentially address this challenge. One class of assays relies on the detection of chromosomal abnormalities, or aneuploidy. Aneuploidy is an early event in neoplasia, present in over 90% of tumors of any given cancer type, and highly specific for cancer (42, 43). It can be assessed either via shallow whole-genome sequencing or through PCR-based strategies. One recently developed PCR-based strategy, called repetitive element aneuploidy sequencing (REALSeq), amplifies short interspersed nuclear elements throughout the genome with a single primer pair, enabling the identification of aneuploidy from lower DNA quantities and with less sequencing than is possible with whole-genome approaches (44). In an initial study, REALSeq was able to detect 71% of cases with variant allele fractions $\geq$0.5% ctDNA (44). The sensitivity of aneuploidy analysis could potentially be improved through the preanalytic or informatic enrichment of DNA sequences under 150 base pairs, which have been shown in multiple studies to contain a higher tumor fraction (45–47).

An alternative genome-wide strategy focuses on unique fragment end profiles. It is thought that cfDNA is released by apoptosis (among other mechanisms), a process involving the digestion of DNA between nucleosomes. Since the pattern of nucleosomal packaging differs between cancerous and normal cells, the unique fragment ends created by this digestion may be used to identify cancer, with potentially thousands of informative fragment ends spread throughout the genome (48). Additionally, since the pattern of nucleosomal packaging differs between different tissue types, this strategy may also be used to provide information on the tissue of origin of identified tumors (49, 50).

A final genome-wide strategy focuses on patterns of methylation. Since genome-wide methylation patterns vary between cancer and normal tissues and between different tissue types, this strategy may also be used not only to identify the presence of cancer but also to get insight into its tissue of origin. Methylation may be assessed through bisulfite conversion prior to sequencing or through targeted capture of methylated base positions (13, 51, 52).

Genome-wide fragment end and methylation patterns have less preestablished specificity for cancer than do oncogenic mutations or aneuploidy. Identifying patterns specific for cancer, therefore, relies on machine learning across large sample sets, a process that holds increased risk for overfitting and batch effects, particularly in studies without independent validation cohorts. While

Personalis EX2166

preliminary studies have shown the potential of these strategies (34, 41, 49, 51), independent follow-up studies with larger cohorts will be necessary to confirm their ability to achieve comparable or superior performance relative to mutation-based strategies.

Mutant Allele Enrichment

Various strategies have been developed to enrich for tumor DNA prior to the application of the above methods, a process often referred to as mutant allele enrichment. One class of strategies takes advantage of differential melting temperatures created by mismatches between mutant and wild-type DNA. Coamplification at lower denaturation temperature PCR selectively amplifies mutant molecules by exploiting the lower denaturation temperature of heteroduplexes that form between mutant strands and wild-type DNA (53). Differential strand separation at critical temperature relies on multiple rounds of hybridization to biotinylated oligonucleotide probes, with preferential melting and amplification of mutant molecules that mismatch reference sequence probes (54). A separate strategy involves the hybridization of short reference sequence probes to the DNA followed by digestion with a double-stranded DNA nuclease that is unable to digest a duplex with a mismatch (55). Another class of strategies exploits the base-pair resolution specificity of Cas9. In one study, Cas9 was used to selectively digest the wild-type allele of *KRAS* (56). In another, deactivated Cas9 was used to selectively bind a mutant allele prior to assessment by qPCR (57).

These methods have demonstrated tens-fold to hundreds-fold enrichment of tumor DNA. However, they typically are limited to the assessment of single loci or small numbers of loci, require additional steps that may result in a loss of mutant molecules, and can make it harder to quantify the tumor proportion in the original sample. As a result, their application in clinical settings has been limited.

TUMOR DNA IN PLASMA

Tumor Genotyping

Multiple studies have shown high concordance between mutations found in ctDNA and tumor biopsies, making it possible to genotype patient tumors from a blood draw. One study examined matched tumor biopsies and cfDNA from a series of 41 metastatic prostate and breast cancer patients with at least 10% tumor-derived cfDNA. Using exome sequencing, the authors found 88% concordance for clonal somatic mutations and 80% concordance for copy number alterations (58). These findings are comparable to those of studies using more focused panels (33, 34, 59). Concordance is typically higher for mutations that are clonal in the tumor than for those that are subclonal (34, 60). Mutations detected in cfDNA but not present in the tumor may represent subclonal tumor mutations not sampled in the matched tumor specimen, technical errors (as discussed above), or benign clonal proliferations (discussed below) (61). Many of the latter variants can be eliminated through paired sequencing of white blood cells (59).

Because cancer DNA is released from tumors throughout the body, it can better capture intratumor genetic heterogeneity than a single tissue biopsy. The extent of this heterogeneity at the time of diagnosis and its clinical importance continues to be investigated. For example, in the 2017 TRACERx study, authors sequenced 327 regions from 100 pretreatment early-stage tumors from non–small cell lung cancer (NSCLC) patients. Mutations in some driver genes (*EGFR*, *MET*, *BRAF*, and *TP53*) were almost always clonal, but subclonal driver gene mutations with evidence of occurrence late in tumor evolution were also observed in 75% of patients (including mutations

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

in *PIK3CA* and *NF1* and copy number alterations) (62). These results contrast with the results of a 2018 study of 76 pretreatment metastases from 20 patients with a variety of solid tumor types. The authors found that the vast majority of driver gene mutations were shared among metastases and that driver genes mutations that were not shared were unlikely to be functionally significant (63). Notably, tumor samples were sequenced to a median depth of 65× in the 2018 study (J. Reiter, personal communication) compared to 431× in the TRACERx study (62), potentially contributing to some of the differences in observations.

One arena in which tumor genotyping from cfDNA offers an advantage is in identifying subclonal mutations conferring resistance to targeted therapy. Although these may be missed by a single tumor biopsy in patients with a high burden of disease, multiple studies have documented their emergence over the course of targeted therapy using cfDNA assays (23, 25, 64, 65). Examples of clinical scenarios in which targeted cfDNA genotyping is often performed include documentation of the emergence of oncogenic *KRAS* variants in CRC patients undergoing treatment with targeted anti-*EGFR* antibodies such as cetuximab or of the *EGFR* T790M variant in NSCLC patients undergoing treatment with anti-*EGFR* tyrosine kinase inhibitors (14). However, given the potential risk of false negatives due to low ctDNA concentration, standard clinical practice usually involves reflexing to tissue biopsies in patients with negative plasma genotyping results (66).

In addition to detecting known resistance mutations, several studies have illustrated the potential of serial cfDNA genotyping to discover new mechanisms of resistance. An early study used serial exome sequencing of cfDNA from six patients with advanced ovarian, breast, and lung cancers to document the emergence of *PIK3CA* mutations during treatment with paclitaxel, *RB1* inactivation during treatment with cisplatin, and *MED1* inactivation during treatment with tamoxifen and trastuzumab (67). In a more recent example, a hybrid capture panel focused on genes mutated in NSCLC was used to analyze cfDNA from 43 patients over the course of treatment with the third-generation *EGFR* inhibitor rociletinib. The authors identified novel *EGFR* variants and *MET* amplification as the most common resistance mechanisms, with additional recurrent variants identified in *PIK3CA*, *ERBB2*, *KRAS*, and *RB1* (68). A similar study in patients treated with the *EGFR* inhibitor osimertinib identified *EGFR* C797S mutations as a common mechanism of treatment resistance to this drug (69).

Tumor Burden Assessment

In addition to documenting tumor genome evolution over the course of therapy, cfDNA also enables assessment of tumor burden and response to therapy. Multiple studies have shown a strong correlation between cfDNA levels and tumor burden, and some have shown that alterations in levels of ctDNA (which has a half-life in the bloodstream of roughly 30 min to 2 h, depending on the study) provide an earlier measure of treatment response than imaging or protein assays (70–73).

One early study in this area compared ctDNA, cancer antigen 15-3, and circulating tumor cells at sequential time points in 30 women with metastatic breast cancer undergoing chemotherapy. After whole-genome sequencing of the patients' tumors, personalized assays were designed to monitor ctDNA levels in each patient and compared to the other biomarkers. The study found that ctDNA levels demonstrated better correlation with tumor burden, and over a wider range. It also detected treatment response earlier than the other biomarkers (74). Multiple additional studies, using a variety of different methods, have shown correlation between ctDNA levels, tumor volume as assessed by imagining, and protein biomarkers, with ctDNA outperforming alternative methods for tracking tumor burden (33, 73).

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

After attempted definitive treatment for cancer with surgery or radiation, a common question is whether or not the patient was truly cured or if the patient has minimal residual disease. If the patient was not cured, he or she may benefit from further systemic therapy. This question can be difficult to answer by imaging, as inflammation or fibrosis can be hard to distinguish from recurrent tumor. Protein biomarkers, likewise, may not be available for a particular tumor type or may correlate poorly with patient outcome. However, ctDNA has shown significant promise in addressing this problem.

Using a form of digital PCR called BEAMing, a 2008 study of 18 patients with CRC demonstrated that ctDNA outperformed carcinoembryonic antigen (CEA) in predicting recurrence-free survival following surgery. Notably, patients with undetectable ctDNA had 100% recurrence-free survival. The authors also demonstrated correlation between ctDNA levels and patient treatment course, with levels decreasing following surgery and roughly tracking with findings on imaging and with CEA levels (73). A 2016 study from the same group examined serial ctDNA levels from 230 patients with stage II CRC. Using the single-locus next-generation sequencing strategy Safe-SeqS, the authors found that the presence of ctDNA after curative-intent surgery had a hazard ratio of 18 (75).

Similar findings have been noted in other solid tumor types. A 2015 study used personalized digital PCR assays to detect ctDNA in patients with early-stage breast cancer following neoadjuvant chemotherapy. The presence of ctDNA had a hazard ratio for recurrent disease of 25, with a median lead time over clinical relapse of 7.9 months (76). A 2017 study using CAPP-Seq targeted recurrently mutated regions in NSCLC. The authors examined cfDNA from 40 patients treated with curative-intent radiation for stage I–III NSCLC. They detected ctDNA at the first posttreatment time point in 94% of patients who developed recurrent disease, with a median lead time over radiographic recurrence of 5.2 months. By using a panel covering a median of six mutations per patient in NSCLC, this strategy avoided the need to design personalized assays for each patient. Additionally, the authors found that their strategy of tracking multiple mutations per patient resulted in an increased likelihood of detecting ctDNA, likely due to the reduced susceptibility of this approach to stochastic sampling error (38). Additional studies have confirmed the value of ctDNA analysis to detect minimal residual disease in NSCLC and other tumor types (24, 36, 37, 77–79).

Tumor Detection and Classification

Perhaps the most potentially impactful, but also the most challenging, application of ctDNA is for the early detection of cancer. Every available treatment modality for cancer—surgery, radiation, chemotherapy, immunotherapy—is more likely to lead to a cure if the tumor is detected early (66). But using cfDNA to screen for cancer faces two challenges: The test must have very high specificity and be able to provide information on the tumor's tissue of origin (80, 81).

Information on a tumor's tissue of origin is needed to effectively treat it, as almost all available treatment strategies are tailored to the tumor type. Localized tumors, in particular, are typically treated (and can often be cured) with either surgical resection or radiation. The need for high specificity can be illustrated by considering the prevalence of cancer in different populations. Globally, the prevalence of cancer among those between 15 and 49 years old is 0.64% (82). If one were to test for cancer within this age group with a test that could detect every type of cancer and had 70% sensitivity and 98% specificity, roughly four out of every five positive results would be a false positive, a positive predictive value of about 20%. In actual practice, most tests can only detect a subset of cancers, reducing the effective prevalence in the population tested and thereby reducing the positive predictive value of the test. The positive predictive value of the test can be

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

improved by either improving the test's specificity or testing a higher-risk population. The global prevalence of cancer among those aged 50–69 years old, for example, is 3.32%. Within this population, with the above hypothetical test, roughly one out of every two people with a positive result would have cancer, giving a positive predictive value of approximately 50% (82).

One initial example of a screening strategy using cfDNA examined Epstein-Barr virus (EBV) DNA in the plasma as a biomarker for nasopharyngeal carcinoma, with which EBV is causally associated. The authors applied a single-locus qPCR-based test to 20,174 middle-aged men in Southeast Asia, a demographic with an elevated prevalence of the disease, and increased the specificity of their test by requiring a second test to confirm initial positive results. With this strategy, the authors achieved a sensitivity of 97.1% and specificity of 98.6%. They detected nasopharyngeal carcinomas at an earlier stage with superior outcomes compared to a historic cohort, albeit with a positive predictive value of only 11% (83). A follow-up study using hybrid capture and next-generation sequencing found this positive predictive value could be improved to 19.6% by considering the length and abundance of detected EBV molecules, even while eliminating the requirement for a confirmatory test (84). A further study adding consideration of methylation differences to the test increased the positive predictive value to 35.1% (85).

Other studies have examined the possibility of using multilocus cfDNA assays targeting recurrently mutated driver genes to screen for many different solid tumor types. In one such study, authors designed an 81-kilobase hybrid capture panel and evaluated its performance in a workflow called targeted error correction sequencing across 44 healthy controls and 200 patients with early-stage colorectal, breast, lung, or ovarian cancers. They achieved sensitivities ranging from 59% to 71%. No mutations related to solid tumors were detected among the healthy controls, although the study was limited by the absence of independent control and validation cohorts and an inability to provide information on the tissue of origin of the tumor (34).

A separate study examined the performance of a 61-amplicon panel across localized tumors of eight different cancer types. A key distinction of this study was the addition of protein biomarkers in an attempt to increase sensitivity without degrading specificity as well as to potentially give insight into the tumor's tissue of origin. In theory, protein and ctDNA biomarkers could be complementary due to different mechanisms of biomarker release, with tumor DNA released primarily by dying cells and proteins released primarily by active secretion from living cells. Across 1,005 patients and 812 healthy controls, the authors' workflow (called CancerSEEK) achieved a median sensitivity of 70% with a specificity of 99%. Additionally, the test localized the tumor signal to one of two tissue of origin types in 83% of cases (41). When considered individually, the protein component of the assay outperformed the ctDNA component (44). Across 153 cases with a high-confidence ctDNA mutation and available tumor tissue, the authors found that the mutation observed in plasma was also present in the paired tumor in 90% of cases (41). Limitations of this study included the absence of an independent validation cohort, lack of matched leukocytes for exclusion of clonal hematopoiesis variants in the majority of cases positive for ctDNA, and the use of cases and controls not fully representative of the general population.

Other approaches have attempted to apply genome-wide strategies for cancer detection. These include the aforementioned REALSeq approach, which, when combined with the CancerSEEK assay, increased the latter's sensitivity to 80% (44). Alternative genome-wide strategies for early detection have focused on the unique fragment end profiles generated by digestion of DNA between nucleosomes prior to release, as noted above. One study showed that fragment end distribution patterns could be used to infer the cell types contributing to the pool of cfDNA (50). Another study demonstrated stereotypic fragment ends in cfDNA from hepatocellular carcinoma patients that distinguished them from patients with cirrhosis, those with hepatitis B, and healthy individuals. The authors demonstrated that fragment end profiles may point not only to the tissue of origin

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

of cfDNA but also to the presence or absence of cancer (48). A recent study further extended a related approach to other tumor types (49). However, each of these studies lacked prospectively collected, independent validation cohorts, and thus further work is needed to confirm key test parameters such as sensitivity and specificity.

A variation on this concept looked at differential read depth at transcription start sites to infer expressed genes. An initial iteration of this strategy demonstrated its ability to infer gene expression patterns within tumors from cfDNA, albeit only at very high mutant allele fractions (86). A follow-up study from the same group demonstrated the ability of a similar strategy to infer transcription factor binding across several cancers and applied the technique to detect early-stage colorectal carcinomas, although with relatively low specificity (87).

A final genome-wide strategy assessed methylation patterns of cfDNA. An early study employed genome-wide bisulfite sequencing of plasma cfDNA and inferred the contributing tissues by referencing the methylation profiles of a variety of tissue types. This study demonstrated the ability to infer cfDNA contributions from placental tissue, hepatocellular carcinoma, and transplanted tissue. Intriguingly, the authors were also able to use methylation patterns within cfDNA regions demonstrating aneuploidy to infer the tissue type that contributed the aneuploid region (13). This principle was extended by another group to infer cfDNA contributions from an even wider range of highly specific tissue types, including pancreatic beta cells, oligodendrocytes, neuronal and glial cells, and the exocrine pancreas (52). A more recent study employed anti-5-methylcytidine antibodies to immunoprecipitate methylated regions, thereby enabling more focused sequencing of differentially methylated regions than is possible by genome-wide bisulfite sequencing (51). As with the other approaches in this area, further validation studies will be required to robustly assess the potential of methylation-based approaches for early cancer detection.

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

TUMOR DNA IN BODY CAVITY FLUIDS

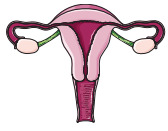
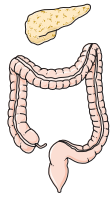
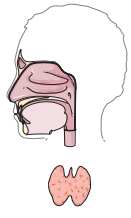
One strategy to surmount the challenges posed to using plasma cfDNA for early detection and other applications is to analyze lower-volume fluids in closer proximity to the tumor of interest. These fluids can include urine for genitourinary cancers, Pap smear fluid for gynecologic cancers, saliva for head and neck cancers, stool for gastrointestinal cancers, cerebrospinal fluid (CSF) for brain tumors and leptomeningeal metastases, and cyst fluid for the classification of pancreatic, thyroid, and ovarian cancers. Such fluids typically have higher mutant allele fractions for their associated tumors than plasma has, and many are easily accessed. Applications in each of these fluids are discussed below (**Table 2**).

Urine

The first demonstration of the concept of using tumor-specific point mutations as biomarkers for cancer detected *TP53* mutations in the cell pellet of patients with bladder cancer, using PCR followed by oligomer-specific hybridization (2). More than two decades later, the same group employed Safe-SeqS to demonstrate the ability of *TERT* promoter mutations in urinary cell pellets to serve as a biomarker for bladder cancer (88). Since the *TERT* promoter is mutated in 60–70% of bladder cancers at one of two hotspots that can be assessed in a single amplicon, it is one of the most valuable biomarkers for bladder cancer (60). A more recent study by the same group developed a multilocus, PCR-based panel (UroSEEK) including the *TERT* promoter region, regions from ten other genes (*CDKN2A*, *ERBB2*, *FGFR3*, *HRAS*, *KRAS*, *MET*, *MLL*, *PIK3CA*, *TP53*, and *VHL*), and a PCR-based test for aneuploidy. In a cohort of 570 patients undergoing diagnostic evaluation for bladder cancer, the test achieved a sensitivity of 83% (95% when combined with cytology) and

Personalis EX2166

Table 2 Tumor detection from body cavity fluids

Anatomic system[a]		Fluids	Cancers	References
Genitourinary		Urine	Urothelial Prostate Renal Systemic	2, 60, 88–96, 98, 103, 104
Gynecologic		Pap smear Uterine lavage Ovarian cyst	Ovarian Endometrial	30, 107–110
Gastrointestinal		Pancreatic cyst Ductal brushings Stool	Pancreatic Bile duct Colorectal	111–118
Head and neck		Saliva Thyroid cyst	Squamous cell Thyroid	119–124
Central nervous system		Cerebrospinal fluid	Brain Leptomeningeal metastases	126–137

[a]Images adapted from Servier Medical Art (CC BY 3.0 Unported).

a specificity of 93%. In a cohort of 322 patients undergoing surveillance following surgery, the test achieved a sensitivity of 68% (71% when combined with cytology) and a specificity of 80% (89).

While the above studies evaluated DNA extracted from the cell pellet, other recent studies have analyzed cfDNA in the supernatant. An advantage of this approach is that it does not compete with cytology for diagnostic material, as the supernatant of urine cytology specimens is discarded

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

as part of the clinical workflow. One recent study adapted the hybrid capture–based CAPP-Seq workflow to address the unique challenges posed by urinary cfDNA, including the extraction of cfDNA from large volumes of fluid and heterogeneous fragment lengths. The resulting protocol, called urine tumor DNA CAPP-Seq or uCAPP-Seq, was applied to urinary cfDNA from 67 healthy adults and 118 patients with early-stage, mostly nonmuscle invasive bladder cancer who were undergoing either diagnostic evaluation or surveillance for recurrent disease. Employing a 311-kilobase hybrid capture panel targeting recurrently mutated regions in bladder cancer, the study found *PLEKHS1* promoter mutations in 46% of cases and demonstrated their utility as a biomarker. It achieved approximately 84% sensitivity in both diagnostic and surveillance contexts while maintaining cytology's high specificity. Notably, the test found cell-free tumor DNA in the supernatant of every cancer case that cytology detected and in 82% of cases that cytology missed (60). An alternative strategy to analyzing urinary cfDNA for the detection of bladder cancer employed shallow-depth whole-genome sequencing on bisulfite converted DNA, analyzing both methylation and copy number aberrations. The authors evaluated urine from 46 bladder cancer patients and 39 controls with hematuria, achieving 93.5% sensitivity (84.2% for low-grade nonmuscle invasive disease) and 95.8% specificity (90).

Aside from practical considerations, whether cell pellets or cfDNA provides a better sample for detecting bladder cancer is unclear. Some studies have found higher mutant allele fractions in the urinary supernatant of bladder cancer patients while others have not found a significant difference between the compartments (91–93). However, either compartment provides a superior sample to that provided by plasma for the detection of bladder cancer (91). Preliminary studies have demonstrated that urinary DNA can also be used for the detection of prostate and kidney cancers, although whether it can achieve diagnostic performance similar to that for bladder cancer or provides a better substrate than plasma does for these tumor types remains unclear (92, 94).

One intriguing application of urinary cfDNA is for the detection of nongenitourinary tumors. Several studies have provided evidence that some circulating DNA from the bloodstream is excreted in the urine. This includes fetal DNA in pregnant women and tumor DNA in patients with cancer (95–100). Although it is often assumed that this occurs through glomerular filtration, microhematuria or extravasation and diffusion directly into the bladder cavity have not been ruled out as excretion mechanisms (101). An initial study suggested that this DNA, called transrenal DNA in some studies, was the same size as cfDNA in the plasma, but follow-up studies have shown that this DNA is in fact very short (102). One study of fetal DNA in maternal urine found peak DNA sizes ranging from 29 to 45 base pairs by next-generation sequencing, a finding congruent with other studies (95, 103). Aside from avoiding a venous blood draw, the main theoretical advantage of using urine for the detection of nongenitourinary cancers is the ability to obtain a much larger quantity of cfDNA. One study found a median of 3.6 ng of cfDNA per mL of urine in healthy controls, comparable to cfDNA concentrations in the plasma of healthy subjects (60). Whereas 2–20 mL of plasma is typically obtainable for cfDNA studies, however, several hundred milliliters of urine can be obtained in a single collection (60). The increased quantity of DNA available could enable much deeper sequencing, potentially leading to the detection of earlier-stage tumors.

The fact that more cfDNA is obtainable from urine, however, does not necessarily mean that more mutant molecules are present. If the mutant DNA from nongenitourinary tumors is far more dilute in urine than in plasma, the theoretical advantage of having more starting cfDNA could be nullified. One study using a next-generation sequencing–based assay for *KRAS* hotspot mutations found that advanced cancer patients with biopsy-confirmed *KRAS* mutations had more than an order of magnitude fewer mutant molecules in a median of 60 mL of urine than in 1.5–4 mL of plasma collected at the same time (104). These findings are congruent with the results of a separate study of NSCLC patients that found the mutant allele fraction of the *EGFR* T790M

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

mutation (as assessed by BEAMing) was roughly an order of magnitude lower in the urine than in paired plasma samples (96). Although some studies have shown that biopsy-proven mutations are occasionally detected in urine that were not found in paired plasma samples, this could be explained by stochastic sampling error (98). These findings cast doubt on the value of analyzing urinary cfDNA over plasma for the detection of mutations in nongenitourinary tumors.

Gynecologic Fluids

Pap smears are one of the best illustrations of the promise of reducing cancer mortality through early detection. Populations screened by Pap smears have seen a 75% decrease in mortality from cervical cancer since the test was introduced (105). The value of molecular testing in Pap smears is also well established, such that testing for the human papillomavirus (HPV) is now integrated into the standard of care for cervical cancer screening (106).

Building on the observation by cytopathologists that endometrial and ovarian cancer cells are occasionally seen in Pap smears, one group investigated whether molecular testing of Pap smears could be used to improve the detection of these tumor types. In a proof-of-concept study, the authors designed a multilocus Safe-SeqS PCR panel targeting the most commonly mutated genes in these tumor types and applied it to 24 endometrial and 22 ovarian tumor specimens, finding mutations in every tumor specimen. They then searched for those mutations in paired Pap smear fluid, finding at least one mutation from the associated tumor in all 24 of the endometrial cancer cases and 9 of the 22 ovarian cancer cases (107). In a follow-up study, the authors employed the same technology with an 18-gene panel and a PCR-based assay for aneuploidy to screen for endometrial and ovarian cancers, without prior knowledge of tumor mutation status. Across 382 endometrial and 245 ovarian cancer patients, the authors achieved sensitivities of 81% and 33%, respectively, with a specificity (across 714 women without cancer) of 98.6%. They also found that sensitivity for both cancers increased by about 10% by intrauterine sampling with a Tao brush and that the sensitivity for ovarian cancer increased to 63% when Pap smear and plasma testing were combined (108).

Other studies have extended the concept to detect gynecologic cancers from other fluids. In one study, authors sequenced DNA from uterine lavage fluid, finding mutations in 18 of 30 patients with ovarian cancer and 5 of 5 patients with endometrial cancer. However, mutations were also detected with this approach in 8 of 27 patients with benign gynecologic lesions, potentially reflecting benign clonal expansions (discussed below). Notably, *TP53* mutations were only found in cancer, whereas 6 of the 8 positive benign cases had a *KRAS* mutation. This finding suggests that in certain contexts, some mutations may have a higher specificity for cancer than others (109). Another study sequenced tumor DNA from ovarian cyst fluid, examining 10 nonneoplastic, 12 benign, 31 invasive-cancer, and 24 borderline cysts. Using a multilocus PCR-based panel, the authors found tumor-associated mutations in 87% of the latter two cyst types (which require surgical removal) but no mutations in the former two cyst types (which do not require surgery). Notably, the type of mutation present shed light on the nature of the cyst, with borderline tumors harboring only *KRAS* or *BRAF* mutations but type II cancers (typically, high-grade ovarian serous carcinomas) harboring almost exclusively *TP53* mutations (110).

A final study used the ultrasensitive duplex sequencing method to analyze DNA extracted from peritoneal fluid of 17 women with high-grade ovarian serous carcinoma and 20 women without cancer. Although the authors detected *TP53* mutations in 16 of the 17 women with cancer, they also detected *TP53* mutations (most at much lower median mutant allele fractions) in nearly all of the control patients. They were able to create a threshold that would achieve 82% sensitivity and 90% specificity across this group, but their results demonstrate a background rate of *TP53*

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

mutation in normal tissues that may set a lower limit on the gains that can be achieved in diagnostic performance with more sensitive mutation-based assays (30).

Gastrointestinal Fluids

As with ovarian cysts, analyzing pancreatic cyst fluid tumor DNA can provide information on the cyst subtype. There are four main types of neoplastic cysts of the pancreas: intraductal papillary mucinous neoplasms (IPMNs), mucinous cystic neoplasms (MCNs), solid pseudopapillary neoplasms (SPNs), and serous cystadenomas (SCAs). Accurately subtyping these cysts from biopsy material can spare patients surgery, as SCAs and low- and intermediate-grade IPMNs can be followed, whereas high-grade IPMNs, SPNs, and MCNs require surgical excision to prevent the development of invasive carcinoma. Several studies have shown that the presence of mutations in certain genes can provide insight into the type of cyst present. Mutations in *KRAS*, *GNAS*, and *RNF43*, for example, are very common in IPMNs but rarely present in SCAs or SPNs, which almost always have mutations affecting *VHL* and *CTNNB1*, respectively. MCNs may have mutations in *KRAS* but typically lack *GNAS* variants. This information can be combined with cytology, radiographic imaging, and analysis of protein levels (e.g., CEA, vascular endothelial growth factor A) for added diagnostic performance (111–114). One study showed that using an algorithm incorporating these features would have resulted in patient management superior to that of a standard-of-care diagnostic workup and would have prevented unnecessary surgery in 50% of patients in the study (111). In addition to its value in subtyping pancreatic cysts, the finding of mutant DNA in brushing specimens from strictures of the bile or pancreatic ducts may indicate that the stricture is caused by a neoplastic process (as opposed to inflammation) (115).

Other studies have examined the use of stool for the detection of CRC. One study analyzed stool from 25 patients with CRC using BEAMing. In 23 of 25 cases, the authors found mutations present in stool that matched a mutation present in the patient's tumor specimen. By contrast, tumor mutations were detected in paired plasma in only 50% of cases (116). Another study examined *KRAS* mutation status and the methylation of four genes in stool, integrating the information with stool hemoglobin results. The sensitivity of the test was directly proportional to tumor size, identifying 54% of patients with adenomas greater than or equal to 1 cm and 85% patients with CRC, with 90% specificity (117). A follow-up study demonstrated that the test significantly outperformed a commercially available test based on plasma cfDNA *SEPT9* methylation status (118).

Head and Neck Fluids

Testing of thyroid cyst fluid has shown promise for assessing cysts with indeterminate cytology following fine-needle aspiration. ThyroSeq is the name of an influential multilocus next-generation sequencing assay used to reduce surgery in patients with indeterminate cytological findings (119). It tests for mutations in genes commonly involved in thyroid cancer and provides information on the presence or absence of cancer and, if positive, the type of cancer present. The *BRAF* V600E variant or the *RET-PTC* translocation, for example, is indicative of papillary carcinoma, whereas mutations in *PTEN* or rearrangements of *PAX8-PPARγ* are found in follicular carcinoma (120). In its most recent iteration, ThyroSeq has a negative predictive of value of 97% (121).

Other studies have looked at the use of tumor DNA in saliva as a biomarker for head and neck squamous cell carcinoma. An early study using PCR followed by hybridization with radiolabeled oligonucleotides found *TP53* mutations in the saliva of 5 of 7 patients with head and neck squamous cell carcinoma prior to surgery (122). After the discovery of the causal role of HPV in this

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

cancer type, other studies examined the use of HPV in the saliva, plasma, or both as a biomarker. Two independent studies found that testing both saliva and plasma for HPV (and other genetic biomarkers as well in one study) resulted in increased sensitivity over testing either fluid alone. One of the studies found that oral cavity tumors are most readily detected from saliva, whereas tumors arising in the oropharynx, hypopharynx, or larynx are more easily detected from plasma. Both studies also found that the presence of tumor DNA in saliva posttreatment predicted recurrent disease (123, 124).

Multiple studies have also examined the use of mutant *KRAS* DNA, *TP53* DNA, or both in sputum for the diagnosis of lung cancer. Although an initial study found mutations in one or both genes in 8 of 10 patients who went on to develop lung cancer (4), follow-up studies have demonstrated lower sensitivities (13–46%) (125).

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Cerebrospinal Fluid

Proximal fluid sampling can be especially valuable when DNA is used to assess lesions arising in the brain. Such tumors are among the most difficult to detect in plasma, presumably due to the blood-brain barrier (64, 126). Several studies have documented the ability to detect tumor DNA from primary brain tumors and leptomeningeal metastases in CSF, with sensitivities ranging from 38% to 92% for primary brain tumors and 63% to 100% for metastatic disease (126–137). Methods used include digital PCR, Safe-SeqS, hybrid capture–based multilocus panels, and shallow whole-genome sequencing. In the case of metastatic disease, two studies have compared the sensitivity of testing the cellular pellet versus the supernatant for cancer, with both achieving higher sensitivity in the supernatant (134, 136). CSF sampling can enable the detection of clinically informative genetic mutations in tumors that cannot otherwise be detected without a potentially morbid craniotomy procedure to directly sample tissue. These include mutations that provide diagnostic and prognostic information, such as codeletion of chromosome arms 1p/19q (characteristic of oligodendroglioma), isocitrate dehydrogenase (*IHD*) mutation status (with mutations conferring a superior outcome), and *TERT* promoter mutations status (with mutations conferring an inferior outcome) (138). In the case of leptomeningeal disease, CSF sampling may provide insight into the development of resistance mutations to targeted therapy, such as *EGFR* variants in NSCLC or *PIK3CA* variants in breast cancer (136). Finally, such sampling may shed light on the nature of abnormal findings on imaging that could reflect neoplastic or inflammatory processes, with mutant DNA supporting the presence of neoplasia.

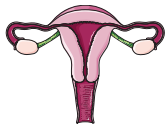
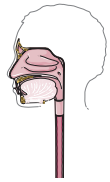
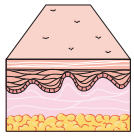
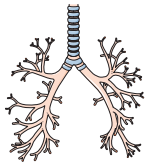
BENIGN CLONAL EXPANSIONS

The use of mutant DNA as a biomarker for cancer is based on the premise that because oncogenic mutations cause cancer rather than merely being associated with it, they should have exquisite specificity. This premise has recently been challenged by the discovery of oncogenic mutations in a variety of benign processes and normal tissues (139, 140) (**Table 3**).

Most notably, several studies have documented oncogenic mutations in the white blood cells of otherwise healthy individuals that increase in prevalence with age. This phenomenon is termed clonal hematopoiesis of indeterminate potential (CHIP) (141, 142). The mutations described predominate in genes associated with hematopoietic neoplasms, such as *DNMT3A*, *TET2*, and *ASXL1*, although they are also found in *TP53*. Initial studies based on low-depth, whole-exome sequencing data estimated the prevalence of CHIP to be roughly 3% among those aged 50–59 years old, 5% among those aged 60–69, and 10% among those aged 70–79 (142). A recent study using more targeted sequencing and UMIs for highly sensitive mutation detection,

Personalis EX2166

Table 3 Benign clonal expansions

Tissue type[a]		Mutated genes	References
White blood cells		*DNTM3A* *TET2* *ASXL1* *TP53*	141–143
Uterine cavity		*KRAS* *PIK3CA* *ARID1A*	109, 144, 145
Esophagus		*NOTCH1* *TP53*	147, 148, 150
Skin		*NOTCH1* *NOTCH2* *TP53*	146, 150
Bronchial epithelium		*NOTCH1* *TP53* *ARID2*	149

[a]Images adapted from Servier Medical Art (CC BY 3.0 Unported).

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

however, found mutations associated with CHIP in 95% of individuals aged 50–60 years old (143). CHIP is typically accounted for in blood-based liquid biopsy applications through paired sequencing of white blood cells. Mutations found in the white blood cells but not present (or present at a much lower variant allele fraction) in the plasma are likely due to CHIP (59). Other strategies for addressing CHIP for solid tumor applications include discounting mutations in genes associated primarily with hematopoietic neoplasms and using genome-wide biomarkers that may be less commonly present in CHIP (aneuploidy) or have unique patterns that can be distinguished from those associated with CHIP (fragment end and methylation profiles) (34, 44, 49).

Although CHIP presents the biggest challenge across liquid biopsy applications, benign clonal expansions in nonhematopoietic tissues may pose problems for the analysis of certain body cavity fluids. Two independent studies of uterine lavage fluid, for example, revealed oncogenic mutations

Personalis EX2166

in 8 of 27 and 51 of 95 of subjects without cancer, including mutations in *KRAS*, *PIK3CA*, and *PTEN* (109, 144). In the latter study, these mutations were significantly associated with increased patient age and postmenopausal status (144). Although the etiology of these mutations remains unclear, one study provides clues. The authors demonstrated the presence of driver mutations in deeply infiltrating endometriotic lesions. Across 39 lesions in 27 patients, the authors found driver gene mutations (including in *ARID1A*, *PIK3CA*, *KRAS*, and *PP2R1A*) in 10 lesions. All mutations were present in the epithelial component of the lesions but not the stromal component (145). It is conceivable that some of these mutations arose in normal endometrial epithelium and enabled it to acquire the cancer-like properties of deeply infiltrating endometriosis.

In fact, several recent studies have demonstrated a prevalent landscape of driver mutations in apparently normal tissues. One study sequenced 74 cancer genes in millimeter-scale biopsies from the normal-appearing eyelids of four individuals, finding mutations with evidence of positive selection in 18–32% of skin cells. These included mutations in driver genes associated with squamous cell carcinoma, including *NOTCH1*, *NOTCH2*, and *TP53* (146). Two other studies demonstrated a similar pattern in histologically normal esophageal epithelium, with a patchwork of clonal expansions carrying driver mutations in genes such as *NOTCH1* and *TP53*. These clonal expansions were associated with smoking, alcohol consumption, and age, covering nearly the entire esophageal mucosa in the elderly (147, 148). A final study demonstrated the presence of driver mutations in single cells grown as colonies from human bronchial epithelium, revealing recurrent mutations in driver genes *NOTCH1*, *TP53*, and *ARID2*. These mutations increased with age and were more prevalent in smokers than in ex-smokers and those who never smoked (149). While these studies illustrate microscopic clonal expansions or mutations in single cells, a recent study documented the presence of macroscopic expansions with clonal driver gene mutations across multiple normal tissue types. The authors evaluated 6,700 samples across 29 normal tissues and found the highest mutation burdens in sun-exposed skin, esophagus, and lung, reflecting the higher exposure of these tissue types to external mutagens (150).

These clonal expansions may reflect the earliest stage of cancer initiation. Due to their small size, they are unlikely to cause false positive results with most plasma-based tumor DNA assays. However, they may complicate tumor-detection strategies based on sequencing DNA from body cavity fluids or brushings specimens.

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

CONCLUSION

Despite the caveats introduced by benign clonal expansions, the studies surveyed in this review demonstrate that analysis of tumor DNA in plasma and body cavity fluids remains a very promising strategy to genotype, monitor, and detect cancer. Most of the studies surveyed here, however, are proof-of-concept studies. At present, there are few clinical scenarios in which cfDNA analysis has demonstrated clinical utility, and these are primarily in advanced cancers (14). The challenge for the field is to translate the promise of initial studies across a variety of applications into clinically validated assays with concrete evidence of improving patient care.

DISCLOSURE STATEMENT

J.C.D. has served as a consultant for ClearView Healthcare Partners and Sound Healthcare Communications. M.D. reports research funding from Varian Medical Systems and Illumina Inc.; ownership interest in CiberMed Inc. and Foresight Diagnostics; patent filings related to cancer biomarkers; paid consultancy from Roche, AstraZeneca, BioNTech, Novartis, Gritstone Oncology, and Illumina Inc.; and travel/honoraria from RefleXion Medical Inc.

Personalis EX2166

ACKNOWLEDGMENTS

J.C.D. is the Gordon Family Physician-Scientist of the Damon Runyon Cancer Research Foundation (PST-27-20). This work was supported by the National Cancer Institute (R01CA188298, R01CA244526, and R01CA254179 to M.D.), the US National Institutes of Health Director's New Innovator Award Program (1-DP2-CA186569 to M.D.), the Virginia and D.K. Ludwig Fund for Cancer Research (M.D.), the Bakewell Foundation (M.D.), and the CRK Faculty Scholar Fund (M.D.).

LITERATURE CITED

1. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, et al. 2017. Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat. Rev. Cancer* 17:223–38
2. Sidransky D, Von Eschenbach A, Tsai YC, Jones P, Summerhayes I, et al. 1991. Identification of p53 gene mutations in bladder cancers and urine samples. *Science* 252:706–9
3. Sidransky D, Tokino T, Hamilton SR, Kinzler KW, Levin B, et al. 1992. Identification of *ras* oncogene mutations in the stool of patients with curable colorectal tumors. *Science* 256:102–5
4. Mao L, Hruban RH, Boyle JO, Tockman M, Sidransky D. 1994. Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. *Cancer Res.* 54:1634–37
5. Lo YMD, Corbetta N, Chamberlain PF, Rai V, Sargent IL, et al. 1997. Presence of fetal DNA in maternal plasma and serum. *Lancet* 350:485–87
6. Bianchi DW, Chiu RWK. 2018. Sequencing of circulating cell-free DNA during pregnancy. *N. Engl. J. Med.* 379:464–73
7. Fan HC, Gu W, Wang J, Blumenfeld YJ, El-Sayed YY, Quake SR. 2012. Non-invasive prenatal measurement of the fetal genome. *Nature* 487:320–24
8. Chan KCA, Jiang P, Sun K, Cheng YKY, Tong YK, et al. 2016. Second generation noninvasive fetal genome analysis reveals de novo mutations, single-base parental inheritance, and preferred DNA ends. *PNAS* 113:E8159–68
9. Lo YMD, Chan KCA, Sun H, Chen EZ, Jiang P, et al. 2010. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci. Transl. Med.* 2:61ra91
10. Kitzman JO, Snyder MW, Ventura M, Lewis AP, Qiu R, et al. 2012. Noninvasive whole-genome sequencing of a human fetus. *Sci. Transl. Med.* 4:137ra76
11. Camunas-Soler J, Lee H, Hudgins L, Hintz SR, Blumenfeld YJ, et al. 2018. Noninvasive prenatal diagnosis of single-gene disorders by use of droplet digital PCR. *Clin. Chem.* 64:336–45
12. De Vlaminck I, Valantine HA, Snyder TM, Strehl C, Cohen G, et al. 2014. Circulating cell-free DNA enables noninvasive diagnosis of heart transplant rejection. *Sci. Transl. Med.* 6:241ra77
13. Sun K, Jiang P, Chan KCA, Wong J, Cheng YKY, et al. 2015. Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *PNAS* 112:E5503–12
14. Merker JD, Oxnard GR, Compton C, Diehn M, Hurley P, et al. 2018. Circulating tumor DNA analysis in patients with cancer: American Society of Clinical Oncology and College of American Pathologists joint review. *Arch. Pathol. Lab. Med.* 142:1242–53
15. Heitzer E, Haque IS, Roberts CES, Speicher MR. 2019. Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nat. Rev. Genet.* 20:71–88
16. Diaz LA Jr., Bardelli A. 2014. Liquid biopsies: genotyping circulating tumor DNA. *J. Clin. Oncol.* 32:579–86
17. Kopreski MS, Benko FA, Kwee C, Leitzel KE, Eskander E, et al. 1997. Detection of mutant K-*ras* DNA in plasma or serum of patients with colorectal cancer. *Br. J. Cancer* 76:1293–99
18. Sorenson GD, Pribish DM, Valone FH, Memoli VA, Bzik DJ, Yao SL. 1994. Soluble normal and mutated DNA sequences from single-copy genes in human blood. *Cancer Epidemiol. Biomark. Prev.* 3:67–71
19. Kwapisz D. 2017. The first liquid biopsy test approved. Is it a new era of mutation testing for non-small cell lung cancer? *Ann. Transl. Med.* 5:46
20. Vogelstein B, Kinzler KW. 1999. Digital PCR. *PNAS* 96:9236–41

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

21. Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B. 2003. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *PNAS* 100:8817–22
22. Taly V, Pekin D, Benhaim L, Kotsopoulos SK, Le Corre D, et al. 2013. Multiplex picodroplet digital PCR to detect *KRAS* mutations in circulating DNA from the plasma of colorectal cancer patients. *Clin. Chem.* 59:1722–31
23. Oxnard GR, Paweletz CP, Kuang Y, Mach SL, O'Connell A, et al. 2014. Noninvasive detection of response and resistance in *EGFR*-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. *Clin. Cancer Res.* 20:1698–705
24. Groot VP, Mosier S, Javed AA, Teinor JA, Gemenetzis G, et al. 2019. Circulating tumor DNA as a clinical test in resected pancreatic cancer. *Clin. Cancer Res.* 25:4973–84
25. Watanabe M, Kawaguchi T, Isa S, Ando M, Tamiya A, et al. 2015. Ultra-sensitive detection of the pretreatment EGFR T790M mutation in non–small cell lung cancer patients with an EGFR-activating mutation using droplet digital PCR. *Clin. Cancer Res.* 21:3552–60
26. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. 2011. Detection and quantification of rare mutations with massively parallel sequencing. *PNAS* 108:9530–35
27. Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, Loeb LA. 2012. Detection of ultra-rare mutations by next-generation sequencing. *PNAS* 109:14508–13
28. Schmitt MW, Fox EJ, Prindle MJ, Reid-Bayliss KS, True LD, et al. 2015. Sequencing small genomic targets with high efficiency and extreme accuracy. *Nat. Methods* 12:423–25
29. Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, et al. 2016. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat. Biotechnol.* 34:547–55
30. Krimmel JD, Schmitt MW, Harrell MI, Agnew KJ, Kennedy SR, et al. 2016. Ultra-deep sequencing detects ovarian cancer cells in peritoneal fluid and reveals somatic *TP53* mutations in noncancerous tissues. *PNAS* 113:6005–10
31. Tewhey R, Warner JB, Nakano M, Libby B, Medkova M, et al. 2009. Microdroplet-based PCR enrichment for large-scale targeted sequencing. *Nat. Biotechnol.* 27:1025–31
32. Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, et al. 2012. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci. Transl. Med.* 4:136ra68
33. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, et al. 2014. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* 20:548–54
34. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, et al. 2017. Direct detection of early-stage cancers using circulating tumor DNA. *Sci. Transl. Med.* 9:eaan2415
35. Lanman RB, Mortimer SA, Zill OA, Sebisanovic D, Lopez R, et al. 2015. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. *PLOS ONE* 10:e0140712
36. Moding EJL, Liu Y, Nabet BY, Chabon JJ, Chaudhuri AA, et al. 2020. Circulating tumor DNA dynamics predict benefit from consolidation immunotherapy in locally advanced non-small-cell lung cancer. *Nat. Cancer* 1:176–83
37. Azad TD, Chaudhuri AA, Fang P, Qiao Y, Esfahani MS, et al. 2019. Circulating tumor DNA analysis for detection of minimal residual disease after chemoradiotherapy for localized esophageal cancer. *Gastroenterology* 158:494–505.e6
38. Chaudhuri AA, Chabon JJ, Lovejoy AF, Newman AM, Stehr H, et al. 2017. Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling. *Cancer Discov*. 7:1394–403
39. Chen G, Mosier S, Gocke CD, Lin MT, Eshleman JR. 2014. Cytosine deamination is a major cause of baseline noise in next-generation sequencing. *Mol. Diagn. Ther.* 18:587–93
40. Gerstung M, Beisel C, Rechsteiner M, Wild P, Schraml P, et al. 2012. Reliable detection of subclonal single-nucleotide variants in tumour cell populations. *Nat. Commun.* 3:811
41. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, et al. 2018. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* 359:926–30
42. Knouse KAD, Davoli T, Elledge SJ, Amon A. 2017. Aneuploidy in cancer: seq-ing answers to old questions. *Annu. Rev. Cancer Biol.* 1:335–54

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

43. Douville C, Springer S, Kinde I, Cohen JD, Hruban RH, et al. 2018. Detection of aneuploidy in patients with cancer through amplification of long interspersed nucleotide elements (LINEs). *PNAS* 115:1871–76
44. Douville CC, Cohen JD, Ptak J, Popoli M, Schaefer J, et al. 2020. Assessing aneuploidy with repetitive element sequencing. *PNAS* 117:4858–63
45. Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, et al. 2018. Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci. Transl. Med.* 10:eaat4921
46. Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, et al. 2016. Fragment length of circulating tumor DNA. *PLOS Genet*. 12:e1006162
47. Jiang P, Chan CW, Chan KCA, Cheng SH, Wong J, et al. 2015. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *PNAS* 112:E1317–25
48. Jiang P, Sun K, Tong YK, Cheng SH, Cheng THT, et al. 2018. Preferred end coordinates and somatic variants as signatures of circulating tumor DNA associated with hepatocellular carcinoma. *PNAS* 115:E10925–33
49. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, et al. 2019. Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature* 570:385–89
50. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. 2016. Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. *Cell* 164:57–68
51. Shen SY, Singhania R, Fehringer G, Chakravarthy A, Roehrl MHA, et al. 2018. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* 563:579–83
52. Lehmann-Werman R, Neiman D, Zemmour H, Moss J, Magenheim J, et al. 2016. Identification of tissue-specific cell death using methylation patterns of circulating DNA. *PNAS* 113:E1826–34
53. Li J, Wang L, Mamon H, Kulke MH, Berbeco R, Makrigiorgos GM. 2008. Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nat. Med.* 14:579–84
54. Guha M, Castellanos-Rizaldos E, Liu P, Mamon H, Makrigiorgos GM. 2013. Differential strand separation at critical temperature: a minimally disruptive enrichment method for low-abundance unknown DNA mutations. *Nucleic Acids Res*. 41:e50
55. Song C, Liu Y, Fontana R, Makrigiorgos A, Mamon H, et al. 2016. Elimination of unaltered DNA in mixed clinical samples via nuclease-assisted minor-allele enrichment. *Nucleic Acids Res*. 44:e146
56. Gu W, Crawford ED, O'Donovan BD, Wilson MR, Chow ED, et al. 2016. Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol*. 17:41
57. Aalipour A, Dudley JC, Park SM, Murty S, Chabon JJ, et al. 2018. Deactivated CRISPR associated protein 9 for minor-allele enrichment in cell-free DNA. *Clin. Chem*. 64:307–16
58. Adalsteinsson VA, Ha G, Freeman SS, Choudhury AD, Stover DG, et al. 2017. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. *Nat. Commun.* 8:1324
59. Razavi P, Li BT, Brown DN, Jung B, Hubbell E, et al. 2019. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat. Med.* 25:1928–37
60. Dudley JC, Schroers-Martin J, Lazzareschi DV, Shi WY, Chen SB, et al. 2019. Detection and surveillance of bladder cancer using urine tumor DNA. *Cancer Discov*. 9:500–9
61. Stetson DA, Ahmed A, Xu X, Nuttall BRB, Lubinski TJ, et al. 2019. Orthogonal comparison of four plasma NGS tests with tumor suggests technical factors are a major source of assay discordance. *JCO Precis. Oncol.* 3. **https://doi.org/10.1200/PO.18.00191**
62. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, et al. 2017. Tracking the evolution of non-small-cell lung cancer. *N. Engl. J. Med.* 376:2109–21
63. Reiter JG, Makohon-Moore AP, Gerold JM, Heyde A, Attiyeh MA, et al. 2018. Minimal functional driver gene heterogeneity among untreated metastases. *Science* 361:1033–37
64. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, et al. 2014. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* 6:224ra24
65. Diaz LA Jr., Williams RT, Wu J, Kinde I, Hecht JR, et al. 2012. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 486:537–40

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

66. Mattox AK, Bettegowda C, Zhou S, Papadopoulos N, Kinzler KW, Vogelstein B. 2019. Applications of liquid biopsies for cancer. *Sci. Transl. Med.* 11:eaay1984
67. Murtaza M, Dawson SJ, Tsui DW, Gale D, Forshew T, et al. 2013. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* 497:108–12
68. Chabon JJ, Simmons AD, Lovejoy AF, Esfahani MS, Newman AM, et al. 2016. Circulating tumour DNA profiling reveals heterogeneity of EGFR inhibitor resistance mechanisms in lung cancer patients. *Nat. Commun.* 7:11815
69. Thress KS, Paweletz CP, Felip E, Cho BC, Stetson D, et al. 2015. Acquired *EGFR* C797S mutation mediates resistance to AZD9291 in non–small cell lung cancer harboring iEGFR T790M. *Nat. Med.* 21:560–62
70. Yao W, Mei C, Nan X, Hui L. 2016. Evaluation and comparison of in vitro degradation kinetics of DNA in serum, urine and saliva: a qualitative study. *Gene* 590:142–48
71. To EW, Chan KCA, Leung SF, Chan LY, To KF, et al. 2003. Rapid clearance of plasma Epstein-Barr virus DNA after surgical treatment of nasopharyngeal carcinoma. *Clin. Cancer Res.* 9:3254–59
72. Chen K, Zhao H, Shi Y, Yang F, Wang LT, et al. 2019. Perioperative dynamic changes in circulating tumor DNA in patients with lung cancer (DYNAMIC). *Clin. Cancer Res.* 25:7058–67
73. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, et al. 2008. Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* 14:985–90
74. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, et al. 2013. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N. Engl. J. Med.* 368:1199–209
75. Tie J, Wang Y, Tomasetti C, Li L, Springer S, et al. 2016. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci. Transl. Med.* 8:346ra92
76. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, et al. 2015. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci. Transl. Med.* 7:302ra133
77. Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, et al. 2017. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature* 545:446–51
78. Phallen J, Leal A, Woodward BD, Forde PM, Naidoo J, et al. 2019. Early noninvasive detection of response to targeted therapy in non–small cell lung cancer. *Cancer Res.* 79:1204–13
79. Kurtz DM, Esfahani MS, Scherer F, Soo J, Jin MC, et al. 2019. Dynamic risk profiling using serial tumor biomarkers for personalized outcome prediction. *Cell* 178:699–713.e19
80. Kalinich M, Haber DA. 2018. Cancer detection: seeking signals in blood. *Science* 359:866–67
81. Aravanis AM, Lee M, Klausner RD. 2017. Next-generation sequencing of circulating tumor DNA for early cancer detection. *Cell* 168:571–74
82. Inst. Health Metr. Eval. (IHME). 2018. *Global Burden of Disease Study 2017*. Seattle, WA: IHME
83. Chan KCA, Woo JKS, King A, Zee BCY, Lam WKJ, et al. 2017. Analysis of plasma Epstein-Barr virus DNA to screen for nasopharyngeal cancer. *N. Engl. J. Med.* 377:513–22
84. Lam WKJ, Jiang P, Chan KCA, Cheng SH, Zhang H, et al. 2018. Sequencing-based counting and size profiling of plasma Epstein-Barr virus DNA enhance population screening of nasopharyngeal carcinoma. *PNAS* 115:E5115–24
85. Lam WKJ, Jiang P, Chan KCA, Peng W, Shang H, et al. 2019. Methylation analysis of plasma DNA informs etiologies of Epstein-Barr virus-associated diseases. *Nat. Commun.* 10:3256
86. Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, et al. 2016. Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat. Genet.* 48:1273–78
87. Ulz P, Perakis S, Zhou Q, Moser T, Belic J, et al. 2019. Inference of transcription factor binding from cell-free DNA enables tumor subtype prediction and early detection. *Nat. Commun.* 10:4666
88. Kinde I, Munari E, Faraj SF, Hruban RH, Schoenberg M, et al. 2013. TERT promoter mutations occur early in urothelial neoplasia and are biomarkers of early disease and disease recurrence in urine. *Cancer Res.* 73:7162–67
89. Springer SU, Chen CH, Rodriguez Pena MDC, Li L, Douville C, et al. 2018. Non-invasive detection of urothelial cancer through the analysis of driver gene mutations and aneuploidy. *eLife* 7:e32143

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

90. Cheng THT, Jiang P, Teoh JYC, Heung MMS, Tam JCW, et al. 2019. Noninvasive detection of bladder cancer by shallow-depth genome-wide bisulfite sequencing of urinary cell-free DNA for methylation and copy number profiling. *Clin. Chem.* 65:927–36
91. Patel KM, van der Vos KE, Smith CG, Mouliere F, Tsui D, et al. 2017. Association of plasma and urinary mutant DNA with clinical outcomes in muscle invasive bladder cancer. *Sci. Rep.* 7:5554
92. Ge G, Peng D, Guan B, Zhou Y, Gong Y, et al. 2019. Urothelial carcinoma detection based on copy number profiles of urinary cell-free DNA by shallow whole-genome sequencing. *Clin. Chem.* 6:188–98
93. Togneri FS, Ward DG, Foster JM, Devall AJ, Wojtowicz P, et al. 2016. Genomic complexity of urothelial bladder cancer revealed in urinary cfDNA. *Eur. J. Hum. Genet.* 24:1167–74
94. Xia Y, Huang CC, Dittmar R, Du M, Wang Y, et al. 2016. Copy number variations in urine cell free DNA as biomarkers in advanced prostate cancer. *Oncotarget* 7:35818–31
95. Tsui NB, Jiang P, Chow KC, Su X, Leung TY, et al. 2012. High resolution size analysis of fetal DNA in the urine of pregnant women by paired-end massively parallel sequencing. *PLOS ONE* 7:e48319
96. Goldman JW, Karlovich C, Sequist LV, Melnikova V, Franovic A, et al. 2018. *EGFR* genotyping of matched urine, plasma, and tumor tissue in patients with non-small-cell lung cancer treated with rociletinib, an *EGFR* tyrosine kinase inhibitor. *JCO Precis. Oncol.* 2. **https://doi.org/10.1200/PO.17.00116**
97. Lu T, Li J. 2017. Clinical applications of urinary cell-free DNA in cancer: current insights and promising future. *Am. J. Cancer Res.* 7:2318–32
98. Reckamp KL, Melnikova VO, Karlovich C, Sequist LV, Camidge DR, et al. 2016. A highly sensitive and quantitative test platform for detection of NSCLC *EGFR* mutations in urine and plasma. *J. Thorac. Oncol.* 11:1690–700
99. Umansky SR, Tomei LD. 2006. Transrenal DNA testing: progress and perspectives. *Expert Rev. Mol. Diagn.* 6:153–63
100. Su YH, Wang M, Block TM, Landt O, Botezatu I, et al. 2004. Transrenal DNA as a diagnostic tool: important technical notes. *Ann. N.Y. Acad. Sci.* 1022:81–89
101. Botezatu I, Serdyuk O, Potapova G, Shelepov V, Alechina R, et al. 2000. Genetic analysis of DNA excreted in urine: a new approach for detecting specific genomic DNA sequences from cells dying in an organism. *Clin. Chem.* 46:1078–84
102. Su YH, Wang M, Brenner DE, Ng A, Melkonyan H, et al. 2004. Human urine contains small, 150 to 250 nucleotide-sized, soluble DNA derived from the circulation and may be useful in the detection of colorectal cancer. *J. Mol. Diagn.* 6:101–7
103. Shekhtman EM, Anne K, Melkonyan HS, Robbins DJ, Warsof SL, Umansky SR. 2009. Optimization of transrenal DNA analysis: detection of fetal DNA in maternal urine. *Clin. Chem.* 55:723–29
104. Fujii T, Barzi A, Sartore-Bianchi A, Cassingena A, Siravegna G, et al. 2017. Mutation-enrichment next-generation sequencing for quantitative detection of *KRAS* mutations in urine cell-free DNA from patients with advanced cancers. *Clin. Cancer Res.* 23:3657–66
105. Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, et al. 2010. European Guidelines for quality assurance in cervical cancer screening. Second edition—summary document. *Ann. Oncol.* 21:448–58
106. Cibas ES, Ducatman BS. 2014. *Cytology: Diagnostic Principles and Clinical Correlates*. Philadelphia: Elsevier
107. Kinde I, Bettegowda C, Wang Y, Wu J, Agrawal N, et al. 2013. Evaluation of DNA from the Papanicolaou test to detect ovarian and endometrial cancers. *Sci. Transl. Med.* 5:167ra4
108. Wang Y, Li L, Douville C, Cohen JD, Yen TT, et al. 2018. Evaluation of liquid from the Papanicolaou test and other liquid biopsies for the detection of endometrial and ovarian cancers. *Sci. Transl. Med.* 10:eaap8793
109. Maritschnegg E, Wang Y, Pecha N, Horvat R, Van Nieuwenhuysen E, et al. 2015. Lavage of the uterine cavity for molecular detection of Müllerian duct carcinomas: a proof-of-concept study. *J. Clin. Oncol.* 33:4293–300
110. Wang Y, Sundfeldt K, Mateoiu C, Shih M, Kurman RJ, et al. 2016. Diagnostic potential of tumor DNA from ovarian cyst fluid. *eLife* 5:e15175
111. Springer S, Masica DL, Dal Molin M, Douville C, Thoburn CJ, et al. 2019. A multimodality test to guide the management of patients with a pancreatic cyst. *Sci. Transl. Med.* 11:eaav4772
112. Springer S, Wang Y, Dal Molin M, Masica DL, Jiao Y, et al. 2015. A combination of molecular markers and clinical features improve the classification of pancreatic cysts. *Gastroenterology* 149:1501–10

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

113. Rosenbaum MW, Jones M, Dudley JC, Le LP, Iafrate AJ, Pitman MB. 2017. Next-generation sequencing adds value to the preoperative diagnosis of pancreatic cysts. *Cancer Cytopathol.* 125:41–47
114. Jones M, Zheng Z, Wang J, Dudley J, Albanese E, et al. 2016. Impact of next-generation sequencing on the clinical diagnosis of pancreatic cysts. *Gastrointest. Endosc.* 83:140–48
115. Dudley JC, Zheng Z, McDonald T, Le LP, Dias-Santagata D, et al. 2016. Next-generation sequencing and fluorescence in situ hybridization have comparable performance characteristics in the analysis of pancreaticobiliary brushings for malignancy. *J. Mol. Diagn.* 18:124–30
116. Diehl F, Schmidt K, Durkee KH, Moore KJ, Goodman SN, et al. 2008. Analysis of mutations in DNA isolated from plasma and stool of colorectal cancer patients. *Gastroenterology* 135:489–98
117. Ahlquist DA, Zou H, Domanico M, Mahoney DW, Yab TC, et al. 2012. Next-generation stool DNA test accurately detects colorectal cancer and large adenomas. *Gastroenterology* 142:248–56
118. Ahlquist DA, Taylor WR, Mahoney DW, Zou H, Domanico M, et al. 2012. The stool DNA test is more accurate than the plasma septin 9 test in detecting colorectal neoplasia. *Clin. Gastroenterol. Hepatol.* 10:272–77.e1
119. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. 2013. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J. Clin. Endocrinol. Metab.* 98:E1852–60
120. Nikiforov YE, Nikiforova MN. 2011. Molecular genetics and diagnosis of thyroid cancer. *Nat. Rev. Endocrinol.* 7:569–80
121. Steward DL, Carty SE, Sippel RS, Yang SP, Sosa JA, et al. 2019. Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: a prospective blinded multicenter study. *JAMA Oncol.* 5:204–12
122. Boyle JO, Mao L, Brennan JA, Koch WM, Eisele DW, et al. 1994. Gene mutations in saliva as molecular markers for head and neck squamous cell carcinomas. *Am. J. Surg.* 168:429–32
123. Wang Y, Springer S, Mulvey CL, Silliman N, Schaefer J, et al. 2015. Detection of somatic mutations and HPV in the saliva and plasma of patients with head and neck squamous cell carcinomas. *Sci. Transl. Med.* 7:293ra104
124. Ahn SM, Chan JY, Zhang Z, Wang H, Khan Z, et al. 2014. Saliva and plasma quantitative polymerase chain reaction-based detection and surveillance of human papillomavirus-related head and neck cancer. *JAMA Otolaryngol. Head Neck Surg.* 140:846–54
125. Hubers AJ, Prinsen CF, Sozzi G, Witte BI, Thunnissen E. 2013. Molecular sputum analysis for the diagnosis of lung cancer. *Br. J. Cancer* 109:530–37
126. De Mattos-Arruda L, Mayor R, Ng CKY, Weigelt B, Martinez-Ricarte F, et al. 2015. Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. *Nat. Commun.* 6:8839
127. Miller AM, Shah RH, Pentsova EI, Pourmaleki M, Briggs S, et al. 2019. Tracking tumour evolution in glioma through liquid biopsies of cerebrospinal fluid. *Nature* 565:654–58
128. Pan C, Diplas BH, Chen X, Wu Y, Xiao X, et al. 2019. Molecular profiling of tumors of the brainstem by sequencing of CSF-derived circulating tumor DNA. *Acta Neuropathol.* 137:297–306
129. Zheng MM, Li YS, Jiang BY, Tu HY, Tang WF, et al. 2019. Clinical utility of cerebrospinal fluid cell-free DNA as liquid biopsy for leptomeningeal metastases in *ALK*-rearranged NSCLC. *J. Thorac. Oncol.* 14:924–32
130. Juratli TA, Stasik S, Zolal A, Schuster C, Richter S, et al. 2018. *TERT* promoter mutation detection in cell-free tumor-derived DNA in patients with *IDH* wild-type glioblastomas: a pilot prospective study. *Clin. Cancer Res.* 24:5282–91
131. Panditharatna E, Kilburn LB, Aboian MS, Kambhampati M, Gordish-Dressman H, et al. 2018. Clinically relevant and minimally invasive tumor surveillance of pediatric diffuse midline gliomas using patient-derived liquid biopsy. *Clin. Cancer Res.* 24:5850–59
132. Mouliere F, Mair R, Chandrananda D, Marass F, Smith CG, et al. 2018. Detection of cell-free DNA fragmentation and copy number alterations in cerebrospinal fluid from glioma patients. *EMBO Mol. Med.* 10:e9323
133. Martinez-Ricarte F, Mayor R, Martinez-Sáez E, Rubio-Pérez C, Pineda E, et al. 2018. Molecular diagnosis of diffuse gliomas through sequencing of cell-free circulating tumor DNA from cerebrospinal fluid. *Clin. Cancer Res.* 24:2812–19

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

134. Li YS, Jiang BY, Yang JJ, Zhang XC, Zhang Z, et al. 2018. Unique genetic profiles from cerebrospinal fluid cell-free DNA in leptomeningeal metastases of *EGFR*-mutant non-small-cell lung cancer: a new medium of liquid biopsy. *Ann. Oncol.* 29:945–52
135. Huang TY, Piunti A, Lulla RR, Qi J, Horbinski CM, et al. 2017. Detection of Histone H3 mutations in cerebrospinal fluid-derived tumor DNA from children with diffuse midline glioma. *Acta Neuropathol. Commun.* 5:28
136. Pentsova EI, Shah RH, Tang J, Boire A, You D, et al. 2016. Evaluating cancer of the central nervous system through next-generation sequencing of cerebrospinal fluid. *J. Clin. Oncol.* 34:2404–15
137. Wang Y, Springer S, Zhang M, McMahon KW, Kinde I, et al. 2015. Detection of tumor-derived DNA in cerebrospinal fluid of patients with primary tumors of the brain and spinal cord. *PNAS* 112:9704–9
138. Foote MB, Papadopoulos N, Diaz LA Jr. 2015. Genetic classification of gliomas: refining histopathology. *Cancer Cell* 28:9–11
139. Risques RA, Kennedy SR. 2018. Aging and the rise of somatic cancer-associated mutations in normal tissues. *PLOS Genet*. 14:e1007108
140. Kato S, Lippman SM, Flaherty KT, Kurzrock R. 2016. The conundrum of genetic "drivers" in benign conditions. *J. Natl. Cancer Inst.* 108:djw036
141. Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, et al. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* 371:2477–87
142. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, et al. 2014. Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* 371:2488–98
143. Young AL, Challen GA, Birmann BM, Druley TE. 2016. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat. Commun.* 7:12484
144. Nair N, Camacho-Vanegas O, Rykunov D, Dashkoff M, Camacho SC, et al. 2016. Genomic analysis of uterine lavage fluid detects early endometrial cancers and reveals a prevalent landscape of driver mutations in women without histopathologic evidence of cancer: a prospective cross-sectional study. *PLOS Med*. 13:e1002206
145. Anglesio MS, Papadopoulos N, Ayhan A, Nazeran TM, Noe M, et al. 2017. Cancer-associated mutations in endometriosis without cancer. *N. Engl. J. Med.* 376:1835–48
146. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, et al. 2015. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* 348:880–86
147. Martincorena I, Fowler JC, Wabik A, Lawson ARJ, Abascal F, et al. 2018. Somatic mutant clones colonize the human esophagus with age. *Science* 362:911–17
148. Yokoyama A, Kakiuchi N, Yoshizato T, Nannya Y, Suzuki H, et al. 2019. Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature* 565:312–17
149. Yoshida K, Gowers KHC, Lee-Six H, Chandrasekharan DP, Coorens T, et al. 2020. Tobacco smoking and somatic mutations in human bronchial epithelium. *Nature* 578:266–72
150. Yizhak K, Aguet F, Kim J, Hess JM, Kubler K, et al. 2019. RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science* 364:eaaw0726

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

Annual Review of Pathology: Mechanisms of Disease

Volume 16, 2021

Contents

The Accidental Pathologist: A Curiosity-Driven Journey from Plant Evolution to Innate Immunity
Vinay Kumar 1

Lethal Infectious Diseases as Inborn Errors of Immunity: Toward a Synthesis of the Germ and Genetic Theories
Jean-Laurent Casanova and Laurent Abel 23

Animal Models and Their Role in Understanding the Pathophysiology of Cystic Fibrosis–Associated Gastrointestinal Lesions
Katherine N. Gibson-Corley and John F. Engelhardt 51

Molecular Pathogenesis of Merkel Cell Carcinoma
James A. DeCaprio 69

Monocytes in the Tumor Microenvironment
Stefano Ugel, Stefania Canè, Francesco De Sanctis, and Vincenzo Bronte 93

The Spectrum of *Helicobacter*-Mediated Diseases
Karen Robinson and John C. Atherton 123

Genetic Disease and Therapy
Theodore L. Roth and Alexander Marson 145

Opposing Roles of Type I Interferons in Cancer Immunity
Giselle M. Boukhaled, Shane Harding, and David G. Brooks 167

Detection and Diagnostic Utilization of Cellular and Cell-Free Tumor DNA
Jonathan C. Dudley and Maximilian Diehn 199

Immune Checkpoint Inhibitors for the Treatment of Cancer: Clinical Impact and Mechanisms of Response and Resistance
Sreya Bagchi, Robert Yuan, and Edgar G. Engleman 223

Gut Microbiota in Intestinal and Liver Disease
Rheinallt M. Jones and Andrew S. Neish 251

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166

Complement in Neurologic Disease
Nicholas E. Propson, Manasee Gedam, and Hui Zheng 277

The Hippo Pathway in Liver Homeostasis and Pathophysiology
Jordan H. Driskill and Duojia Pan 299

Metabolic Gatekeepers of Pathological B Cell Activation
Teresa Sadras, Lai N. Chan, Gang Xiao, and Markus Müschen 323

Genetic Insights into Alzheimer's Disease
Caitlin S. Latimer, Katherine L. Lucot, C. Dirk Keene, Brenna Cholerton, and Thomas J. Montine 351

Perspectives and Advances in the Understanding of Tuberculosis
Rachel L. Kinsella, Dennis X. Zhu, Gregory A. Harrison, Anne E. Mayer Bridwell, Jerome Prusa, Sthefany M. Chavez, and Christina L. Stallings 377

When a House Is Not a Home: A Survey of Antimetastatic Niches and Potential Mechanisms of Disseminated Tumor Cell Suppression
Sarah B. Crist and Cyrus M. Ghajar 409

Pathogenesis of Cholangiocarcinoma
Pedro M. Rodrigues, Paula Olaizola, Nuno A. Paiva, Irene Olaizola, Alona Agirre-Lizaso, Ana Landa, Luis Bujanda, Maria J. Perugorria, and Jesus M. Banales 433

The Membrane Interactions of Synuclein: Physiology and Pathology
Gautam Runwal and Robert H. Edwards 465

The Complex Clinical and Genetic Landscape of Hereditary Peripheral Neuropathy
Soumitra Ghosh and Warren G. Tourtellotte 487

Errata

An online log of corrections to *Annual Review of Pathology: Mechanisms of Disease* articles may be found at http://www.annualreviews.org/errata/pathmechdis

Annu. Rev. Pathol. Mech. Dis. 2021.16:199-222. Downloaded from www.annualreviews.org
Access provided by 12.51.190.113 on 09/07/23. For personal use only.

Personalis EX2166