BIOMARKER DISCOVERIES AND POTENTIAL BIOMARKERS IN LUNG CANCER: A NARRATIVE REVIEW.

PENEMUAN BIOMARKER DAN POTENSIAL BIOMARKER PADA KANKER PARU: TINJAUAN NARASI

Andika Dwi Mahendra1*, Muhammad Hatta Prabowo1, Aisyah Fitriannisa Prawiningrum2
1 Department of Pharmacy, Faculty of Mathematics and Natural Science, Universitas Islam Indonesia, Daerah Istimewa Yogyakarta, Indonesia.
2 Bioinformatics Core Facilities, Indonesian Medical Education and Research Institute (IMERI), Faculty of Medicine, University of Indonesia, Jakarta, Indonesia
*Email: andikadwimahendra@gmail.com

Abstract

Lung cancer is the second most common cause of death worldwide after breast cancer. This is associated with late diagnosis, which reduces survival rate. Early diagnosis is one approach that may be employed to improve outcome for the patient with lung cancer. When a small level of cancer is present in the body, a biomarker can be utilized as a diagnostic marker. Liquid biopsy is minimal invasive which can be done to obtain multi-biomarkers in lung cancer. Through liquid biopsy, various biomarkers of lung cancer can be acquired and utilized as early diagnosis tools. The types of biomarker and changes of DNA during lung cancer development can be obtained from this liquid such as cell-free DNA (cfDNA), Micro-RNA (miRNA), methylated DNA, and Histone Epigenetics. To explore in-depth, we will discuss about potential biomarkers for lung cancer early diagnosis in this article.

Keywords: lung cancer, biomarker, early diagnosis, potential biomarker

INTRODUCTION

One of the factors limiting life expectancy is cancer, which is the leading cause of mortality worldwide (Sung et al., 2021). The five malignancies with the highest estimated incidence rates, which include cases in both sexes and account for more than 44% of global cases, are female breast cancer (11.7%), lung cancer (11.4%), colorectal cancer (10.0%),
prostate cancer (7.3%), and stomach cancer (5.6%) (Sung et al., 2021). Despite advances in therapy, the majority of malignancies are diagnosed at an advanced stage and have a poor prognosis (Pal et al., 2022).

With 18.4% of all cancer-related mortality worldwide, Lung Cancer (LC) is the most common type of cancer. According to the most recent statistics, 2.2 million new lung cancer cases are expected to be diagnosed globally in 2020 (Ferlay et al., 2021). Lung cancer is typically divided into two main categories: Small Cell Lung Cancer (SCLC; 15-20% of LC cases) and Non-Small Cell Lung Cancer (NSCLC; 80-85% of LC cases). Adenocarcinoma (ADC), which has a 40% prevalence, is the most prevalent histologic subtype of NSCLC (Osmani et al., 2018). Squamous cell lung cancer (SQC) is a subtype of NSCLC, which accounts for 20–30% of all incidents of LC (Hoang and Landi, 2022; Osmani et al., 2018). At the molecular level, it describes and unfolds tumors (Mahindroo et al., 2009; Sawyers, 2008). Although characterization is challenging, the utilization of current and developing molecular biomarkers as a diagnostic and prognostic tool is critical. Some biomarkers are effectively used to diagnose cancer in early stages or as it progresses (Muinao et al., 2018).

Based on the issues above, it is important to conduct a review related to cancer biomarkers, especially in LC. In this article, we will review articles related to discoveries and potential biomarkers in LC.

RESULT AND DISCUSSION
Criteria of Cancer Biomarkers

A biomarker is a measurable characteristic change that can distinguish normal and abnormal conditions in an organism through analysis, such as protein, RNA, DNA, or other modifications of biomolecules (Goossens et al., 2015; Šutić et al., 2021). DNA, metabolites, RNA, transcription factors, and enzymes can be used as cancer biomarkers. These biomarkers can be obtained from cancer cells or the host in response to cancer and be used as an initial diagnosis or cancer prognosis (Pal et al., 2022). In the clinical field, biomarkers are very useful in diagnosis, prediction, survival and outcome, response to treatment, and prognosis (Khomiak et al., 2020; Townsend et al., 2018; Verma et al., 2011). Diagnostic biomarkers can detect and differentiate one cancer from others or one subtype from others (Šutić et al., 2021). Predictive biomarkers make it possible to know the patient's response to a given therapy. If a biomarker is used to determine progression and irrespective of treatment, it is called a prognostic biomarker (Goossens et al., 2015; Šutić et al., 2021). There are several ways to determine the development of cancer through several changes. These changes are genetic mutations, changes in gene expression, changes in post-translational proteins, and changes in metabolites. These changes will lead to uncontrolled proliferation of cancer cells, inhibition of cancer cell apoptosis, changes in metabolism, and the formation of angiogenesis (Dlamini et al., 2021; Kamel and Al-Amodi, 2016).

The main challenge in cancer treatment is early detection of cancer diagnosis in order to increase drug efficacy and patient survival. Molecular biomarkers are one of the tools that can be utilized to diagnose and estimate the prognosis. It is expected that the treatment given is in accordance with the patient's condition. Early cancer detection is made possible by several discovered biomarkers. (Das et al., 2017; Lu et al., 2018). Ideally, cancer biomarkers
should be able to differentiate between cancer from benign and aggressive tumors from insignificant ones as well as have high specificity and sensitivity (2001; Levenson, 2004). Various criteria that can be taken into consideration in relation to potential cancer biomarkers include short half-life biomarker to reflect actual changes in any stage of cancer appears at low levels in early cancer development closely linked to the prevalence and stage of cancer high specificity and sensitivity can be acquired using non-invasive techniques cost-effective validation, standardization, quantitative, and reproducible checks can be performed distinguishes metastases and non-metastasis (Dlamini et al., 2021; Kamel and Al-Amodi, 2016; Ray et al., 2010; Verma et al., 2011)

As there is a tumor or the body's response to a tumor occurs, cancer biomarkers will be released in the body through serum, plasma, blood, urine, sputum, and CSF. This makes it possible to perform non-invasive procedures and periodically (Kamel and Al-Amodi, 2016). Liquid biopsy is one approach that can be employed to obtain multi-biomarkers in cancer (Herath et al., 2022). This procedure has the advantage of being minimally invasive, capable of collecting systemic changes to provide disease progression, including local primary tumor and distant metastases, treatment responses across many sites, and the capability to capture tumor heterogeneity across time (Barefoot et al., 2021; Chen et al., 2017). Extracellular vesicles, ctDNA, miRNA, mRNA, methylated DNA, protein, and circulating tumor cells are some of the biomarkers that can be utilized to identify cancer (Kamel and Al-Amodi, 2016; Pal et al., 2022; Pudil et al., 2020).

**Types of biomarker in lung cancer cell-free DNA (cfDNA)**

cfDNA is a potential biomarker for cancer diagnosis and prognosis (Sánchez-Herrero et al., 2022), it is usually found in healthy patients with concentrations of 0 - 100 nanograms/mL and rises to more than 1000 ng/mL in cancer patients (Schwarzenbach et al., 2011; Thierry et al., 2016). cfDNA consists of small fragments of mitochondrial DNA (mtDNA) and double strands that are coding or non-coding with a length of about 40-200 base pairs (Kustanovich et al., 2019; Schwarzenbach et al., 2011). The proportion of cfDNA in the blood is closely linked to the severity of cancer. cfDNA generated from cancer tissue is called ctDNA which comes from either a cell death or non-cell death mechanism (Amintas et al., 2021; Chen and Zhao, 2019; Garcia-Saenz et al., 2017; Provencio et al., 2018, 2017). Gene mutations, chromosomal rearrangements, copy number aberrations, methylation, DNA fragment lengths, gene expression, and virus sequence identification are some features of using cfDNA (Keller et al., 2021).

The length of the cfDNA fragment indicates the mechanism involved in generating cfDNA. Fragment length <200bp is possible due to apoptosis (Mouliere et al., 2011) and <100 bp is possible for ctDNA and mtDNA produced by tumors (Vagner et al., 2018). Higher ctDNA concentrations were seen in patients with melanoma, lung cancer, and metastatic colorectal cancer compared with healthy individuals (El Messaoudi et al., 2016). Base length >200 bp originates from the mechanism of necrosis, but in the process of necrosis, there are 2 processes that may cause shorter fragment lengths. The first mechanism is the DNA of cells undergoing necrosis will be degraded by DNase and the necrotic cells will be engulfed by macrophages resulting in shorter fragment lengths (Jiang et al., 2003; Stephan et al., 2014).
The estimated correlation of the length of the fragment with the source of the mechanism that occurs can be one of the markers of the presence of cancer in the body. In a normal body, cell death occurs through the process of apoptosis, but in cancer cells, it can occur through other mechanisms such as necrosis or autophagy (Leon et al., 1977). The autophagy process is activated by cancer cells to produce alternative energy so that ctDNA will eventually be formed (Sato et al., 2007).

Several studies have been carried out using cfDNA in diagnosing several cancers such as thyroid cancer (Salvianti et al., 2017), colorectal cancer (Bhangu et al., 2017), hepatocellular carcinoma (Chan et al., 2008), hematological malignancies (Quan et al., 2015; Zorofochian et al., 2018), nasoparingeal carcinoma (Chan et al., 2017), and breast cancer (Li et al., 2020). cfDNA fragments are not a non-random mechanism. cfDNA fragment is part of the genetic sequence in a tissue of origin (Jiang et al., 2018). In addition, ctDNA base sequence obtained can be used as a diagnostic by matching it with the cfDNA fragment of tumor cells and comparing it with the base sequence of healthy human DNA (Barefoot et al., 2021; Serpas et al., 2019). The quantity of ctDNA has been associated with the size, stage, and metastasis of the tumor (Sánchez-Herrero et al., 2022). ctDNA profiling can be done using NGS in combination with optimizing library preparation and bioinformatics even at low levels of ctDNA. This combination of the technologies will improve ctDNA analysis as a cancer biomarker (Lin et al., 2021; Sánchez-Herrero et al., 2022).

**Micro-RNA (miRNA)**

Extracellular vesicles (EVs) are a source of potential cancer biomarkers. Blood, urine, mucus or bronchial fluid, amniotic fluid, and saliva are all possible sources of EVs. EV is a bilayer particle secreted by cells and contains DNA, RNA, and protein. Cancer cells also produce EVs in the pathogenesis process (Möller and Lobb, 2020; Pal et al., 2022). It contributes to angiogenesis, metastasis, cell immortality, proliferation, and migration (Möller and Lobb, 2020). Several findings of cancer biomarkers obtained through EVs (miRNA, non-long coding RNA, mRNA) such as miR-21 in esophageal cancer (Tanaka et al., 2013), CD24, EpCAM, TGF-beta1 in ovarian cancer (Magdalena Derbis, 2012; Runz et al., 2007), miR-4772-3p in colon cancer (Liu et al., 2016), LRG1 in lung cancer (Li et al., 2011), and fibronectin in breast cancer (Moon et al., 2016). To support the discovery of these biomarkers, the use of technology such as NGS is recommended. NGS allows for analysis of gene changes including copy number variation, a fusion of multiple genes, and mutations (Pruneri et al., 2021).

MicroRNA (miRNA) is a part of EV. miRNA is a single-stranded, non-coding RNA with an arrangement of about 22 nucleotides (Chen et al., 2016). miRNA is expressed from the gene strand by RNA polymerase II which produces pri-miRNA. miRNA is partially cleaved by DGC58/Drosha protein being pre-miRNA which is 70-100 nucleotides long. Next, the pre-miRNA is released via Exportin 5 protein to the cytoplasm and is bound to Dicer which functions to cut the hairpin loop on the pre-miRNA. This cleavage produces a miRNA double strand 20–25 nucleotides long. The miRNA double strands are then converted into single strands by helicase. miRNA will bind to Argonautes proteins (AGO1 and AGO2) resulting in the formation of the miRNA-RISC complex (RNA-induced silencing complex). This complex aims to stabilize when miRNA binds with mRNA which will regulate post-translational gene
expression (Iorio and Croce, 2012). Differentiating between cancer and normal tissue, identifying the tendency for cancer development, distinguishing cancer subtypes, and identifying the origin of cancer development are some applications of miRNA in the diagnosis of cancer (Paranjape et al., 2009).

In the human body, there are over 1000 miRNAs encoded by more than 3% of the human genome (Bentwich et al., 2005). miRNA can regulate post-transcriptional gene expression (Kosaka et al., 2010; Paranjape et al., 2009; Qu et al., 2015, p. 25) by providing the target RNA 3' untranslated region (UTR) (Ghosh et al., 2016; Shah et al., 2016). miRNA has biological functions such as cell proliferation, apoptosis, and cell growth (Bartel, 2004; V. N. Kim et al., 2009). Dysregulations of miRNA formation cause cancer in humans. This can lead to the stimulation of cell proliferation, overexpressed genes, metastasis, tumor-suppressing genes, and angiogenesis associated with oncomiRs (Svoronos et al., 2016). OncomiRs function in cancer cell growth (tumorgenesis) is preventing cancer cell apoptosis, increasing metastatic ability, cell proliferation, inhibiting T cell activity, and reducing NK cell activity as an antitumor (Otmani et al., 2022).

miR-21 is a oncomiRs which has functions in cancer development in several cancers such as colorectal cancer, gastric cancer, prostate cancer, breast cancer, pancreatic cancer, and lung cancer (Menon et al., 2022; Wu et al., 2021). In cases of colorectal cancer and gastric cancer, the level of this biomarker shows the severity and progression of the disease and can be utilized as a prognostic marker (Chen et al., 2016; Simonian et al., 2018; Toiyama et al., 2013). Upregulated miR-181 is formed in oral squamous cell carcinoma, breast, prostate, and gastric cancer. An increase in this biomarker is associated with lymph node metastasis, poor survival rate, and angiogenesis of oral squamous cell carcinoma (Yang et al., 2011). Some examples of miRNAs which play a role in tumorgenesis activation such as miR-125, miR-155 (sustained angiogenesis), miR-200, miR-21 (tissue invasion and metastasis), miR-21, miR-17-92 (evading apoptosis), miR -128, miR-221/222 (limitless replication potential), miR-23b, miR-155 (tumor-promoting inflammation), and miR-25, miR-17 (insensitive to anti-growth signals) (Menon et al., 2022).

Methylated DNA

Epigenetics is the process of chemical changes to the nucleic acids in DNA in regulating gene expression which occurs during the process of cell development or proliferation without changing the DNA sequence (Das and Singal, 2004; Xue et al., 2020). DNA methylation is one of the most common epigenetic processes that can be observed in human plasma (Juergens et al., 2011). This process occurs with the addition of a methyl (CH₃) to cytosine at the fifth carbon. Most of the methylation occurs at the 5' CG3' sequence (also called CpG), sometimes also occurring at CpA and CpT (Xue et al., 2020). However, DNA methylation changes can point to a connection with the development of cancer, which can be used for both diagnosis and prognosis (Müller and Györffy, 2022).

One of the biomarkers used in oncology is DNA methylation (Leygo et al., 2017; Tang et al., 2016). It will have an impact on the expression of essential genes to the proliferation of cancer (Gagliardi et al., 2018). Hypomethylation that occurs in the whole genome as well as
hypermethylation in specific genes causes the growth of malignant tumors (Cao et al., 2020; Tang et al., 2016).

### Table 1. miRNA biomarkers in SCC

<table>
<thead>
<tr>
<th>Patient</th>
<th>Specimen</th>
<th>miRNA</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCC</td>
<td>Plasma</td>
<td>miR-20a</td>
<td>97</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-223</td>
<td>95</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-21</td>
<td>91</td>
<td>90</td>
<td>(Geng et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-155</td>
<td>92</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-145</td>
<td>95</td>
<td>95</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2. Significant upregulation of miRNA in NSCLC compared to control

*P-value, significant*

<table>
<thead>
<tr>
<th>miRNA</th>
<th>miRNA regulation</th>
<th>Log (fold change) (NSCLC vs Control)</th>
<th>P value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-22</td>
<td>Upregulated</td>
<td>0.25 ± 0.18 vs -0.06 ± 0.02</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-423-5p(st)</td>
<td>Upregulated</td>
<td>6.41 ± 0.22 vs 2.50 ± 0.30</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-200b</td>
<td>Upregulated</td>
<td>0.73 ± 0.11 vs -0.05 ± 0.03</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-221</td>
<td>Upregulated</td>
<td>0.20 ± 0.04 vs 0.12 ± 0.02</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-340</td>
<td>Upregulated</td>
<td>0.45 ± 0.02 vs 0.12 ± 0.03</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-203</td>
<td>Upregulated</td>
<td>0.15 ± 0.02 vs 0.05 ± 0.03</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-15b-5p</td>
<td>Upregulated</td>
<td>3.33 ± 0.82 vs 1.32 ± 0.12</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-141</td>
<td>Upregulated</td>
<td>0.32 ± 0.07 vs -0.16 ± 0.02</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-193b</td>
<td>Upregulated</td>
<td>0.24 ± 0.04 vs -0.06 ± 0.02</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-301</td>
<td>Upregulated</td>
<td>0.15 ± 0.02 vs -0.05 ± 0.03</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-210</td>
<td>Upregulated</td>
<td>0.85 ± 0.03 vs 0.55 ± 0.02</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-19a</td>
<td>Upregulated</td>
<td>0.25 ± 0.06 vs -0.10 ± 0.12</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-15a</td>
<td>Down-regulated</td>
<td>-0.01 ± 0.002 vs 0.20 ± 0.25</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-195</td>
<td>Down-regulated</td>
<td>-0.04 ± 0.03 vs 0.22 ± 0.03</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-486-5p</td>
<td>Down-regulated</td>
<td>2.40 ± 0.16 vs 4.49 ± 0.14</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-20b</td>
<td>Down-regulated</td>
<td>-0.19 ± 0.10 vs 0.31 ± 0.07</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-126</td>
<td>Down-regulated</td>
<td>-0.06 ± 0.03 vs 0.24 ± 0.06</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-98</td>
<td>Down-regulated</td>
<td>-0.05 ± 0.03 vs 1.31 ± 0.14</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-146a</td>
<td>Down-regulated</td>
<td>0.05 ± 0.03 vs 0.55 ± 0.02</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-20a-5p</td>
<td>Down-regulated</td>
<td>-0.18 ± 0.017 vs 0.19 ± 0.06</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-17b-5p</td>
<td>Down-regulated</td>
<td>2.40 ± 16 vs 3.05 ± 0.02</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-16-5b</td>
<td>Down-regulated</td>
<td>0.03 ± 0.01 vs 0.22 ± 0.01</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-19-3p</td>
<td>Down-regulated</td>
<td>-0.02 ± 0.012 vs 2.17 ± 0.07</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>miR-339-5p</td>
<td>Down-regulated</td>
<td>0.06 ± 0.02 vs 0.53 ± 0.29</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>let-7a</td>
<td>Down-regulated</td>
<td>2.42 ± 0.37 vs 5.16 ± 0.09</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>let-7c</td>
<td>Down-regulated</td>
<td>3.05 ± 0.03 vs 6.92 ± 0.01</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>let-7d</td>
<td>Down-regulated</td>
<td>1.52 ± 0.19 vs 4.77 ± 0.23</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>let-7e</td>
<td>Down-regulated</td>
<td>3.80 ± 0.04 vs 7.41 ± 0.82</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>let-7f</td>
<td>Down-regulated</td>
<td>-0.04 ± 0.03 vs 0.34 ± 0.08</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>let-7g</td>
<td>Down-regulated</td>
<td>0.08 ± 0.03 vs 0.55 ± 0.03</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td>let-7i</td>
<td>Down-regulated</td>
<td>0.37 ± 0.05 vs 0.93 ± 0.03</td>
<td>0.001*</td>
<td></td>
</tr>
</tbody>
</table>
Epigenetics in cancer occurs with the addition of methyl in DNA which causes interference in the process of cell development, genomic stability, and transcription processes. These DNA methylation aberrations promote cancer development and progression (Galamb et al., 2016). In the process of carcinogenesis, methylation on the CpG promoter is mediated by DNA methyltransferase which causes suppression of Tumor-Suppressor Genes (TSGs) (Grady et al., 2021). In a systematic review, the process of promoter hypermethylation in CpG in the early development of colorectal cancer can be used as an initial screening, prognosis, and prediction of cancer recurrence (Rasmussen et al., 2016).

Hypermethylation in the promoter area increases the progression of cancer cells. Hypermethylation in the promoters of APC, SFRP1, SFRP2, SFRP5, WIFI, DKK3, ITIH5, and RASSF1A is associated with breast cancer growth (Feng et al., 2010) and epigenetics in APC and RASSF1A genes are biomarkers in early detection of breast cancer (Cao et al., 2018; Coyle et al., 2007; Kloten et al., 2013). Hypermethylation is seen in the PTGER4 and ZNF43 promoters in colorectal patients compared to normal tissue leading to shorter survival followed up to 5 years (Chen et al., 2022). In another research, cancer growth can be characterized by hypermethylation of the TSGs (Baylin and Jones, 2011; Esteve-Puig et al., 2020; Hon et al., 2012; Tavares et al., 2022) in the CAV1 (Li et al., 2015, p. 48), CDH13 (Feng et al., 2010), ID4 (Umetani et al., 2005), and SCGB3A1 (Benevolenskaya et al., 2016) genes. Currently, there are 2 FDA-approved diagnostic biomarkers based on the DNA methylation process for colorectal cancer. The two biomarkers are SEPT9 and a combination of NDRG4 and BMP3 (Müller and Győrffy, 2022).

Table 3. Methylated DNA region causing cancer progression (Liang et al., 2021)

<table>
<thead>
<tr>
<th>Methylation region</th>
<th>Application</th>
<th>Methylation level</th>
<th>outcomes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS3ST2</td>
<td>Progression and metastasis</td>
<td>Increase</td>
<td>NSCLC tends to metatasize</td>
<td>(Hwang et al., 2013)</td>
</tr>
<tr>
<td>TMEM88</td>
<td>Progression and metastasis</td>
<td>Increase</td>
<td>NSCLC tends to metatasize</td>
<td>(Rongna et al., 2017)</td>
</tr>
<tr>
<td>DAL-1</td>
<td>Progression and metastasis</td>
<td>Increase</td>
<td>NSCLC tends to metatasize</td>
<td>(Zhang et al., 2012)</td>
</tr>
<tr>
<td>ELMO3</td>
<td>Progression and metastasis</td>
<td>Increase</td>
<td>NSCLC tends to metatasize</td>
<td>(Søes et al., 2014)</td>
</tr>
<tr>
<td>HMLH1</td>
<td>Prognosis</td>
<td>Increase</td>
<td>NSCLC rises against resistance</td>
<td>(Wu et al., 2015)</td>
</tr>
<tr>
<td>IGFBP-3</td>
<td>Prognosis</td>
<td>Increase</td>
<td>NSCLC rises against resistance</td>
<td>(Pernía et al., 2014)</td>
</tr>
<tr>
<td>RASSF1A</td>
<td>Prognosis</td>
<td>Increase</td>
<td>Patients with NSCLC may benefit more from gemcitabine.</td>
<td>(Fischer et al., 2007)</td>
</tr>
<tr>
<td>MGMT</td>
<td>Prognosis</td>
<td>Increase</td>
<td>Patients with NSCLC may benefit more from Temozolomide.</td>
<td>(Hiddinga et al., 2022)</td>
</tr>
<tr>
<td>TMEM 196</td>
<td>Prognosis</td>
<td>Increase</td>
<td>Low survival rates and a poor prognosis are present.</td>
<td>(Liu et al., 2018)</td>
</tr>
<tr>
<td>HERC5</td>
<td>Prognosis</td>
<td>Increase</td>
<td>Low survival rates and a poor prognosis are present.</td>
<td>(Wrage et al., 2015)</td>
</tr>
<tr>
<td>GRK6</td>
<td>Prognosis</td>
<td>Increase</td>
<td>The metastasis of LUAD cells</td>
<td>(Yao et al., 2019)</td>
</tr>
<tr>
<td>FAM83A</td>
<td>Prognosis</td>
<td>Decrease</td>
<td>The prognosis for patients with LUAD is poor.</td>
<td>(Yu et al., 2020)</td>
</tr>
</tbody>
</table>
Histone Epigenetics

DNA is packaged as chromatin, which contains nucleosomes, in eukaryotic cells. Each nucleosome is made up of 147 base pairs of DNA wrapped around an octamer of four core histones (H3, H4, H2A, and H2B) (Bannister and Kouzarides, 2011). A variety of post-translational modifications (PTMs) are concentrated in the N-terminal tails of molecules. PTMs can also appear on the lateral surface of the histone nucleosome core areas which are in touch with the DNA (Lawrence et al., 2016). Modification of the tail and core of histones will change the structure of chromatin through changes in the charge of histones or inter-nucleosomal interactions (Bannister and Kouzarides, 2011). Histone acetylation is a fundamental epigenetic modification that regulates gene expression by determining whether a gene is accessible to transcriptional factors or not. Histone Acetyltransferases (HATs) and Histone Deacetylases (HDACs) are two sets of enzymes that work together to regulate histone acetylation (Gujral et al., 2020).

Histone tail methylation and acetylation (H4K5/8/12/20/31/79, H3K4/9/14/23/27/36/56/64/79/122, H2AK5/13, and H2BK5/46/108) are the most frequent epigenetic processes. Some amino acids, such as H4K8me/ac/pr/bio/cr, H4K12me/ac/pr/bu/bio, H4K5me/ac/pr/bu/la, H3K18me/ac/la/cr, H3K9me/ac/cr, H3K14me/ac/pr/bu, and H2AK13me/ac/ar/bio, can be changed into more than 2 type of epigenetics (Yang et al., 2022). Histones with the functions of activating transcription, such as H3K4me3 and H3K36me3, as well as those with roles in gene suppression, such as H3K27me3, H3K9me2/3, and H4K20me3, are being well-studied in the context of cancer development (Zhao and Shilatifard, 2019).

The methylation of histones is carried out by the enzyme Arginine Methyltransferases (PRMTs) and Lysine Methyltransferases (KMTs). Lysine (K) and arginine (R) in the histone tails are modified by the addition of methyl group from S-adenosyl methionine (SAM) throughout this process. As opposed to this, Lysine Demethylases (KDMs) function by removing methyl groups from histones (Saleh et al., 2020). Histone-binding proteins can bind to histones or prevent them from binding, which results in histone methylation or not. The H3K4me3 promoter gene will get methylated when transcription factors, a protein activator, are bound to it. Repressor proteins, such as the nucleosome remodeling and deacetylase (NuRD) complex, can bind to this process to stop it (Bannister and Kouzarides, 2011; Xhemalce et al., 2011). The instability of the genome may be impacted by abnormal histone modifications, which can also change gene expression patterns and result in a variety of disorders including cancer (Akman and Erkek-Ozhan, 2022; Yang et al., 2022). In cancer cells, changes in the genes or histone modifier enzymes are lead to the growth and metastasis of cancer. The methylation pattern of the chromatin will change as a result of this. According to the analysis of The Cancer Genome Atlas (TCGA) database, various enzyme mutations related to the methylation process cause the development of cancer (Dawson and Kouzarides, 2012; Sharma et al., 2021).

Acetylation and deacetylation of histones are other epigenetic processes. Histone Acetyltransferases (HATs) or histone deacetylases (HDACs) activity is involved in this process during acetyl attachment and removal, respectively. DNA replication, RNA transcription, and DNA damage repair are all impacted by this process (Chan and Maze, 2020; Choudhary et al.,
P300 and CBP, components of HATs, act as tumor suppressors in hematological malignancies (Rebel et al., 2002). HATs are also found in patients with breast, lung, colorectal, gastric, ovarian, and pancreatic cancer cases which have mutations (Bryan et al., 2002; Kishimoto et al., 2005; Muraoka et al., 1996). Currently, there are 4 drugs that have been approved by the FDA in inhibiting HDAC in T cell lymphoma and multiple myeloma patients. These drugs are vorinostat, istodax, beleodap, and panobinostat (Jones et al., 2016).

In Non-Small Cell Lung Cancer (NSCLC), the histone methyltransferase G9a-mediated H3K9me2 binds transcription factors HP1 and DNMT1 to suppress the expression of TSGs such APC2 and WIF1. This results in Wnt activation and cancer development. This cancer-promoting impact was reversed when G9a was targeted, indicating that G9a may be a useful therapeutic target for the treatment of lung cancer (Zhang et al., 2018). G9a (histone methyltransferase) is an aggressive lung Tumor-Propagating Cells (TPCs). G9a depletion during carcinogenesis enriches tumors in TPCs and accelerates disease progression metastasis. This leads to a decrease in H3K9me2, which allowed the expression of the genes of MMP10, KRAS, and ECM to promote lung adenocarcinoma (Rowbotham et al., 2018).

Next-Generation Sequencing (NGS) in molecular diagnostic

Genomic technologies such Next-Generation Sequencing (NGS) are one of the technologies which can be employed in the detection and treatment of cancer (Chen and Zhao, 2019). NGS is an advanced technology used to perform massive parallel sequencing to determine gene sequences. This technology has the ability to detect genetic changes/abnormalities in a very long gene strand or whole gene, in a cost-effective way (Zhao et al., 2015). NGS analysis can be used to conduct DNA sequencing (DNA-seq) and RNA sequencing (RNA-seq). Application of RNA-seq makes it possible to determine changes in gene expression, detect post-transcriptional modifications, identify small and long noncoding RNAs, and nucleotide mutations. On DNA-seq includes targeted sequencing, whole-genome, and whole-exome (Serrati et al., 2016). The information obtained by the NGS can be used as a diagnosis, decision enforcement in providing personalized therapy, or to analyze biological process between DNA and transcription factors (TSs) in a disease (Mundade et al., 2014; Tkachev et al., 2020). These processes can be performed by combining NGS and bioinformatics to obtain accurate and consistent results which can be applied in clinical decisions (Kanzi et al., 2020) through RNA-seq analysis, DNA-seq analysis (Buzdin et al., 2020) or ChIP-seq (Muhammad et al., 2019). In addition, optimization of gene panels is a choice to get maximum outputs (Japanese Society of Medical Oncology et al., 2021).

Nowaday, the use of NGS and informatics analysis in clinical oncology applications has become increasingly popular by speeding up early diagnosis and shifting traditional cancer therapy to the direction of precision medicine through pharmacogenomics (Hussen et al., 2022). The combination of both approaches can be employed to treat patients with cancer to obtain personalized therapy which extended the patient's survival (Poddubskaya et al., 2019).
<table>
<thead>
<tr>
<th>Histones</th>
<th>Tail modification</th>
<th>Modifiers</th>
<th>Binding factors</th>
<th>Functions</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2A</td>
<td>K119Ub1</td>
<td>RING1A/B</td>
<td>BAP1, USP16, USP21, USP22</td>
<td>Transcriptional repression</td>
<td>(J. Kim et al., 2009; Wood et al., 2003)</td>
</tr>
<tr>
<td>H2AX</td>
<td>S139p(gamma H2AX)</td>
<td>ATM, ATR, DNA-PPK,</td>
<td>Wip1, PP6, AND PP4</td>
<td>DNA repair</td>
<td>(Hake et al., 2005)</td>
</tr>
<tr>
<td>H2B</td>
<td>K120Ub1</td>
<td>RNF20, RNF40</td>
<td>USP3, USP7, USP22</td>
<td>Transcriptional activation, DNA damage response</td>
<td>(Zhu et al., 2005)</td>
</tr>
<tr>
<td>H3</td>
<td>K4me1/2/3</td>
<td>SET1A/B, MLL 1-4</td>
<td>KDM1A/B, KDM2A/B, KDM5A/B/C/D</td>
<td>Transcriptional activation</td>
<td>(Meeks and Shilatifard, 2017; Shilatifard, 2012; Sze and Shilatifard, 2016)</td>
</tr>
<tr>
<td>K9me1/2/3</td>
<td>Suv39H1/2, G9a, GLP, SETDB1</td>
<td>KMD1A, KMD3A/B, KMD4A/B/C/D</td>
<td>HP1, ATRX</td>
<td>Transcriptional activation (K9me1), repression (K9me2/3), X-inactivation and imprinting (K9me2)</td>
<td>(Clos et al., 2008)</td>
</tr>
<tr>
<td>S10p</td>
<td>Aurora B, MSK/RSK/Jil-1</td>
<td>PP2A, PP1</td>
<td>14-3-3ζ</td>
<td>Mitosis, meiosis, transcriptional activation</td>
<td>(Prigent and Dimitrov, 2003)</td>
</tr>
<tr>
<td>R26me2</td>
<td>CARM1</td>
<td>PADI4</td>
<td>HDACs</td>
<td>Transcriptional activation</td>
<td>(Litt et al., 2009)</td>
</tr>
<tr>
<td>K27ac</td>
<td>CBP/p300</td>
<td>HDACs</td>
<td>BRDs</td>
<td>Transcriptional activation</td>
<td>(Wang and Filippakopoulos, 2015)</td>
</tr>
<tr>
<td>K27me1/2/3</td>
<td>EZH1/2</td>
<td>KDM6A/B, KDM4A/B/C/D, JHDM1A</td>
<td>EED, PC, CBX7</td>
<td>Transcriptional activation (K27me1); Transcriptional silencing, X-inactivation, bivalent genes / gene poising (K27me2/3)</td>
<td>(Margueron and Reinberg, 2011)</td>
</tr>
<tr>
<td>K36me1/2/3</td>
<td>NSD1-3, SETD2/3, SETMAR, SMYD2</td>
<td>KMD2A/B, KDM4A/B/C/D, JHDM1A</td>
<td>ZYMND11, Phf19, LEDGF</td>
<td>Transcriptional elongation, repression, DNA repair</td>
<td>(Clos et al., 2008)</td>
</tr>
<tr>
<td>K79me1/2/3</td>
<td>DOT1L</td>
<td>unknown</td>
<td>P53BP1</td>
<td>Transcriptional activation</td>
<td>(Nguyen and Zhang, 2011)</td>
</tr>
<tr>
<td>H4</td>
<td>K20me1</td>
<td>PR-Set7</td>
<td>LSD1n</td>
<td>CRB2, p53BP1</td>
<td>Transcriptional activation</td>
</tr>
<tr>
<td>K20me2/3</td>
<td>SUV4-20H1/2</td>
<td>LSD1n, DPY-21</td>
<td>CRB2, p53BP1, JMJD2</td>
<td>Transcriptional silencing, Heterochromatin</td>
<td>(Brejc et al., 2017; Lu et al., 2008)</td>
</tr>
<tr>
<td>K16ac</td>
<td>MOF</td>
<td>HDACs, Sirt2</td>
<td>BRDs</td>
<td>Transcriptional activation, DNA repair</td>
<td>(Creighton et al., 2010; Seligson et al., 2009)</td>
</tr>
</tbody>
</table>
Another study which combined NGS and DNA-seq/RNA-seq analysis resulted detection of significant mutation genes (KRAS, TP53, EGFR, PIK3CA, BRAF, NRAS, JAK3, CTNNB1, CDKN2A) in patients with Non-Small Cell Lung Cancer (NSCLC) (Zhao et al., 2015). On further research, incorporation NGS and RNA-seq/ChIP-seq analysis in patients with lung cancer, makes it possible to obtain biomarkers in the disease (Ong et al., 2021).

Numerous researches are employed to find oncology-related biomarkers. Patients with Esophageal Squamous Cell Cancer (ESCC) have undergone RNAseq research to assess the prognosis of biomarkers. Patients with ESCC have elevated levels of the Non-Coding RNA (NC) LINC01614, which is an indicator of a poor prognosis. This NC RNA suppression dramatically reduces cancer to metastasis (Tang et al., 2020). hsa_circ_0007843 was discovered in SW480 cells taken from colon cancer patients. A circular RNA called hsa_circs0007843 interacts with miRNA miR-518-5p on matrix metallopeptidase 2 (MMP2). Increased expression of NC RNA promotes tumor cell proliferation, invasion, and metastasis (He et al., 2020). After cancer has spread to other organs, DNA-seq can also be utilized to make a diagnosis based on the primary origin. Using clinical imaging, histology, and tumor markers couldn’t support to diagnosis of a patient with thyroid cancer and lung adenocarcinoma of uncertain cancer origin. After DNA-seq analysis, there was a G12C mutation in the KRAS gene in both cancers, but the incidence of the mutation was higher in lung cancer. Therefore, cancer originated from the lung (Matrone et al., 2019).

Conclusion

Even with the use of advanced therapy, cancer has a poor prognosis when it is diagnosed too late. The first step in optimizing treatment outcomes and survival rates is early cancer diagnosis. When exploring epigenetic regulators or potential biomarkers in cancer for the development of future biomarkers, it is essential to comprehend the cellular mechanism in particular tumor populations in-depth. Biomarkers are a tool that can be used to diagnose cancer earlier and used for personalized therapy and response to therapy.

References


Hiddinga, B., Zwaenepoel, K., Janssens, A., Van Meerbeeck, J., Pauwels, P., 2022. Are anaplastic lymphoma kinase (ALK) and O6-methylguanine-DNA methyltransferase (MGMT) promoter methylation driver biomarkers of pulmonary neuroendocrine tumors (NETs) and carcinomas (NECs)? Oncotarget 13, 800–809. https://doi.org/10.18632/oncotarget.28240


