Circulating DNA-based lung cancer diagnostics and follow-up: looking for epigenetic markers

Elena Y. Rykova1,2, Anastasia A. Ponomaryova3,4, Ivan A. Zaporozhchenko1,5, Valentin V. Vlassov1, Nadezhda V. Cherdyntseva2,6, Pavel P. Laktionov1,5

1Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Science, Novosibirsk, Russia; 2Department of Technologies Derived Ecological Problems, Novosibirsk State Technical University, Novosibirsk, Russia; 3Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia; 4Department of Applied Physics, Tomsk Polytechnical University, Tomsk, Russia; 5Meshalkin Siberian Federal Biomedical Research Center, Novosibirsk, Russia; 6Laboratory of Translational Cell and Molecular Biomedicine, Tomsk State University, Tomsk, Russia

Contributions: (I) Conception and design: EY Rykova, PP Laktionov; (II) Administrative support: PP Laktionov, VV Vlassov, NV Cherdyntseva; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: EY Rykova, AA Ponomaryova, IA Zaporozhchenko; (V) Data analysis and interpretation: EY Rykova, AA Ponomaryova, IA Zaporozhchenko; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Elena Y. Rykova, PhD. Institute of Chemical Biology and Fundamental Medicine Siberian Branch of the Russian Academy of Science, Novosibirsk 630090, Russia. Email: rykova.elena.2014@gmail.com.

Abstract: This review outlines the major trends in the investigation of circulating cell-free DNA-based methylated markers, including key technical aspects of the circulating cell-free DNA methylation analysis, recent advancements in methylation detection techniques, and most successful examples of their combined use. The approaches to discovery and validation of epigenetic markers are discussed, highlighting the advantages and disadvantages of these methods. Special attention is devoted to cell-free DNA methylated markers, circulating in blood, which could be used for minimally invasive lung cancer (LC) diagnostics, prediction of antitumor treatment efficiency and disease prognosis.

Keywords: Lung cancer (LC); diagnosis; prognosis; oncomarkers; methylation; circulating DNA

Submitted Aug 30, 2017. Accepted for publication Feb 08, 2018. doi: 10.21037/tcr.2018.02.08

View this article at: http://dx.doi.org/10.21037/tcr.2018.02.08

Introduction

Lung cancer (LC) remains the world’s leading cause of cancer-associated mortality (1). Due to the late manifestation of symptoms and low sensitivity of chest radiography used as screening technique most LCs (about 85%) are diagnosed in patients with either locally advanced disease and/or developed distant metastases, which results in a tremendously low 5-year survival rate (4% in case of a metastatic disease) (2). By contrast, 5-year survival rate of early stage LC patients exceeds 50%. Recently, low dose computed tomography (LDCT) was shown to be a more promising screening tool, but its implementation into the routine practice is hurdled by high false positive rate (fewer than 5% of discovered nodules are malignant), high cost and unexplored health risks (3). False positive CT results entail follow up procedures that are invasive, costly, and have associated morbidity and mortality. Likewise, magnetic resonance imaging (MRI) is a costly method with high false positive rates and is mostly applied to confirm existing diagnoses. Histological analysis of samples collected using bronchoscopy biopsies and fine needle aspirations provides valuable diagnostic information but has drawbacks due to procedure invasiveness.

Alongside early diagnostics another major battlefield in the war against LC is the assignment of appropriate and effective treatment regimes. LC is a diverse group of histologically and clinically distinct malignancies and therefore the effect of therapeutic measures differs dramatically on a case-to-case basis even for superficially similar tumors (4). Up to 30% of surgically treated stage
I patients are lost to recurrent disease (5). Identification and clinical application of multiple noninvasive or tissue-based molecular biomarkers can assist in understanding the taxonomy and molecular landscape of LC necessary to identify (stage I) patients at high-risk for recurrence who may benefit from adjuvant chemotherapy or innovative immunotherapy regimens (5). Similarly, immunotherapies targeted at specific genetic alterations such as EGFR mutations and ALK fusions are known to improve outcomes for a subset of patients even with advanced stage of LC (6,7). Recent promising results demonstrate that genomic landscape of lung tumours shapes response to antibody targeting programmed cell death-1 (anti-PD-1) therapy, suggesting that exome-guided neoantigen identification may improve treatment responses (8).

Thus, new strategies for preclinical screening and monitoring of post therapy relapses based on selected “omics” biomarkers may be instrumental in relieving the burden of LC.

**Cell-free circulating DNA (cfDNA) biomarkers**

Complementary diagnostics using blood-derived molecular DNA markers is one attractive solution to the problem. cfDNA isolated from the blood of cancer patients contains DNA fragments shed from tumor cells and provides a convenient and minimally invasive access to the molecular portrait of cancer (9,10). Recently, cell-surface-bound fraction of circulating DNA (csbDNA) was proposed as an additional source of material (11-13). Cell-free circulating tumour DNA (ctDNA) in the blood plasma/serum presents a surrogate for the entire cancer genome (including the clonal landscape of the primary tumour and metastases) which gives it a unique appeal as liquid biopsy that provides “real-time” information about the evolution of tumor genome essential for planning the precision treatment regime. Our knowledge of cfDNA origins, mechanism and rate of release is still incomplete. The main sources are proposed to be the necrotic and apoptotic cell death (either before or post macrophage engulfment and digestion) along with active secretion by living cells (9,14,15). In any case, all tissues of the multicellular body have the potential to contribute to the collective information of the circulating genome. Another reason to consider ctDNA for liquid biopsy is concerned with its rapid clearance from bloodstream with half-life of approximately one hour followed by a slow phase with half-life of 13 hours as shown by the study of fetal cfDNA kinetics (16). These properties of cfDNA make it a useful tool for detection of residual cancer disease and recurrence during the post-treatment follow-up.

**KRAS** oncogene mutation was the first cfDNA marker detected in blood of cancer patients (17). Diagnostic tests for EGFR and KRAS mutations in cfDNA are now commercially available and can be used as prognostic indicators in advanced NSCLC patients. Meta-analysis of recent data demonstrated that EGFR mutations in ctDNA predicted better progression-free survival in advanced NSCLC patients treated by EGFR-TKIs (18). The feasibility of using tumour-specific mutations in ctDNA in diagnosis of cancer and monitoring of response to therapy has been widely demonstrated, although, the high individual variability of mutational landscape of tumour DNA makes it a labor-intensive approach (19).

**Advantages of ctDNA methylated markers**

Methylation of cytosines on position 5 in CpG dinucleotides is one of the most important epigenetic modifications in DNA of eukaryotic cells necessary for cellular differentiation and normal development of tissues and organs. Changes in DNA methylation are a common feature of most cancer types and occur early and consistently in cancer development, thus making aberrant DNA methylation a broadly applicable marker of ctDNA in blood alongside mutations, deletions, loss of heterozygosity, etc. (20). However, the consistency of methylation events in cancer suggests them a somewhat more reliable option. Age-, gender- and smoking-associated changes in methylation profile are also known to occur and should be considered when methylation alterations are used for diagnostic purposes (21). Implications of these findings for the study of ctDNA methylation are currently being reviewed.

Previously, aberrant DNA methylation has been strongly implicated in cancer initiation and progression (21,22). Hypermethylation of CpG-rich regions in gene promoters called CpG islands (CGIs) may induce repression of individual tumour suppressor genes, while global hypomethylation contributes to tumourigenesis through the promotion of genomic instability and activation of oncogenes. In non-small cell lung cancer (NSCLC) hypermethylation of tumour suppressor genes can serve as an indication of stage and histological type, aiding in assessing survival prognosis, disease progression and recurrence (23). Most of the studies to date have been conducted on cfDNA to validate previous discoveries of
CpG island hypermethylation of certain single-copy genes (24-26). An alternative strategy is to look at hypomethylation of genetic mobile elements, which are scattered around the human genome as multiple repeated copies and normally silenced by methylation (27,28). Detection of repetitive DNA should have improved sensitivity compared with a single-copy gene assay. Hypomethylation of LINE-1 retrotransposons in lung tumours (29-31) and blood cells of LC patients (32,33) has been described as one of the key hallmarks of carcinogenesis. Moreover, the degree of LINE-1 hypomethylation was associated with clinical features of the disease and survival prognosis (34,35). Therefore, hypomethylation of repetitive elements could be used as a screening, diagnostic and prognostic biomarker in LC.

As mentioned before, cfDNA represents a mixture of molecules originating from different tissues, with ctDNA accounting for as low as 0.05% of total cfDNA or less in many cancer patients, especially at the early stages of the disease (36). High accuracy of discrimination between unmethylated and methylated alleles is possible due to the development of the bisulfite conversion technique based on chemical treatment of DNA strands by sodium bisulphite that converts unmethylated cytosines to uracils. Methylation-specific PCR (MSP) of bisulphite converted DNA allows for a highly specific detection of a single methylated allele over a background of more than 1,000-fold excess of unmethylated alleles, which is crucially important for the study of ctDNA.

These properties make aberrant methylation of ctDNA a promising cancer biomarker, and recent high throughput investigations have demonstrated the correspondence between the changes in methylation profiles of ctDNA and DNA from paired tumor tissue (37-40). However, DNA methylation-based biomarkers have not been incorporated into commercially available assays for in vitro diagnostics until very recently, and currently Epi ProColon (Epigenomics AG, Berlin, Germany) is the only FDA approved ctDNA assay for colorectal cancer screening based on qPCR detection of hypermethylated Septin9 in the ctDNA derived from blood plasma. A prospective LC screening assay from the same company named Epi ProLung BL is based on qPCR detection of hypermethylated SHOX2 and PTGER4 in DNA from bronchoalveolar fluid obtained from patients during bronchoscopy. Notably, the authors have also recently demonstrated the potential utility of methylated SHOX2 and PTGER4 in plasma cfDNA for LC screening (41) (Table 1).

Technical aspects of the methylated cfDNA analysis

The long and winding road of cfDNA-based methylated liquid biopsy markers into clinical practice can be attributed to the challenges inherent to the nature of cfDNA. It is present in the circulation in a low quantity of less than 10 ng per mL of blood plasma, and low quality, inasmuch as it is highly fragmented and mostly consists of roughly 180 bp molecules. So far, several reviews aimed to determine the optimal pre-analytical considerations necessary to optimize cfDNA yield and quality, establish protocols for cfDNA analysis and suggest guidelines for translation of cfDNA analysis into routine clinical practice (9,59-61). These reports defined various parameters for optimal blood sample handling before cfDNA isolation based on literature data and proposed confirmatory experiments as a first step in this direction. In brief, when developing a methylated ctDNA marker assay one should first standardize sample preparation: a choice must be made between plasma and serum, storage temperature and time interval between blood sampling and sample processing should be optimized, suitable DNA extraction and purification methods should be selected. In order to minimize the loss of cfDNA fragments during isolation, a broad range of commercial kits designed to recover low molecular weight nucleic acids is available, although when compared they demonstrate vastly different efficiency and reproducibility (62-64). Hulbert et al developed the methylation-on-beads (MOB) technique, which reduces sample loss thereby potentially increasing sensitivity (65).

Cytosine methylation detection techniques based on bisulfite conversion

Second step in setting up a cfDNA assay is to select an effective method for discriminating between methylated and unmethylated CpG sites. The treatment of cfDNA with sodium bisulfite has become the most widely used method, easily combined with various downstream detection technologies, including PCR-based assays, pyrosequencing, dideoxy-sequencing, high resolution melting, single-strand conformation polymorphism analysis, microarrays, next-generation sequencing (NGS), MALDI-TOF mass spectrometry, etc. Quantitative MSP assays and digital PCR-based approaches have high analytical sensitivity of detection of rare methylated alleles in the presence of excess unmethylated cfDNA, making them suitable for the analysis...
<table>
<thead>
<tr>
<th>Gene</th>
<th>Marker</th>
<th>Ref</th>
<th>Source</th>
<th>Changes in lung cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>Adenomatous polyposis coli</td>
<td>(39,42)</td>
<td>Plasma</td>
<td>Significantly changes of methylation level in lung cancer patients compared with healthy subjects</td>
</tr>
<tr>
<td>CDH13</td>
<td>Cadherin 13</td>
<td>(43)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>DCLK1</td>
<td>Doublecortin like kinase 1</td>
<td>(38)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>DLEC1</td>
<td>Deleted in lung and esophageal cancer 1</td>
<td>(44)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>LINE-1</td>
<td>LINE-1 retrotransposable element 1</td>
<td>(45)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>P16 (CDKN2A)</td>
<td>Cyclin-dependent kinase inhibitor 2A</td>
<td>(46-48)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>RARB2</td>
<td>Retinoic acid receptor beta 2</td>
<td>(49)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>SEPT9</td>
<td>Septin 9</td>
<td>(50)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>SHOX2</td>
<td>Short stature homeobox 2</td>
<td>(51)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>BRMS1</td>
<td>Breast cancer metastasis suppressor 1</td>
<td>(52)</td>
<td></td>
<td>Negative impact on survival</td>
</tr>
<tr>
<td>DCLK1</td>
<td>Doublecortin like kinase 1</td>
<td>(27)</td>
<td></td>
<td>Negative impact on survival</td>
</tr>
<tr>
<td>LINE-1</td>
<td>LINE-1 retrotransposable element 1</td>
<td>(53)</td>
<td></td>
<td>Dynamic changes of methylation level in response to antitumor therapy</td>
</tr>
<tr>
<td>APC</td>
<td>Adenomatous polyposis coli</td>
<td>(42)</td>
<td>Serum</td>
<td>Significantly changes of methylation level in lung cancer patients compared with healthy subjects</td>
</tr>
<tr>
<td>CDH1</td>
<td>Cadherin 1</td>
<td>(54)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>DAPK</td>
<td>Death-associated protein kinase</td>
<td>(55,56)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>DCC</td>
<td>DCC netrin 1 receptor</td>
<td>(54)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>GSTP1</td>
<td>Glutathione S-transferase pi 1</td>
<td>(55)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>MGMT</td>
<td>O-6-methylguanine-DNA methyltransferase</td>
<td>(55)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>P16 (CDKN2A)</td>
<td>Cyclin-dependent kinase inhibitor 2A</td>
<td>(55)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>RASSF1A</td>
<td>Ras association domain family 1 isomorphism A</td>
<td>(54,56)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>TMS1</td>
<td>PYD and CARD domain containing</td>
<td>(56)</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>CHFR</td>
<td>Checkpoint with forkhead and ring finger domains</td>
<td>(57)</td>
<td></td>
<td>Negative impact on survival with second-line EGFR-TKIs, compared to chemotherapy</td>
</tr>
<tr>
<td>SFN</td>
<td>Stratfin</td>
<td>(58)</td>
<td></td>
<td>Positive impact on survival with platinum-based chemotherapy</td>
</tr>
</tbody>
</table>

* The information is the same as above written.

Of ctDNA. MSP initially only allowed for a qualitative estimation of methylation of the region of interest. Quantitative MSP using a Taqman probe (MethyLight) enabled the detection of methylated allele in a 10,000-fold excess of unmethylated alleles thus providing a 10-fold increase of sensitivity compared to conventional MSP. A further evolution of this approach is digital MethyLight (dMethyLight) which is based on compartmentalization of DNA templates over multiple reaction wells, allowing for detection of single methylated alleles and quantitative analysis via counting of positive wells (66). Droplet digital MethyLight (ddMethyLight) further improves on the sensitivity of MethyLight assay by using droplets instead of wells used in dMethyLight, increasing effective dilution factor by an order of magnitude (67). In the recent years other highly sensitive technologies have been developed, such as Methyl-BEAMing combining emulsion dPCR with magnetic beads and flow cytometry to achieve sensitivity equivalent to dPCR (68), or RainDrop digital PCR (69). An approach called DREAMing (Discrimination of Rare
Exemplary analysis of methylated CpGs is pyrosequencing—sequencing-by-synthesis system that relies on the luminometric detection of pyrophosphates released as nucleotides are incorporated into the extended strand. This technique provides quantitative data of the methylation status of multiple individual CpGs within a region of interest in bisulfite-treated DNA (71). Usefulness of pyrosequencing-based methods on ctDNA samples is limited compared with PCR-based assays by the gradual decay of the accuracy of evaluation of methylation status for CpG sites based on the distance from the 3’ end of the primer.

Nevertheless, there are several significant drawbacks associated with the chemical conversion of cytosines. One of the main drawbacks of bisulfite conversion is the introduction of breaks in the DNA, leading to a dramatic loss of already highly fragmented cfDNA. Bisulfite treatment reduces sequence complexity, puts constrains on primer design for PCR amplification, and disallows the distinction between 5-methylcytosines and 5-hydroxymethylcytosines (72), leading to false-positives in downstream analyses. Finally, incomplete conversion of either methylated or unmethylated cytosines leads to false-positive and false-negative results even when using commercially available kits (73). Therefore, a number of bisulfite-independent methods have been developed aiming to improve the locus-specific CpG methylation detection.

**Other methods and their combinations**

Affinity enrichment-based methods utilize specific antibodies recognizing methylated cytosines or methyl-binding proteins to selectively recover methylated DNA. Examples include MethylCpG Binding Domain MBD2 proteins (MBD, also termed Methyl Cap), methylated DNA immunoprecipitation (MeDIP) and methylated CpG island recovery assay (MIRA) (74-76). MBD fusion proteins bind specifically to dsDNA that is methylated at CpG sites on both strands. They demonstrate a bias for high CpG densities and preferentially recover methylated CGIs (60). Commercially available MBD affinity kits use MBD proteins fused with glutathione-S-transferase (GST) (MethylMagnet, RiboMed; MethylQuest, Millipore; Methyl-Cap, Diagenode) or with His6 (hexahistidine-tag or polyhistidine-tag) affinity tags immobilized to magnetic beads (MethylMiner, Life Technologies; MethylCollector, Active Motif). These techniques can be combined with quantitative PCR to perform locus-specific assessment of methylation or with sequencing to obtain whole-genome methylation profiles. Low recovery of methylated DNA is the main disadvantage of this approach, alongside its inability to target single CpG sites.

Another group of methods employs methylation-sensitive restriction enzymes that cleave at sites with unmethylated CpGs, therefore in the following PCR only the methylated alleles are amplified. Methylation-sensitive restriction enzymes can be used in combination with different detection techniques, such as combined bisulfite restriction analysis (COBRA), differential methylation hybridization (DMH), HpaII tiny fragment enrichment by ligation-mediated PCR (HELP), and methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) (77). In order to circumvent the major disadvantage of this approach, stemming from the ability of the enzyme to only recognize a set pattern of CpG sites, another approach called Methylation Specific Nuclease-assisted Minor-allele Enrichment (MS-NaME) was recently developed (78). This technique utilizes oligonucleotide probes that direct double-strand-specific DNA nuclease (DSN) to multiple targets in bisulfite-treated DNA followed by amplification of the targeted regions, resulting in the enrichment of multiple methylated or unmethylated rare epigenetic alleles. Additional examples of the integrative approach are the nucleosome occupancy and methylome sequencing (NOMe-seq) method which allows to simultaneously determine the nucleosome occupancy and DNA methylation. In NOMe-seq, nuclei are treated with GpC Methyltransferase (M.CviPI) before bisulfite conversion and sequencing.

To sum up, a plethora of methods and their modifications utilised for the detection of DNA methylation provides an array of potential solutions for biomarker research, addressing the needs of a wide variety of studies in LC, but simultaneously rendering the comparison of results very difficult, and creating a major obstacle for conducting meta-analyses (79).

**Discovery and validation of epigenetic markers**

Third critical step in the development of an effective methylated
ctDNA-based assay is to select the valid markers (61). Two approaches for marker selection have been utilized so far. One approach entails the selection of tumor tissue-associated candidate genes, followed by development and validation of sensitive and specific assay using representative groups of patients and controls. Another approach relies on genome-wide methylation profiling techniques, such as array hybridization and NGS of cfDNA. The former method has instigated the evaluation of a number of potentially useful epigenetic signatures as markers applicable for LC diagnosis, staging, prognosis and assessment of the response to therapy. Some of the candidates are currently undergoing validation studies. Notably, this approach was used in the development of the successful Epi ProColon assay. In order to select the proper strategy for the development of novel biomarkers it is imperative to understand if the benefits, potential biases and resource requirements of a particular technique are suitable to meet the research objectives of the study (61,80,81).

**Genome-wide methylation profile analysis**

Studies based on genome-wide methylation profiling of cfDNA remain scarce due to technical challenges such as low starting amounts of cfDNA, further aggravated by the loss of material during sample processing, DNA extraction, bisulfite conversion and library preparation.

Assessment of DNA methylation on a genome level was initially conducted by a combination of microarray hybridization with either digestion of DNA by methylation sensitive enzymes, affinity purification or bisulfite conversion (82). One of the most widely used array-based approaches for methylation profiling is the Infinium HumanMethylation450 BeadChip Array (Illumina) or 450k, which is able to detect the methylation status of 485,000 individual CpGs from reference genes and CGIs (83). The reason behind the popularity of the 450k platform are the low cost of batch processing of the samples and ease of data analysis using publicly available R packages. Array-based methodologies are ideal for first-pass methylation profiling; however, the requirement for a large amount of input DNA (hundreds of ng) precludes their use for widespread assessment of DNA methylation in liquid biopsies. Still, encouraging results from a recent pilot study showed that genome-wide profiles of DNA methylation in ctDNA are consistent with corresponding tumor tissues (84).

Whole genome bisulfite sequencing (WGBS) is perhaps the most powerful method to explore the methylome with the potential to identify the methylation status of every CpG site in the genome. WGBS libraries are prepared similarly to regular whole genome sequencing, with the additional step of bisulfite conversion (85). The critical determinants for applying high-throughput WGBS to epigenome-wide association studies for cfDNA-based cancer marker discovery are reviewed elsewhere (81). Targeted sequencing using panels consisting of tens to hundreds of select genes is a more appealing and cost-effective solution for ctDNA analysis. Recently whole-genome bisulfite sequencing from nanogram quantities of input cfDNA has been enabled by novel methods such as T-WGBS (transposase-based library construction) and PBAT (post-bisulfite adaptor tagging, which can be performed with as little as 125 pg of input DNA) (86-89). Study of thousands of hypermethylated CGIs in cfDNA using methylated CpG tandems amplification and sequencing (MCTA-Seq) was performed by Wen et al. in tissue and plasma samples from hepatocellular carcinoma (HCC) patients and healthy controls, resulting in the identification of a panel of four cancer-specific genes (RGS10, STSLA6, RUNX2 and VIM) (Table 2) (89).

Distinct advantage of genome-wide DNA methylation assays is the option to apply innovative bioinformatic solutions to data analysis, as evidenced by a great number of recent investigations reporting newly developed approaches for evaluating the proportion and the tissue-of-origin of tumor-derived cell-free DNA in blood samples with the general aim of developing a universal cancer detection technique (Table 2). The underlying principle of cfDNA tissue mapping is based on the genome-wide bisulfite sequencing of plasma cfDNA and methylation deconvolution of the sequencing data (88). Stretches of adjacent CpG sites are used as tissue-specific methylation markers to increase signal-to-noise ratio and improve the sensitivity of the assay. Sun et al. [2015] (88) estimated the proportion of cfDNA contributed by different tissues and showed that an abnormally high presence of cfDNA from a specific tissue may indicate it as the tumor site. Lehmann-Werman et al. [2016] (90) applied the same rationale to the diagnosis of pancreatic cancer (90). With the help of the Illumina 450k methylation arrays (HM450K) platform they showed that roughly 50% of pancreatic cancer patients demonstrated a substantial over-representation of pancreas-derived cfDNA fragments compared with healthy subjects. Recent study from Kang et al. proposed a probabilistic method, CancerLocator, which exploits the diagnostic potential of cfDNA to determine not just the presence
Table 2 Examples of high-throughput approach to cfDNA methylation analysis

<table>
<thead>
<tr>
<th>Name</th>
<th>Input data</th>
<th>Method</th>
<th>Results</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methylated CpG tandems amplification and sequencing (MCTA-Seq)</td>
<td>Sequencing on Illumina HiSeq2000/2500 system</td>
<td>After bisulfite conversion, cfDNA is amplified using polymerase with displacement activity and MCTA-Seq primer A consisting of a semi-random sequence (RS) containing one CpG at the 3’-end, a unique molecular identifier linker and an anchor sequence is at the 5’-end. Second primer MCTA-Seq primer B, contains the CpG tandem sequence “CGCGCGG” at the 3’-end followed by a 4-bp sequence of A, T or G. This selectively amplifies the methylated CpG tandem sites. Last, exponential PCR amplification is performed using primers against the anchor sequences.</td>
<td>The method could detect as little as 7.5 pg of methylated DNA, with sensitivity of detecting methylated allele with as frequencies low of 0.25%. Panel of four genes (RGS10, STSIA6, RUNX2 and VIM) was identified as marker of hepatocellular carcinoma.</td>
<td>(89)</td>
</tr>
<tr>
<td>cfDNA tissue mapping</td>
<td>Genome-wide bisulfite sequencing, Illumina HiSeq and NextSeq</td>
<td>Identification of tissue-specific methylation markers by dividing CpG-rich genomic loci into 500-bp units, and studying their variability across 14 tissues. Then using these signatures methylation deconvolution using quadratic programming can be used to assess contribution of different tissues to the cfDNA pool.</td>
<td>Accurate estimated proportion of cfDNA from different tissues and fetal cfDNA in pregnant women, showed increased DNA input from tumour sites.</td>
<td>(88)</td>
</tr>
<tr>
<td>Tissue-specific cfDNA methylation patterns</td>
<td>Illumina Infinium HumanMethylation450 BeadChip Array</td>
<td>Regional nature of CpG methylation was used to increase the signal-to-noise ratio, by taking into account the methylation status of 4 to 9 adjacent CpG sites. These expanded sites were used to construct tissue-specific methylation signatures.</td>
<td>Identification of beta-cell cfDNA in T1D patients, oligodendrocyte cfDNA in multiple sclerosis patients, brain derived cfDNA in patients with brain injuries, and over-representation of pancreas-derived cfDNA fragments in patients with pancreatic cancer.</td>
<td>(90)</td>
</tr>
<tr>
<td>CancerLocator</td>
<td>Whole-genome bisulfite sequencing data from other studies and Illumina Infinium HumanMethylation450 BeadChip Array data from TCGA repository</td>
<td>CancerLocator simultaneously infers the proportion and tissue of origin of ctDNA in a blood sample using genome-wide DNA methylation data. The model is first learned on TCGA methylation data grouped into clusters of at least 3 CpG sites. Then the algorithm can use the informative features to estimate the fraction of ctDNAs in the plasma and the likelihood that the detected ctDNAs come from each tumor type. Based on these data it returns the probability of patient having a tumor and the predicted primary cancer site.</td>
<td>The method outperformed RandomForest and Support Vector Machine based methods on simulated plasma cfDNA samples, and showed promising results in the classification of real plasma samples of liver, lung and breast cancer patients.</td>
<td>(91)</td>
</tr>
<tr>
<td>Methylated haplotype load (MHL)-based tissue mapping</td>
<td>Published and in-house WGBS data from human primary tissues, embryonic stem cells, progenitors and cancer cell lines, reduced-representation bisulfite sequencing (RRBS) data sets from ENCODE and HM450K data from TCGA</td>
<td>The method is also based on the regional nature of methylation and the expanded concept of genetic linkage disequilibrium. First, genome was divided into blocks of CpG sites with tightly coupled methylation (methylation haplotype blocks, MHB). MHL metric was used to define the methylation pattern and depth within one MHB block as the weighted mean of the fraction of fully methylated haplotypes and substrings at different lengths. RandomForest models were used to distinguish between tissue-of-origin of cfDNA.</td>
<td>The authors achieved 90% accuracy for mapping of primary tissue samples, average prediction accuracy of 82.8%, 88.5% and 91.2% for the CRC, LC patients and healthy individuals' plasma samples.</td>
<td>(92)</td>
</tr>
</tbody>
</table>
but also the location of the tumors, based on methylation signatures associated with tissue-specific cancers (91). Guo et al. [2017] performed an exhaustive search of tissue-specific methylation haplotype blocks (MHBs) across the entirety of the genome and proposed a methylated haplotype load (MHL) block-level metric for systematic discovery of informative markers. By using an original analytical framework and identified markers, the authors demonstrated the ability to accurately identify the tissue origin of cfDNA and classify plasma samples from LC and colorectal cancer patients (92).

Another innovative approach to cfDNA tissue mapping comes from the discovery of the direct association between cfDNA fragmentation patterns and nucleosome positioning (93). Ulz et al. [2016] performed whole-genome sequencing of plasma cfDNA and identified two discrete regions containing transcription start sites (TSSs) where nucleosome occupancy resulted in different read depth coverage patterns for expressed and silent genes and used support vector machine models to classify of cancer plasma cfDNA samples and infer gene expression in the primary tumour.

### Early LC screening

So far, a significant number of candidate tumour tissue-associated genes were selected and thoroughly validated, yielding the biomarkers listed in Table 1 and reviewed elsewhere (95,96). The list includes hypermethylated SHOX2 (38,39,41,51), RASSF1A (4,56,97-99), RARB2 (49,97), P16 (46-48,55,98), MGMT (54,55,100), death-associated protein kinase (DAPK) (55,100,101), APC (39,42,54) and DLEC1 (39,44). Notably, detection of methylated markers in blood plasma of smokers with COPD, which have a high risk of developing LC, predicted LC 3–18 months before its clinical diagnosis (99). Two genes, P16 and ESR1, are methylated in the peripheral blood DNA of stage I LC patients, providing higher sensitivity of LC detection in the same patients than CEA (102).

Diagnostic efficiency of individual gene assays is generally lower compared with a panel of two or more methylated markers previously reported to have high diagnostic potential (Table 3). Their potential for cancer screening is at its highest if used as complementary tests to improve the diagnostic accuracy of LC risk assessment by low-

### Table 2 (continued)

<table>
<thead>
<tr>
<th>Name</th>
<th>Input data</th>
<th>Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleosome footprint approach to cfDNA mapping</td>
<td>Deep sequencing of libraries prepared by end-repair and adaptor ligation to cfDNA fragments</td>
<td>A metric for the prediction of nucleosome protection sites was proposed—windowed protection score (WPS), the number of DNA fragments completely spanning a 120 bp window centered at a given genomic coordinate, minus the number of fragments with an endpoint within that same window. Using this metric genome-wide nucleosome protection map was inferred based on cfDNA.</td>
<td>Using the method, the authors showed association between cfDNA fragmentation patterns, nucleosome positioning and known transcription factor binding sites. Also, correlation was reported between nucleosome positioning and tissue-of-origin of cfDNA.</td>
</tr>
<tr>
<td>Tissue mapping by different read coverage of active and silent genes</td>
<td>Paired-end-sequencing of plasma cfDNA and micrococcal nuclease treatment</td>
<td>The patterns of fragmentation were assessed based on read coverage, two discrete regions containing transcription start sites (TSSs) where nucleosome occupancy results in different read depth coverage patterns for expressed and silent genes and used support vector machine models to classify of cancer plasma cfDNA samples and infer gene expression in the primary tumour.</td>
<td>The approach was able to correctly identify and classify expression changes in primary tumours based on copy number variation and nucleosome occupancy inferred from cfDNA. Based on the analysis of 426 plasma samples from patients with metastasized colon, prostate, breast, lung, and other cancers the authors found that 220 (51.6%) of these samples had at least 100 genomic bins (&gt;5.6 Mb) suitable for promoter read depth analysis.</td>
</tr>
</tbody>
</table>
Table 3 Combination of circulating methylated markers for lung cancer screening, diagnostics and prognosis

<table>
<thead>
<tr>
<th>Panel of methylated markers</th>
<th>Detected changes</th>
<th>Source</th>
<th>Clinical application</th>
<th>Method of detection</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><em><em>RASSF1A</em>, RARB2</em>*</td>
<td>Demonstrated sensitivity of 87% and specificity of 75% for lung cancer diagnosis</td>
<td>Plasma</td>
<td>Diagnostics</td>
<td>qMS-PCR</td>
<td>(97)</td>
</tr>
<tr>
<td><em><em>APC</em>, RASSF1A</em>, KLK10, CDH13 <em>and DLEC1</em>*</td>
<td>Demonstrated 84% sensitivity and 74% specificity for lung cancer diagnosis</td>
<td>Plasma</td>
<td>Diagnostics</td>
<td>qMS-PCR</td>
<td>(51)</td>
</tr>
<tr>
<td><em><em>SHOX2</em>, PTGER4</em>*</td>
<td>Demonstrated specificity of 90%, sensitivity of 67% for lung cancer diagnosis</td>
<td>Plasma</td>
<td>Diagnostics</td>
<td>Rt-PCR</td>
<td>(41)</td>
</tr>
<tr>
<td><em><em>APC</em>, RASSF1A</em>, CDH13*, KLK1, DLEC1**</td>
<td>Demonstrated sensitivity of 84% and a specificity of 74% for cancer diagnosis</td>
<td>Plasma</td>
<td>Diagnostics</td>
<td>MS-PCR</td>
<td>(39)</td>
</tr>
<tr>
<td><em><em>RARB2</em>, RASSF1A</em>*</td>
<td>Reduced after neoadjuvant chemotherapy and surgery, increase of methylation level associated with cancer progression (relapses)</td>
<td>Plasma</td>
<td>Prognosis, monitoring</td>
<td>qMS-PCR</td>
<td>(97)</td>
</tr>
<tr>
<td><em><em>RASSF1A</em>, APC</em>*</td>
<td>Elevated methylation level after chemotherapy; correlation with good response to cisplatin</td>
<td>Plasma</td>
<td>Prognosis</td>
<td>qMS-PCR</td>
<td>(103)</td>
</tr>
<tr>
<td><em><em>APC</em>, RASSF1A</em>, CDH13*, CDKN2A**</td>
<td>Association with a higher risk of recurrence and decreased overall survival of patients</td>
<td>Plasma</td>
<td>Prognosis</td>
<td>MS-PCR</td>
<td>(104)</td>
</tr>
<tr>
<td><strong>SOX17, TAC1, HOXA7, CDO1, HOXA9 and ZFP42</strong></td>
<td>First combination: TAC1, HOXA17 and SOX17 The sensitivity and specificity for lung cancer diagnosis using the best individual genes was 65% to 76% and 74% to 84%</td>
<td>Plasma</td>
<td>Diagnostics</td>
<td>qMS-PCR and Methylation on Beads for cancer associated genes</td>
<td>(105)</td>
</tr>
<tr>
<td><strong>Second combination: CDO1, TAC1 and SOX17</strong></td>
<td>A three-gene combination of the best individual genes has sensitivity and specificity of 93% and 62% The 4-marker model yielded an AUC of 0.85 with a sensitivity of 97% and a specificity of 73%.</td>
<td>Plasma/serum</td>
<td>Diagnostics</td>
<td>Genome-wide methylation screening, methylation-sensitive restriction enzyme digestion and enrichment of methylated DNA and parallel qPCR</td>
<td>(106)</td>
</tr>
<tr>
<td><em><em>APC</em> CDH1</em>, MGMT*, DCC, RASSF1A* and AIM1**</td>
<td>Demonstrated sensitivity of 84% and a specificity of 57% for lung cancer diagnosis</td>
<td>Serum</td>
<td>Diagnostics</td>
<td>qMS-PCR</td>
<td>(54)</td>
</tr>
<tr>
<td><strong>AHRR, F2RL3 genes, (2q37.1, 6p21.33 and 12q14.1 loci)</strong></td>
<td>Cpgs hypomethylation in these regions (AHRR, F2RL3, 2q37.1, 6p21.33 and 12q14.1 loci) related to smoking that may raise lung cancer risk/ Hypomethylation of these regions were lowest for current smokers and increased with time since quitting for former smokers</td>
<td>Peripheral blood (DNA samples were extracted from dried blood spots)</td>
<td>Screening</td>
<td>Illumina Infinium HumanMethylation450 beadchip assays</td>
<td>(107)</td>
</tr>
</tbody>
</table>

Table 3 (continued)
dose computed tomography (LDCT) screening (Figure 1). The National Lung Screening Trial (USA) demonstrated a 20% reduction in LC mortality using LDCT screening. If confirmed in a validation study, the developed panels of methylation markers could be used as a companion test for LDCT screening, identifying patients at high risk for LC, reducing false positive results, preventing unnecessary tests, and improving the detection and diagnosis of LC at the early stages. Recently a case-control study of subjects with suspicious nodules revealed by LDCT imaging was conducted using quantitative methylation-specific real-time PCR and Methylation on Beads for six genes (SOX17, TAC1, HOXA7, CDO1, HOXA9, and ZFP42) (105), selected for the study using The Cancer Genome Atlas (TCGA) as having high prevalence of DNA methylation in lung squamous and adenocarcinoma, but not in normal lung

<table>
<thead>
<tr>
<th>Panel of methylated markers</th>
<th>Detected changes</th>
<th>Source</th>
<th>Clinical application</th>
<th>Method of detection</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 lung-cancer-related genes (KLF6, STK32A, TERT, MSH5, ACTA2, GATA2, VT11A and CHRNA5)</td>
<td>Demonstrated the impact of tobacco smoking on DNA methylation at these genes, associated with lung cancer risk</td>
<td>Whole blood samples (extracted whole blood DNA)</td>
<td>Screening</td>
<td>Illumina Infinium Human Methylation 450 BeadChip array</td>
<td>(108)</td>
</tr>
<tr>
<td>CSF3R, ERCC1</td>
<td>Predict higher risk for SCLC and diagnosis of SCLC</td>
<td>Peripheral leukocytes</td>
<td>Diagnostics</td>
<td>Pyrosequencing</td>
<td>(109)</td>
</tr>
<tr>
<td>CDH1*, p16*, MGMT*, DAPK*</td>
<td>Methylation of CDH1 and DAPK occurs in the early stages lung cancer</td>
<td>Peripheral lymphocytes</td>
<td>Diagnostics</td>
<td>MS-PCR</td>
<td>(100)</td>
</tr>
<tr>
<td>p16*, RASSF1A*, FHIT</td>
<td>The vector machine (SVM) and a decision tree (DT) models including combined detection of p16, RASSF1A and FHIT promoter methylation and relative telomere length (RTL) of white blood cells yielded an AUC 0.670–0.810</td>
<td>Peripheral leukocytes</td>
<td>Screening, diagnostics</td>
<td>SYBR Green-based qMS-PCR, qPCR</td>
<td>(98)</td>
</tr>
<tr>
<td>DAPK*, PAX5b, PAX5a, Dal1, GATA5, SULF2, CXCL14</td>
<td>Demonstrated sensitivity of 75% and a specificity of 68% for cancer diagnosis</td>
<td>Sputum</td>
<td>Diagnostics</td>
<td>MS-PCR</td>
<td>(101)</td>
</tr>
<tr>
<td>SOX17, TAC1, HOXA7, CDO1, HOXA9 and ZFP42</td>
<td>First combination: TAC1, HOXA17 and SOX17</td>
<td>Sputum</td>
<td>Diagnostics</td>
<td>qMS-PCR and methylation on beads for cancer associated genes</td>
<td>(105)</td>
</tr>
<tr>
<td></td>
<td>Second combination: CDO1, TAC1 and SOX17</td>
<td>A three-gene combination of the best individual genes has sensitivity and specificity of 98% and 71%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*, full gene names are presented in Table 1. qPCR, quantitative PCR; qMS-PCR, quantitative methyl-specific PCR; Rt-PCR, real-time PCR; SCLC, small-cell-lung cancer; SHOX17, short stature homeobox 17; PTGER4, prostaglandin E receptor 4; HOXA, homeobox box A cluster; HOXD10, homeobox D10; PAX, paired box; PTPRN2, protein tyrosine phosphatase receptor type N2; STAG3, stromal antigen 3; CDO1, cysteine dioxygenase type 1; KLK, kallikrein; AIM1, absent in melanoma 1; DCC, DCC netrin 1 receptor; GATA, GATA binding protein; SULF2, sulfatase 2; CXCL14, C-X-C motif chemokine ligand 14; CSF3R, colony stimulating factor 3 receptor; ERCC1, ERCC excision repair 1; FHIT, fragile histidine triad; TAC1, tachykinin precursor 1; ZFP42, ZFP42 zinc finger protein; AHRR, aryl-hydrocarbon receptor repressor; F2RL3, F2R like thrombin or trypsin receptor 3; KLF6, Kruppel like factor 6; STK32A, serine/threonine kinase 32A; TERT, telomerase reverse transcriptase; MSH5, mutS homolog 5; ACTA2, actin, alpha 2, smooth muscle, aorta; VT11A, vesicle transport through interaction with t-SNAREs 1A; CHRNA5, cholinergic receptor nicotinic alpha 5 subunit; EPB41L3 (Dal1), erythrocyte membrane protein band 4.1 like 3.
A combination of three best individual genes has sensitivity and specificity of 98% and 71% using sputum and 93% and 62% using plasma. Independent blinded random forest prediction models combining gene methylation with clinical information correctly predicted LC in 91% of subjects using TAC1, HOXA17 and SOX17 in sputum and 85% of subjects using CDO1, TAC1 and SOX17 in plasma. Another signature of four effective markers (HOXD10, PAX9, PTPRN2, and STAG3) was reported by Wielscher et al. [2015] (106) who used a successive selection by genome-wide methylation screening, methylation-sensitive restriction enzyme digestion and enrichment of methylated DNA and parallel qPCR for the panel development. The four-marker model yielded an AUC of 0.85 with sensitivity of 97% and a specificity of 73%.

Methylated SHOX2 has been proposed as a putative ctDNA marker for LC diagnosis in a number of studies with median sensitivity (60%) and high specificity (90%) (51). Three independent case-control studies examined a total of 330 plasma ctDNA specimens for methylated (SHOX2), prostaglandin E receptor 4 gene (PTGER4), and forkhead box L2 gene (FOXL2) and a panel of SHOX2 and PTGER4 and yielded promising results (41). Importantly, a validation study with 172 patient samples demonstrated high performance in distinguishing patients with LC from subjects without malignancy. At a fixed specificity of 90%, sensitivity for LC was 67%; at a fixed sensitivity of 90%, specificity was 73%. At maximum sensitivity, this assay shows a reasonably high specificity that can provide substantial reduction in the false-positive rate of LDCT. If adopted with high specificity, the test may prove be useful as a screening test directing patients to LDCT screening.

Compared to hypermethylation of tumour suppressor genes, hypomethylation of repetitive elements as Alu, LINE-1 (31,33) in cfDNA have attracted less attention and has not been studied thoroughly. Many studies have reported LINE-1 hypomethylation in various cancer tissues, including LC tumours (29). Global genome hypomethylation in LC tissues in known to contribute to genomic instability (31), and was associated with poor outcome (79) and overexpression of aberrant transcripts such as ΔNp73 (79). Based on the data presented by Hoshimoto et al., increase of LINE-1 hypomethylation was observed in the cfDNA from serum of stage III or IV malignant melanoma patients, compared to healthy donors (110). Recent study evaluated the potential of LINE-1 hypomethylation in plasma cfDNA as a blood biomarker for early stage CRC detection (111). Due to the low amount of tumour DNA in the circulation, large volume of plasma is needed to ensure the sensitivity of the ctDNA-based assay; for instance, mSEPT9-based test Epi proColon requires 3.5 mL of plasma. Conversely, with the abundance of LINE-1 copies in the human genome, only 0.5 mL of plasma enables to quantify its absolute methylation level (111). Our group performed MIRA coupled with qPCR-based quantitation to assess the integral methylation level of LINE-1 promoters in the csbDNA of NSCLC patients and healthy controls (Table 1). Deep sequencing of amplicons revealed that

---

**Figure 1** CfDNA methylation assay as a complementary test for lung cancer screening. Green arrows—low-invasive circulating tumour DNA markers assay.

---

© Translational Cancer Research. All rights reserved. tcr.americanres.com Transl Cancer Res 2018;7(Suppl 2):S153-S170
hypomethylation of LINE-1 promoters in csbDNA of LC patients is more pronounced for the human-specific L1Hs family. Statistical analysis demonstrated significant difference in L1Hs promoter methylation between cancer patients and healthy individuals (AUC = 0.69).

Reviewing these and earlier studies, we should stress the urgency for the validation of best performing methylation markers selected from different studies in multicentric international collaborations with larger cohorts, potentially resulting in an optimized gene panel. Adequate representation of histologic subgroups, TNM subclasses for LC, and inclusion of non-malignant and pre-cancerous states, such as chronic obstructive pulmonary disease (COPD), fibrotic interstitial lung diseases (ILDs) and other non-cancer patient groups, are necessary for the success of the validation study. Evaluation of selected markers using universal detection protocols in a single validation study should allow assess their true efficiency and adequately compare the potential of different markers, resulting in the development of valid biomarker panels (79,112).

Monitoring toxicity of chemotherapy and tumour resection efficiency

Early detection of resistance to chemotherapy treatment is imperative for improving patient outcomes, by minimizing the unnecessary toxicity and accelerating the transition to alternative therapy regimens. ctDNA is a promising new approach for monitoring of changes in tumor burden in response to therapy. Elevated ctDNA levels were shown to precede clinical establishment of progressive disease (113), thus ctDNA may provide an early marker of cancers resistance to therapy.

The potential of SHOX2 in monitoring treatment effectiveness in LC patients receiving chemotherapy was evaluated (114). Patients responding to therapy (17/36) showed a decrease of methylated SHOX2 in plasma at the first post-treatment blood draw 7–10 days after administration of chemotherapy, while for non-responders (8/19) the decrease was significantly smaller. These data suggest that methylated plasma SHOX2 is able to identify patients who will benefit from chemotherapy early and with a high accuracy. The methylation level of two other tumour suppressor genes—RASSF1A and RARB2—was earlier shown to be associated with LC development by different groups. RASSF1A and RARB2 methylation indices in cfDNA and csbDNA were found to be indicative of treatment response in non-small-cell LC patients (97) (Table 3). Both RASSF1A and RARB2 showed a decrease in methylation index after 2 courses of neoadjuvant chemotherapy, and then a further decrease following surgery. During the 9 months of after-treatment follow-up, 5 out of 26 patients had a relapse, all of them having the plasma methylation of at least one gene return to before treatment levels, while no patients without recurrence showed an increase in the methylation index (97) (Table 3). Decreased tumor volume was associated with the decrease in methylated ctDNA markers when measured 10–15 days post-therapy which evidenced of tumour chemosensitivity. In contrast, when methylation of APC1 and RASSF1A measured in ctDNA of LC patients immediately prior to the administration of chemotherapy and again 24 h after, when DNA released from dying cells peaks, was compared, an increase in circulating methylated RASSF1A/APC1 was shown to be associated with chemosensitivity and complete or partial response to treatment (103). In a very recent study authors demonstrated that quantitative analysis of total plasma cfDNA and plasma APC/RASSF1A methylation collectively provided a real-time synchronous rapid monitoring indicator of therapeutic outcomes of advanced LC. Four parameters were assessed: methylation level before chemotherapy (meth0 h), methylation level 24 h after chemotherapy (meth24 h), total plasma DNA concentration before chemotherapy (DNA0 h), and total plasma DNA concentration 24 h after chemotherapy (DNA24 h) (103). Meth24 h > meth0 h for at least one gene and DNA24 h/DNA0 h ≤ 2 were defined as criteria for better tumor response and fewer adverse events with a high correct prediction rate (84.7%). These fluctuations underscore the importance of carefully characterizing ctDNA kinetics in response to chemotherapy as a part of bringing the biomarkers to the clinic for patient monitoring.

The persistence of cancer DNA in blood after tumor has been removed likely reflects residual tumor tissue in the body and indicates poor prognosis (36). Tumour-specific methylation is less variable across tumors than mutation therefore their quantification in blood plasma provides a less labour-intensive approach that is more appealing for clinical use. The presence of tumor-specific methylated sequences that have been shown to decrease in the blood of LC patients following surgery is compiled in Table 3.

Methylation markers of cancer prognosis

Gene methylation patterns in tumor tissue can be indicative of tumor aggressiveness and the likelihood of recurrence
after surgical resection and/or chemotherapy due to residual disease (115). Numerous studies have linked methylation of individual genes (116-118) and gene panels (119-121) in cancer tissue with patient survival. Recent study from the group of Prof. Zhang evaluated the utility of DNA methylation signatures for differentiating between tumor and normal tissue for four common cancers (breast, colon, liver, and lung) using machine learning on whole-genome methylation data from TCGA (122). The study demonstrated the potential of using methylation signatures to identify cancer tissue of origin and predict prognosis.

cTCTA has a short half-life (~2 h), and its persistence in the blood following surgery has been linked to poor prognosis (36). Prognostic biomarkers are urgently needed to distinguish patients who are cured with surgery alone, from those at high risk of disease recurrence who may benefit from adjuvant chemotherapy (61). The prognostic significance of gene promoter cDNA methylation has been described in several studies. Li et al. (118) performed a meta-analysis of cohort studies to determine whether promoter methylation of the DAPK gene contributes to the pathogenesis of NSCLC. Subgroup analysis based on sample source discovered that DAPK gene methylation was implicated in the pathogenesis of NSCLC in both blood and tissue subgroups (all P<0.05). Detection of methylated breast cancer metastasis suppressor-1 (BRMS1) and (sex determining region Y)-box 17 (SOX17) in operable and advanced NSCLC, was shown to have a negative impact on survival (37,52). DCLK1 methylation was also associated with shorter survival (27). In contrast, SFN (14-3-3 Sigma) promoter methylation correlated with a reduced risk of death (58) (Table 1).

Conclusions

So far, a number of studies have supported the potential of ctDNAs methylation for the development of the minimally invasive tests for cancer diagnostics. Changes in the methylation of cTNA in the blood plasma and serum can be used as sensitive and specific markers for LC diagnostics, prognosis and monitoring of the response to therapy. However, further detailed studies are required to make potential biomarkers a routine tool in the laboratory medicine/clinical practice. Pre-analytical methods and analysis of cTNA methylation profiles should be standardized, automated and certified in order to enable rapid and reliable detection and quantification. As the sensitivity of analytical techniques increases, detection of low levels of methylated DNA becomes possible, giving us the opportunity to detect early stages of cancer or minimal residual disease. Therefore, LC-specific methylation has the potential to be used as a minimally invasive biomarker in combination with LDCT to direct individuals to LDCT screening and minimize false positive diagnoses. The use of high-throughput and robust methodologies in diagnostic laboratories should advance the selection of the informative biomarkers. Recently, a multicenter benchmark study was carried out by 19 research groups from seven countries, assessing the currently available promising DNA methylation assays for biomarker validation studies (123). The results of the study endorse locus-specific DNA methylation assays as mature technology ready for widespread use in biomarker development and clinical applications. Another recent study declared uniquely mappable WGBS data to be the most reproducible and accurate measurement of global DNA methylation levels (124). WGBS was previously considered too costly for epidemiology studies with large cohorts of samples however, using multiplexing by indexed barcodes the costs of WGBS can be lowered significantly to improve the accuracy of global DNA methylation assessment for human studies. Most candidate cTNA methylation markers are discovered in small retrospective cohorts or case-control studies, and few are validated in independent studies. To achieve the required degree of optimization, cTNA methylation marker, panels and/or signatures should be developed and validated in large prospective cohort studies and screening populations. Clinical trials of cTNA methylation as an early indicator of LC, prognostic marker or target for surrogate DNA demethylating activity are currently ongoing (61). Additional research is needed to explore the pathologic significance of the cfDNA, emerging from tumor cells into the circulation.

Acknowledgements

Funding: The study was supported by the fundamental research program of the Presidium of the Russian Academy of Sciences “Fundamental research for the development of biomedical technologies” (No. 2014-208), the program of the Presidium of the Russian Academy of Sciences “Molecular and Cellular Biology” and the Russian Foundation for Basic Research (No. 17-29-06002).
Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References


79. Chatterjee A, Rodger EJ, Morison IM, et al. Tools and


