REVIEW



Present progress in biomarker discovery of endometrial cancer by multi-omics approaches

Yuhao An^{2*†}, Quanxin Feng^{1†}, Li Jia³, Xinrui Sha², Tuanjie Zhang², Linlin Lu⁴, Rui Wang^{2*} and Bin Bai^{1*}

Abstract

Endometrial cancer (EC), a prevalent and intricate disease, is associated with a poor prognosis among gynecological malignancies. Its incidence rising globally underscores the urgent need for biomarkers detection in both research and clinical settings. Over the past decade, we've witnessed rapid advancements in biological methodologies and techniques. A multitude of omics technologies, encompassing genomic/transcriptomic sequencing and proteomic/ metabolomic mass spectrometry, have been extensively employed to analyze both tissue and liquid samples derived from EC patients. The integration of multi-omics data has not only broadened our understanding of the disease but also unearthed valuable biomarkers specific to EC. This review encapsulates the recent progress and future prospects in the application of multi-omics technologies in EC research, emphasizing the potential of multi-omics in uncovering novel biomarkers and enhancing clinical assessments.

Keywords Endometrial cancer, Biomarker discovery, Multi-omics

[†]Yuhao An and Quanxin Feng are Co-first authors.

*Correspondence: Yuhao An anyh@szbl.ac.cn Rui Wang wangrui@szbl.ac.cn Bin Bai baibin513127@163.com ¹Department of Gastrointestinal Surgery, Xijing Hospital of Digestive Diseases, Xijing Hospital, Fourth Military Medical University, Xi'an, Shanxi 710032, China ²Pingshan Translational Medicine Center, Shenzhen Bay Laboratory, Shenzhen, Guangdong 518020, China ³Faculty of Pharmacy and Pharmaceutical Sciences Monash University

³Faculty of Pharmacy and Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia

⁴Joint Laboratory for Translational Cancer Research of Chinese Medicine of the Ministry of Education of the People's Republic of China, International Institute for Translational Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong 510006, China

Introduction

Endometrial cancer (EC) is a type of cancer that originates from the inner epithelial layer of the uterus, which is rising in both incidence and associated mortality [1, 2]. Worldwide, approximately 420,368 women are diagnosed with endometrial cancer each year, and it is estimated that this cancer results in the death of about 97,723 women (https://www.wcrf.org/cancer-trends/ endometrial-cancer-statistics/). When considering all stages collectively, the overall survival rate for 5 years is approximately 80% [3]. Depending on the clinical symptoms of the endocrine-metabolic condition and hyperestrogenic manifestations, Bokhman systematically classified endometrial cancer into two categories for the first time in 1983 [4, 5]. Type I, also known as endometrioid endometrial carcinoma (EEC), is the most common type, accounting for about 80% of all cases. It has low grades and a favorable prognosis, associated with excess estrogen, obesity, and endometrial hyperplasia. On the



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

opposite, "Type II", also named uterine serous carcinoma, is a higher-grade non-endometrioid tumor, accounting for about 10%, associated with atrophic endometrium, exhibiting more aggressive behavior and has a worse prognosis [6]. Endometrial clear cell carcinoma (ECCC) is a relatively rare type of endometrial cancer, typically less than 6% of all endometrial cancers [7]. It shares the features of the former two types, such as it is associated with atrophic endometrium, showing aggressive development and worse prognosis like type II, and the characteristics at the immunohistochemical and molecular levels like type I [8]. The rest type of EC is carcinosarcoma, which is a relatively rare and highly invasive type of endometrial cancer, only account for around 3% of endometrial cancers [9]. The definition as a framework instructs the study of teaching and scientific of endometrial cancer continuously. With the advancement of molecular techniques over the decades, multiple-dimensional characteristics have been used to supplement EC's classification definition. The Cancer Genome Atlas Research Network released a study that profiling EC at the genomic, transcriptomic, and proteomic levels. Based on the mutations and histological features, they classified EC into four types, such as POLE (a catalytic unit of DNA polymerase epsilon) ultramutated, accounting for 7.3% (17 of 232); MSI (microsatellite instability) hypermutated, accounting for 28.0% (65 of 232); low-copy-number, accounting for 38.8% (90 of 232); and high-copy-number, accounting for 25.9% (60 of 232) [10]. Early diagnosis is an ideal strategy to improve the low survival rate, evaluating the risk of EC becoming malignant or recurring, measurable indicators of abnormal biological states or dysfunctional conditions such as biomarkers or molecular features could be extremely useful [10–13]. The alteration detection at the early stage holds the potential to refine the

comprehensive classification system for improved accuracy in diagnosing malignancy, as well as to enhance the effectiveness of corresponding treatments. Herein, we offer an overview of the role multi-omics technologies play in biomarker discovery and discuss the prospective strategies for the early diagnosis of EC.

Multi-bio-layer omic data for prognosis of EC

Sample source

As Fig. 1 shows, advances in novel technique development have enabled the use of many types of samples in biomarker discovery. Tissue is the traditional and primary sample for detecting certain gene products specifically expressed in the tumor, facilitating the identification of specific biomarkers [14]. The testing based on tissue biopsy provides a definitive diagnosis of cancer by observing the morphology of the cells and identifying the malignant features. Although tissue biopsy is a standard for diagnosing and studying cancer, it does have several limitations. Because tumors are heterogeneous, tissue biopsy might not describe the entire cancer profiling. The poor repeatability of sample collection complicates this condition further. Additionally, puncturing to collect biopsies may facilitate cancer cell metastasis and elevate the risk to patients' lives [15]. In EC studies, the sampling procedure also poses challenges with approximately 6-33% of cases being unsuitable for pathological diagnosis due to the limited volume of target areas [16]. Liquid biopsy refers to the testing of blood or bodily secretions for the presence of cancer cells [17], which is a minimally invasive approach without surgical procedures. The low cost and convenience make biofluids ideal samples for tumor research. Compared with tissue biopsies, biofluids focus on the localized area of the tumor, liquid samples



Fig. 1 Omic data generation and computation from various sources at different expression levels

could provide information on the whole body condition by continuous monitoring without any invasive operation [18], helping doctors assess the effectiveness of the therapy, and make the necessary adjustments. The screening approaches for tumor-derived factors present in a variety of body fluids [19], such as blood [20], cervicovaginal fluid [20, 21], urine [22, 23], uterine lavage fluid [24], abdominopelvic washing, and ascites [25], have been used in the study and diagnosis of EC, measuring the characteristics of different levels. In the clinic, blood stands as the most commonly tested specimen for liquid biopsy. Except for that, several other bodily fluids have been shown high valuable as a non-invasive sample source. Cervicovaginal fluid is a complex blend of secretions from the uterus, cervix, and vagina, which has been used for biomarker discovery of gynecological diseases [21, 23]. There are studies indicating that the molecular signatures of endometrial cancer (EC) can be identified in cervical scrapes [26], endometrial brushes [27, 28], vaginal swabs [29, 30], and tampons [31, 32], suggesting that endometrial cell fragments could also be present in urine [22, 23]. Uterine lavage and abdominopelvic washing are procedures in a gynecological study introducing a saline solution into the uterine, abdominal, and pelvic cavities, and collecting the subsequent for detection and evaluation of cancer development. Ascites, a common symptom in late-stage cancer, is a valuable liquid biopsy medium, which often occurs in advanced cancer cases, containing cellular and molecular shed by the tumor. The biofluid contains shed cells [33], cancer cells [25, 34], tissue fragments, and other substances from the endometrial lining [24], such as proteins [35], and exosomes [36], providing a profiling of the tumor microenvironment, offering insights into the progression of abdominal malignancies such as endometrial [25], ovarian [37], pancreatic [38], and gastrointestinal [36] cancers. Exosomes are nanovesicles ranging from 30 to 150 nanometers in size, released by various cell types of both tissue and biofluids, rich in nucleic acids, lipids, and proteins [39–41]. They are extracellular structures enclosed by a lipid bilayer, formed by the outward budding of the plasma membrane and released into the extracellular environment through fusion with the plasma membrane [42], delivering information from tumor cells and various cells to the microenvironment [43–46]. Exosomal compositions reflect the molecular characteristics of the cells of origin and vary depending on distinct dysfunctional status. Consequently, tumor cells release unique exosomal contents, providing specific molecular mediators in extracellular communication [47, 48].

Data expression levels

Whether tissue, biological fluids, or exosomes [49] from them are important carriers of genetic material, full of various cancer biomarkers. In the advance of highthroughput techniques development, large amounts of omic data were generated and analyzed. The advent of next-generation sequencing (NGS) has significantly enhanced the precision and efficiency of genomic and transcriptomic studies, resulting in tons of discoveries related to genetic factors in cancer. The first major type of 'omics' data to be found and measured on the biology and development of cancer came from genomic sequences. Based on NGS, tumor genomic profiling becomes a common tool for classifying different types of cancer and identifying biomarkers that can evaluate and predict disease behavior. Somatic variants during cancer incidents can lead to the development and progression of tumors by affecting genes that control cell growth, division, and death. The distribution of mutation frequencies and types from cancers could be adopted for patient diagnostics and biomarker discovery [50]. Somatic copy number alterations (SCNAs), a type of somatic variant, occur in cancer cells to participate in cancer development by amplifications and deletions of the DNA segments. Through SCNAs, endometrial cancer was classified into different groups with significant prognostics [10], showing the potential of SCNAs as a tool for biomarker detection. Circulating tumor DNA (ctDNA) [51] is a kind of small nucleic acid released into the blood from the apoptosis or necrosis of tumor cells, whose expression level is affected by various immune responses, tumor status, and progression [52].

In contrast to the genome, which offers a relatively stable representation of the genetic characteristics, the transcriptome changes across various tissues, stages of development, and disease conditions [53]. Thus, understanding the variations in the transcriptome is key to knowing how genes respond and adjust under genetic or environmental factor influence. Messenger RNA (mRNA) serves as a template for protein synthesis. The mRNA abundance is often measured to determine the expression of genes associated with cancer. For instance, it has been adopted for prognostic markers of EC, promoting the development of effective targets in EC treatment [54]. In the human genome, the vast majority of genes are transcribed into non-coding RNAs (ncRNAs), which do not translate into proteins but have important biological functions, regulating the initiation and progression of various cancers [55]. Despite ncRNAs not being translated into proteins, there is persuasive evidence that they may act as promising biomarkers for cancer prognosis by regulating other genes [56]. In recent decades, studies of the function, regulatory mechanism, and therapeutic potential of the ncRNAs in EC are rapidly evolving, such as microRNA (miRNA) [57, 58], long ncRNA (lncRNA) [59], and circular RNA (circRNA) [60]. miR-NAs are endogenous, small ncRNAs (19-25 nucleotides)

that regulate gene expression post-transcriptionally [61], which attaches to matching sequences within specific mRNAs, usually resulting in gene silencing through translational repression or target degradation [62]. miR-NAs also connect with EC development, which can regulate the expression of genes involved in adhesion, migration, and invasion [63]. Based on this fact, several studies presented that miRNA could be the biomarker to predict the prognosis of EC [64, 65]. Zhou et al. conducted miRNA biomarker identification using the TCGA database and developed a linear regression model to assess the progression of EC. In their study, 26 miRNAs were found to have characteristics of two groups with different immune response conditions [66]. lncRNAs are longer than 200 nucleotides, also widely involved in a range of physiological activities [67]. Jiang et al. identified 1,931 expressed lncRNAs and established a subgroup classification depending on the discovery, providing the feature patterns to measure the risk of malignancy of EC [68]. Compared with lncRNAs, circRNAs are also more than 200 nt, but with a ring structure. The circular structure makes them more stable than linear RNAs, acting as competitive endogenous RNAs or as "microRNA sponges" to bind microRNAs, and preventing them from binding their target mRNAs and regulating gene expression in this way. Due to their ring-like structure stability and cell-type-specific expression patterns, circRNAs have been investigated as potential biomarkers for EC [69, 70].

Proteins, as the executors of gene functions, play a vital role in cellular activities and biological processes. The proteome refers to the complete set of proteins present in a cell, tissue, or organism at a specific time and under particular conditions. Proteomics is the large-scale study of proteomes, focusing on the comprehensive analysis of all proteins produced by an organism, system, or biological context. Mass spectrometry (MS) emerges as a precise technique for comprehensive proteomic studies, facilitating the measurement of features and alterations in proteomes. Proteomics complements genomics and transcriptomics by providing information about the actual functional molecules within the cell and can uncover changes in protein expression, modifications, localization, and interactions that occur in cancer states. There are two primary methods for acquiring proteomic data in the data generation process: data-dependent acquisition (DDA) and data-independent acquisition (DIA). The DDA method traditionally selects peptides based on the highest intensity signals for mapping to a pre-defined database [71]. While this approach enhances the credibility of the results, its stochastic selection process can lead to the loss of valuable information regarding lowabundant peptides [72]. In contrast, the DIA strategy is built upon the development of high-speed scanning techniques, which collect fragments of individual peptides across a series of mass windows, providing comprehensive peptide profiling with high reproducibility [71]. Both methods were employed for biomarker discovery in endometrial cancer (EC). For instance, in Wang's study [73], 20 patients with endometrial lesions and 7 healthy women were enrolled for specimen acquisition, leading to the identification of 7 biomarker candidates for EC using LC-MS/MS with a DDA strategy. Meanwhile, Jamaluddin's team collected 63 tumor biopsies from 20 patients to profile the expression features of EC using the DIA strategy [74].

The term "epigenetics" was originally coined by Conrad Waddington in 1942 to describe heritable changes in phenotype that occur without alterations in the DNA sequence. Benefiting from the development of various global omic techniques, the complexity and plasticity of epigenetics have become increasingly visible to us. Epigenetics encompasses a variety of reversible modifications to DNA, proteins, and RNA, usually happening on chromatin, that work together to regulate gene expression and cellular function [75]. As the primary mechanism of carcinogenesis at the epigenetic level, hypermethylation at site-specific DNA could silence tumor suppressor genes, while global hypomethylation can lead to chromosomal instability and oncogene activation [76]. In a study, researchers collected urine samples from 42 endometrial cancer (EC) patients and 46 healthy controls, identifying three DNA methylation markers that demonstrated high predictive performance for EC [77]. It has been shown that DNA methylation in urine is an effective source for EC biomarker discovery. Chromatin serves as an instructive DNA scaffold that responds to external factors for DNA regulation. As the principal component of chromatin, histones play crucial functions through post-translational modifications (PTMs). Acetylation, phosphorylation, glycosylation and other modifications as PTMs have gotten more attention for their diverse roles in the regulation of gene expression, protein structure, and molecular interactions. Some studies show that changes in the PTMs of proteins that are essential for prognosis and cell proliferation can cause tumorigenesis [39, 78]. Since protein is the basic functional element of most biological processes, its abundance and PTMs have been studied to provide deeper insights into disease, which is valuable in identifying diagnostic and prognostic markers for EC [79, 80]. In comparison with the modifications of DNA and proteins, relatively few reports indicate that mRNA undergoes extensive chemical modifications that could alter its function [81]. Post-transcriptional modification refers to the various biochemical modifications that RNA undergoes after transcription and before translation. These modifications are essential for the stability, localization, and function of RNA molecules. N6-methyladenosine (m6A) modification is widely

identified in mRNA, which regulates gene expression across several processes related to tumor proliferation, invasion, and epithelial-mesenchymal transition [82]. By studying these modifications, scientists can gain a deeper understanding of the mechanisms of gene regulation and their roles in biology and disease [83].

Publicly available sources

Many types of omic data have been gathered and used for the establishment of publicly available databases, accelerating the step of biomarker discovery processes. As well-known, The Cancer Genome Atlas (TCGA) is a vast collection of genetic data, combining over 11,158 samples across 32 cancer types [84]. This resource compiles detailed genetic information based on microarrays and next-generation sequencing methods, including RNA sequencing (RNAseq), microRNA sequencing (miRNAseq), DNA sequencing (DNAseq), SNP-based platforms, array-based DNA methylation sequencing, and reversephase protein array (RPPA) [85]. According to the TCGA database, Miao's team comprehensively investigated the variations of EC specimens at the genomic, transcriptomic, and epigenetic levels, resulting in 80 cancer-testis antigens (CTAs) genes more abundant in EC than in normal tissues. Among them, highly expressed TTK protein kinase (TTK), critical in driving EMT and chemoresistance, was significantly linked with lower survival in EC patients, showing the potential of being biomarker [86]. The Gene Expression Omnibus (GEO, http://www.ncbi. nlm.nih.gov/geo/) is a global public database committed to the storage of high-throughput microarray and nextgeneration sequencing functional genomic datasets. In Wu's study, TCGA and GEO were integrated to build and validate a prognostic model for EC malignancy prediction [87]. Proteomics offers a way to link gene alterations and cellular physiology. The Clinical Proteomic Tumor Analysis Consortium (CPTAC) provides the first multiomics database that integrates mass spectrometry (MS)based global proteomics data, expanding the support for TCGA samples [84, 88]. Combining the clinicopathological data from the TCGA, GEO, and CPTAC, Zhang found the gene products of S100A2 highly expressed in EC tissue at both the mRNA and protein levels, related to IL-17 signaling pathways [89]. The Human Protein Atlas (HPA) focuses on protein expression, distribution and localization, which aims to map all the human proteins in cells, tissues, and organs using the integration of various omics technologies, providing antibody-based imaging, mass spectrometry-based proteomics, clinical and histopathological details [90]. Through HPA database, Zhu measured the functions of TIMM8A in EC, resulting it could be a biomarker to predict the efficacy of anti-PD-L1 therapy [91].

Bioinformatics analysis of biomarker discovery

The emergence of numerous databases has made integration analysis possible, which is valuable in disease research, providing a more comprehensive view of the biological systems by considering the complex interactions between different types of molecular data. According to the definition of "proteomic biomarker" from the Mischak study [92], we could describe "biological marker" at different expression levels as specific gene products that are linked with specific conditions. Biomarkers could be used to assess the risk of disease malignancy, declining the death ratio of the patients. Understanding the clinical questions is the most important issue that needs to be considered in biomarker discovery. Under the clinical question's definition, the experimental workflow will be built properly, including sample type selection, casecontrol design, matched collection procedure, and decision of research direction underlying biological processes associated with clinical symptoms. Sufficient clinical data and matched information are beneficial to biomarker discovery. Detailed clinical information accompanying the samples is essential, such as sex, age, disease subtypes, disease stage, diagnosis, long-term follow-up data, etc. The factors could affect biomarker levels and have value in the predictive process, reflecting the correlations with disease progression or treatment response.

Here, a general workflow of biomarker discovery has been shown in Fig. 2. With advances in high-throughput techniques, the abundance of omics data generated makes it possible to combine machine learning algorithms, enhancing this field development. The application of machine learning in biomarker discovery is greatly impactful. Machine learning can process and analyze large-scale biomedical datasets, identifying complex patterns and associations in data and mining potential biomarker candidates [93, 94]. The performance of algorithm models is often improved with the increase in training data size. More data can help the model capture more complex patterns, get a better generalization ability, and improve the robustness to face the noise, performing more reliable predictions. Bioinformatics data have unique characteristics, such as high dimensionality, complexity, heterogeneity, noise, and multiscale nature, which make building the data analysis workflow a huge challenge.

To counterwork the difficulty of high dimensional data display, data dimensionality reduction was used to reduce the number of variables in a dataset through a mathematical transformation while preserving data features as much as possible. Principal Component Analysis (PCA) [95], Linear discriminant analysis (LDA) [96], t-Distributed Stochastic Neighbor Embedding (t-SNE) [97], and Uniform Manifold Approximation and Projection (UMAP) [98] are commonly used in data dimension



Fig. 2 Schematic representation of multi-omics approaches toward diagnosis in endometrial cancer

Table 1	Comparison of four dimensionality reduction
algorithr	15

algorithms			Algorithms	Advantages	Disadvantages
Algorithms	Advantages	Disadvantages	RF[100]	Handles large datasets with	Can be computationally
PCA[95]	Simple and easy to under- stand, high computational efficiency, retains global structure	Linear assump- tion, sensitive to outliers, limited interpretability	-	high dimensionality well, robust to overfitting, can handle both classification and regression tasks	intensive and slow to train, less interpretable, may require tuning of hyperparameters for
LDA[96]	Good classification perfor- mance, suitable for labeled data	Linear assumption, sensitive to high- dimensional data	Boruta[101]	Comprehensive feature High cor selection by comparing due to m	optimal performance High computational cost due to multiple model
t-SNE[97]	Excellent local structure pres- ervation, suitable for nonlinear data	High computation- al complexity, sensi- tive to parameters, difficult to interpret		actual features with shadow features, robust to noise and missing values, auto- mated process with minimal	training, sensitive to parameters of the under- lying random forest, may select redundant fea-
UMAP[98]	Preserves both global and local structure, high computa- tional efficiency, flexibility	Sensitive to param- eter selection, lim- ited interpretability	for various data types	datasets, less effective with small sample sizes	
			RFE[102]	Effective for selecting impor-	Computationally ex-

reduction (Table 1). They help people better understand data patterns and distribution by visualization, promoting the development of analytical strategies.

Feature selection is a fundamental phase in numerous machine learning workflows, aiming to streamline the task by discarding extraneous factors that might impact model performance. The gene products were treated as features to distinguish different types of samples in studies [99]. High-quality feature data directly influence the accuracy and trustworthiness of biomarker identification. It improves data utilization efficiency and bolsters the validity and clinical applicability of research findings. There are many algorithms (Table 2) available for feature selection, such as Random Forest (RF) [100] algorithm, Boruta [101] (based on RF algorithm), Recursive Feature

Table 2 Comparison of four feature selection algorithms

RF[100]	high dimensionality well, robust to overfitting, can handle both classification and regression tasks	can be computationally intensive and slow to train, less interpretable, may require tuning of hyperparameters for optimal performance
Boruta[101]	Comprehensive feature selection by comparing actual features with shadow features, robust to noise and missing values, auto- mated process with minimal human intervention, suitable for various data types	High computational cost due to multiple model training, sensitive to parameters of the under- lying random forest, may select redundant fea- tures in highly correlated datasets, less effective with small sample sizes
RFE[102]	Effective for selecting impor- tant features, can improve model performance by reducing overfitting, works well with various types of models	Computationally ex- pensive, especially with large datasets, may not perform well with highly correlated features
LASSO[103]	Performs both variable selection and regularization, can handle multicollinearity by shrinking coefficients, produces sparse models, making interpretation easier	Sensitive to the choice of regularization parameter, may exclude important variables if they are correlated with others, assumes a linear relation- ship between predictors and the response

Elimination (RFE) [102], and Least Absolute Shrinkage and Selection Operator (LASSO) [103].

Clustering or classification algorithms (Table 3), such as Hierarchical clustering (HC) [104], K-means [105],

 Table 3
 Comparison of four algorithms

Algorithms	Advantages	Disadvantages
HC[104]	Does not require the number of clusters to be specified in advance, produce the dendrogram, can capture complex cluster shapes	Computationally expensive, sensitive to noise and outliers, difficult to determine the optimal number of clusters from the dendrogram
K-means[105]	Simple and easy to imple- ment, efficient for large datasets with a known number of clusters	Requires the number of clusters to be speci- fied in advance, sensi- tive to initial centroid placement and outliers
KNN[106, 107]	Simple and intuitive, ef- fective for small datasets and when the decision boundary is complex, no explicit training phase	Primarily used for classification, compu- tationally expensive during prediction, sensitive to irrelevant features and the choice of distance metric
SVM[107, 108]	Effective in high- dimensional spaces, can model complex decision boundaries using kernel functions, robust to overfitting in high- dimensional space	Requires careful tuning of parameters (e.g., ker- nel choice, regulariza- tion), computationally intensive, especially with large datasets

k-Nearest Neighbor (KNN) [106, 107], and Support Vector Machine (SVM) [107, 108], play an important role in feature selection, as they can group genes with similar expression patterns, which helps reduce the dimensionality of the data, extract representative features, and simplify subsequent analysis. Grouped gene products with similar functions could reveal functional modules and potentially related biological pathways. Meanwhile, classifying samples into different subtypes based on their molecular characteristics is helpful in specific pattern identification under different biological conditions, providing new perspectives to study the disease mechanism and promotes the discovery process of biomarkers.

Differential expression analysis is essential for discerning notable differences in gene expression, protein levels, or other molecular measurements across varying conditions, such as between states of health and illness, and distinct treatments with temporal intervals. DESeq2 [109] is one of the algorithms for differential expression analysis in RNA-seq data, which estimates the variance and mean of gene expression through a model to identify differentially expressed genes, using shrinkage estimation for dispersions and fold changes to improve the stability and interpretability of estimates. edgeR [110] is a Bioconductor software package for examining the differential expression of replicated count data, which was constructed by an overdispersed Poisson [111] model and Empirical Bayes [112, 113] methods, improving the reliability of inference. limma [114] is another Bioconductor software package, suitable for microarray and RNA seq data, using empirical Bayesian methods and linear models to evaluate gene expression differences, emphasizing statistical power and multivariate analysis.

Survival information is invaluable in evaluating disease prognosis and treatment response. It can be adopted to measure the relationship between molecular characteristics and disease progression. To address the need, the Cox proportional-hazards model (COX) [115] and Kaplan-Meier survival analysis (KM) [116] have been widely utilized. Cox model is a statistical technique that allows researchers to determine whether there exists a statistically significant association between the feature patterns and the clinical events occurring over time. In contrast to the Cox model, the KM method emphasizes the assessment of survival time, facilitating the graphical representation of survival outcomes for patients with specific feature patterns, complementing the insights from the Cox model. In Li's research [99], univariate Cox regression analysis was employed to assess the impact of individual genes on the incidence of endometrial cancer. Meanwhile, the findings from the KM method were combined to provide Death-associated protein kinase-3 (DAPK3) and recombination signal binding protein for immunoglobulin kappa J region (RBPJ) as the biomarker candidates through Boruta and differential expression analysis. According to biomarker candidates of EC and multi-variate Cox regression, they built a linear model for the patient's prognosis. The methods mentioned above are used both integratively and partially in the discovery of tumor biomarkers.

Biomarker candidate validation and application

Experimental validation is crucial because it could confirm the true connections between biomarkers and clinical outcomes, guarantee the biomarkers' reliability, reproducibility, specificity, sensitivity, and biological relevance, and ensure the biomarker is not only statistically associated with the disease but is also practically useful for diagnosis, prognosis, or monitoring in a real-world clinical setting. The selection of experimental methodologies should depend on the types of gene products, ensuring appropriate validation consistent with the specific molecules under study.

Quantitative Polymerase Chain Reaction (qPCR), also known as real-time PCR, which is used to amplify and quantify a targeted DNA molecule. qPCR has several advantages, such as high sensitivity and specificity, realtime monitoring, high-throughput, and versatility [117]. qPCR can detect small amounts of target sequences and differentiate variants of similar nucleic acid sequences accurately. In Chen's study [118], qPCR has been utilized to identify TMEFF2 as the biomarker of EC. Immunohistochemistry (IHC) is a technique for detecting specific proteins through antigen antibody reactions in tissue sections. It typically uses fluorescent or enzyme labeled antibodies, combined with colorimetric reactions or fluorescent signals, to directly visualize the expression and localization of target proteins on tissue slices. Through IHC and qPCR, CCL25, CXCL10, CXCL12, and CXCL16 were validated for EC biomarker discovery [119]. Flow cytometry is employed to examine the physical and chemical characteristics of cells in a fluid suspension, including cell size, complexity, and the presence of specific surface markers or intracellular molecules. This technique enables the evaluation of thousands of cells per second, providing detailed information about individual cells. Using this method, ERRa was measured and validated in EC tumorigenesis in Su's research [120]. Western Blot (WB) is commonly used to detect specific proteins from tissue homogenate or extract. The high sensitivity of this technique facilitates the identification of specific target proteins for research objectives. Blendi.et al. [121] and Małgorzata.et al. [122] used western blot as a validation method to confirm the ABRACL and HLA-G respectively as potential biomarkers in EC. Enzyme-Linked Immunosorbent Assay (ELISA) is a widely used biochemical technique for detecting and quantitating specific proteins or molecules in a sample. It has high specificity and sensitivity due to antigenantibody reaction, providing exceptional efficiency, and allowing for simultaneous analyses without the need for intricate sample pre-treatment procedures. Nonetheless, it's important to note that ELISA assays typically focus on measuring one protein target per assay, constraining the simultaneous analysis of multiple biomarkers. In the context of endometrial cancer research, ELISA measures diverse biomarkers associated with the disease, such as tumor-associated antigens, cytokines, growth factors, or hormones [123]. For example, leptin was measured by ELISA and IHC in the study about EC progression [124]. Parallel Reaction Monitoring (PRM) is an advanced mass spectrometry-based technique used for targeted proteomics analysis This method enables the simultaneous detection and quantification of multiple proteins or peptides of interest in complex biological samples. PRM offers higher sensitivity, selectivity, and multiplexing capabilities compared to traditional shotgun proteomics approaches [125], making it an invaluable technique in the field of translational research and personalized medicine. After comparing the EC and non-EC tissues, Martinez-Garcia identified PERM, CADH1, SPIT1 and OSTP as the potential biomarkers by PRM [126]. Biomarkers are commonly used to define and identify different molecular subtypes, which are associated with specific biomarkers that can predict patient prognosis and response to treatment. Several biomarkers have been

Table 4	Biomarkers	of e	endometrial	cancer
---------	------------	------	-------------	--------

Biomarkers	Methods
CCL25, CXCL10, CXCL12, and CXCL16[119]	qPCR, IHC
ERRa[120]	Flow cytometry
ABRACL[121] and HLA-G[122]	Western Blot
leptin[124]	ELISA, IHC
PERM, CADH1, SPIT1 and OSTP[126]	PRM
TMEFF2[118, 129]	qPCR

used for prognostic, predictive of treatment response (Table 4) in EC studies through those methods [127, 128].

Conclusion and discussion

Despite the significant advances made by multi-omics technologies in uncovering potential biomarkers in recent years, the transition from discovery research to clinical practice continues to present widespread challenges. To design experiments scientifically, a clear grasp of medical issues and clinical symptoms is crucial, as the foundation for all steps of biomarker research. Multiomics techniques generate and analyze large amounts of data. Selecting proper sample types and matched data could help find the relationships between the molecular features and the corresponding disease symptoms, optimizing the data integrative analysis. The advances in high-throughput technologies and the implementation of novel approaches combining machine learning algorithms make multi-layer comprehensive analysis widely performed in biomarker discovery. Multi-omics aims to detect the complex pathophysiology of disease incidents, identifying risk factors and diagnostic cancer biomarkers. The complex cell signaling pathways leading to cancer development rely on the cumulative impact of changes across various expression levels. Therefore, a comprehensive understanding of the various cancer types and stages necessitates the adoption of a multi-omics approach, accelerating the progression of translation into clinical practice and establishment into precision medicine.

Acknowledgements

This work was supported by the projects of the Natural Science Basic Research Program of Shaanxi (No. 2022JN-475) National Natural Science Foundation of China (82022074, 82274196, 81930114 and 22307084), the Shenzhen Science and Technology Program (JCYJ20230807151059004).

Author contributions

Yuhao An, Bin Bai, and Li Jia wrote the main manuscript text; Xinrui Sha and Tuanjie Zhang prepared the figures; Linlin Lu and Rui Wang revised the maunscript. All authors reviewed the manuscript.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 19 July 2024 / Accepted: 14 January 2025 Published online: 26 April 2025

References

- Henley SJ, Ward EM, Scott S, Ma J, Anderson RN, Firth AU, Thomas CC, Islami F, Weir HK, Lewis DR, et al. Annual report to the nation on the status of cancer, part I: national cancer statistics. Cancer. 2020;126(10):2225–49. https://doi.org /10.1002/cncr.32802.
- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49. https://doi.org/10.3322/caac.21660.
- Amant F, Moerman P, Neven P, Timmerman D, Van Limbergen E, Vergote I. Endometrial cancer. Lancet (London England). 2005;366(9484):491–505. https ://doi.org/10.1016/s0140-6736(05)67063-8.
- Bokhman JV. Two pathogenetic types of endometrial carcinoma. Gynecol Oncol. 1983;15(1):10–7. https://doi.org/10.1016/0090-8258(83)90111-7.
- Suarez AA, Felix AS, Cohn DE. Bokhman Redux: endometrial cancer types in the 21st century. Gynecol Oncol. 2017;144(2):243–9. https://doi.org/10.1016/j. ygyno.2016.12.010
- Ferlay J, Shin H-R, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 2010;127(12):2893–917. https://doi.org/10.1002/ijc.25516
- Olawaiye AB, Boruta DM. Management of women with clear cell endometrial cancer: a Society of Gynecologic oncology (SGO) review. Gynecol Oncol. 2009;113(2):277–83. https://doi.org/10.1016/j.ygyno.2009.02.003
- Bae HS, Kim H, Young Kwon S, Kim K-R, Song JY, Kim I. Should endometrial clear cell carcinoma be classified as type II endometrial carcinoma? Int J Gynecol Pathol. 2015;34(1):74–84. https://doi.org/10.1097/PGP.00000000000 0111
- Lee EJ, Joo HJ, Ryu HS. Malignant mixed Müllerian tumors of the uterus: sonographic spectrum. Ultrasound Obstet Gynecol. 2012;39(3):348–53. https://doi. org/10.1002/uog.10085
- Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, Shen H, Robertson AG, Pashtan I, Shen R, Benz CC, et al. Integrated genomic characterization of endometrial carcinoma. Nature. 2013;497(7447):67–73. https://doi.org/10.103 8/nature12113.
- León-Castillo A, de Boer SM, Powell ME, Mileshkin LR, Mackay HJ, Leary A, Nijman HW, Singh N, Pollock PM, Bessette P, et al. Molecular classification of the PORTEC-3 trial for high-risk endometrial Cancer: impact on prognosis and benefit from adjuvant therapy. J Clin Oncology: Official J Am Soc Clin Oncol. 2020;38(29):3388–97. https://doi.org/10.1200/jco.20.00549.
- Piulats JM, Guerra E, Gil-Martín M, Roman-Canal B, Gatius S, Sanz-Pamplona R, Velasco A, Vidal A, Matias-Guiu X. Molecular approaches for classifying endometrial carcinoma. Gynecol Oncol. 2017;145(1):200–7. https://doi.org/10 .1016/j.ygyno.2016.12.015.
- Talhouk A, McConechy MK, Leung S, Li-Chang HH, Kwon JS, Melnyk N, Yang W, Senz J, Boyd N, Karnezis AN, et al. A clinically applicable molecular-based classification for endometrial cancers. Br J Cancer. 2015;113(2):299–310. https: //doi.org/10.1038/bjc.2015.190.
- Xiao Y, Bi M, Guo H, Li M. Multi-omics approaches for biomarker discovery in early ovarian cancer diagnosis. EBioMedicine. 2022;79:104001. https://doi.org /10.1016/j.ebiom.2022.104001.
- Wang Y, Du J, Lv S, Sui Y, Xue X, Sun C, Zou J, Ma Q, Fu G, Song Q, et al. Vaginal implantation metastasis of endometrial carcinoma: a case report. Oncol Lett. 2016;12(1):513–5. https://doi.org/10.3892/ol.2016.4656.
- Bakour SH, Khan KS, Gupta JK. Controlled analysis of factors associated with insufficient sample on outpatient endometrial biopsy. BJOG. 2000;107(10):1312–4. https://doi.org/10.1111/j.1471-0528.2000.tb11627.x
- Li W, Liu JB, Hou LK, Yu F, Zhang J, Wu W, Tang XM, Sun F, Lu HM, Deng J, et al. Liquid biopsy in lung cancer: significance in diagnostics, prediction, and treatment monitoring. Mol Cancer. 2022;21(1):25. https://doi.org/10.1186/s12 943-022-01505-z.
- Bao Y, Zhang D, Guo H, Ma W. Beyond blood: advancing the frontiers of liquid biopsy in oncology and personalized medicine. Cancer Sci. 2024;115(4):1060– 72. https://doi.org/10.1111/cas.16097.

- Zhao M, Yang Y, Guo Z, Shao C, Sun H, Zhang Y, Sun Y, Liu Y, Song Y, Zhang L. A comparative proteomics analysis of five body fluids: plasma, urine, cerebrospinal fluid, amniotic fluid, and saliva. Proteomics Clin Appl. 2018;12(6):1800008. https://doi.org/10.1002/prca.201800008.
- Njoku K, Pierce A, Chiasserini D, Geary B, Campbell AE, Kelsall J, Reed R, Geifman N, Whetton AD, Crosbie EJ. Detection of endometrial cancer in cervicovaginal fluid and blood plasma: leveraging proteomics and machine learning for biomarker discovery. EBioMedicine. 2024;102:105064. https://doi.org/10.1 016/j.ebiom.2024.105064.
- Zegels G, Van Raemdonck GA, Tjalma WA, Van Ostade XW. Use of cervicovaginal fluid for the identification of biomarkers for pathologies of the female genital tract. Proteome Sci. 2010;8:63. https://doi.org/10.1186/1477-5956-8-6
- Njoku K, Barr CE, Sutton CJJ, Crosbie EJ. Urine CA125 and HE4 for the triage of symptomatic women with suspected endometrial Cancer. Cancers. 2022;14(14). https://doi.org/10.3390/cancers14143306.
- O'Flynn H, Ryan NAJ, Narine N, Shelton D, Rana D, Crosbie EJ. Diagnostic accuracy of cytology for the detection of endometrial cancer in urine and vaginal samples. Nat Commun. 2021;12(1):952. https://doi.org/10.1038/s4146 7-021-21257-6
- Shen Y, Yang W, Liu J, Zhang Y. Minimally invasive approaches for the early detection of endometrial cancer. Mol Cancer. 2023;22(1):53. https://doi.org/1 0.1186/s12943-023-01757-3.
- Li X, Zhu D, Li N, Yang H, Zhao Z, Li M. Characterization of ascitesderived tumor cells from an endometrial cancer patient. Cancer Sci. 2017;108(12):2352–7. https://doi.org/10.1111/cas.13407
- Liew PL, Huang RL, Wu TI, Liao CC, Chen CW, Su PH, Wang HC, Weng YC, Lai HC. Combined genetic mutations and DNA-methylated genes as biomarkers for endometrial cancer detection from cervical scrapings. Clin Epigenetics. 2019;11(1):170. https://doi.org/10.1186/s13148-019-0765-3.
- Yuan J, Mao Z, Lu Q, Xu P, Wang C, Xu X, Zhou Z, Zhang T, Yu W, Dong S et al. Hypermethylated PCDHGB7 as a biomarker for early detection of Endometrial Cancer in Endometrial Brush samples and cervical scrapings. Front Mol Biosci. 2022;8:774215. https://doi.org/10.3389/fmolb.2021.774215
- DeJong SR, Bakkum-Gamez JN, Clayton AC, Henry MR, Keeney GL, Zhang J, Kroneman TN, Laughlin-Tommaso SK, Ahlberg LJ, VanOosten AL et al. Tao brush endometrial cytology is a sensitive diagnostic tool for cancer and hyperplasia among women presenting to clinic with abnormal uterine bleeding. Cancer Med. 2021;10(20):7040–7. https://doi.org/10.1002/cam4.423
- Doufekas K, Hadwin R, Kandimalla R, Jones A, Mould T, Crowe S, Olaitan A, Macdonald N, Fiegl H, Wik E et al. GALR1 methylation in vaginal swabs is highly accurate in identifying women with endometrial cancer. Int J Gynecol Cancer. 2013;23(6):1050–5. https://doi.org/10.1097/IGC.0b013e3182959103
- Weimer J, Hüttmann M, Nusilati A, Andreas S, Röseler J, Tribian N, Rogmans C, Stope MB, Dahl E, Mustea A et al. Fluorescence in situ hybridization test for detection of endometrial carcinoma cells by non-invasive vaginal swab. J Cell Mol Med. 2023;27(3):379–91. https://doi.org/10.1111/jcmm.17658
- Bakkum-Gamez JN, Sherman ME, Slettedahl SW, Mahoney DW, Lemens MA, Laughlin-Tommaso SK, Hopkins MR, VanOosten A, Shridhar V, Staub JK et al. Detection of endometrial cancer using tampon-based collection and methylated DNA markers. Gynecol Oncol. 2023;174:11–20. https://doi.org/10.1016/j. ygyno.2023.04.014
- Bakkum-Gamez JN, Wentzensen N, Maurer MJ, Hawthorne KM, Voss JS, Kroneman TN, Famuyide AO, Clayton AC, Halling KC, Kerr SE et al. Detection of endometrial cancer via molecular analysis of DNA collected with vaginal tampons. Gynecol Oncol. 2015;137(1):14–22. https://doi.org/10.1016/j.ygyno. 2015.01.552
- Rodriguez EF, Monaco SE, Khalbuss W, Austin RM, Pantanowitz L. Abdominopelvic washings: a comprehensive review. Cytojournal. 2013;10:7. https://doi. org/10.4103/1742-6413.111080
- Poonpanichakul T, Shiao M-S, Jiravejchakul N, Matangkasombut P, Sirachainan E, Charoensawan V, Jinawath N. Capturing tumour heterogeneity in pre- and post-chemotherapy colorectal cancer ascites-derived cells using single-cell RNA-sequencing. Biosci Rep. 2021;41(12):BSR20212093. https://doi.org/10.10 42/BSR20212093
- Tanaka N, Sekiya S, Takamizawa H, Kato N, Moriyama Y, Fujimura S. Characterization of a 54 kDa, alpha 1-antitrypsin-like protein isolated from ascitic fluid of an endometrial cancer patient. Jpn J Cancer Res. 1991;82(6):693–700. https ://doi.org/10.1111/j.1349-7006.1991.tb01905.x
- Wei M, Yang T, Chen X, Wu Y, Deng X, He W, Yang J, Wang Z. Malignant ascites-derived exosomes promote proliferation and induce

carcinoma-associated fibroblasts transition in peritoneal mesothelial Cells. Oncotarget. 2017;8(26):42262–71. https://doi.org/10.18632/oncotarget.15040

- Lo C-W, Chen M-W, Hsiao M, Wang S, Chen C-A, Hsiao S-M, Chang J-S, Lai T-C, Rose-John S, Kuo M-L et al. IL-6 trans-signaling information and progression of malignant ascites in ovarian cancer. Cancer Res. 2011;71(2):424–34. https:// doi.org/10.1158/0008-5472.CAN-10-1496
- Liu M-P, Guo X-Z, Xu J-H, Wang D, Li H-Y, Cui Z-M, Zhao J-J, Ren L-N. New tumor-associated antigen SC6 in pancreatic cancer. World J Gastroenterol. 2005;11(48):7671–5. https://doi.org/10.3748/wjg.v11.i48.7671
- O'Driscoll L. Expanding on exosomes and ectosomes in cancer. N Engl J Med. 2015;372(24):2359–62. https://doi.org/10.1056/nejmcibr1503100.
- Thakur BK, Zhang H, Becker A, Matei I, Huang Y, Costa-Silva B, Zheng Y, Hoshino A, Brazier H, Xiang J. Double-stranded DNA in exosomes: a novel biomarker in cancer detection. Cell Res. 2014;24(6):766–9. https://doi.org/10. 1038/cr.2014.44.
- Jiang C, Zhang N, Hu X, Wang H. Tumor-associated exosomes promote lung cancer metastasis through multiple mechanisms. Mol Cancer. 2021;20(1):1– 17. https://doi.org/10.1186/s12943-021-01411-w.
- Mathieu M, Martin-Jaular L, Lavieu G, Théry C. Specificities of secretion and uptake of exosomes and other extracellular vesicles for cell-to-cell communication. Nat Cell Biol. 2019;21(1):9–17. https://doi.org/10.1038/s41556-018-025 0-9.
- Johnstone RM, Adam M, Hammond J, Orr L, Turbide C. Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes). J Biol Chem. 1987;262(19):9412–20. PMID: 3597417.
- Maas SL, Breakefield XO, Weaver AM. Extracellular vesicles: unique intercellular delivery vehicles. Trends Cell Biol. 2017;27(3):172–88. https://doi.org/10.1 016/j.tcb.2016.11.003.
- Skog J, Würdinger T, Van Rijn S, Meijer DH, Gainche L, Curry WT, Carter BS, Krichevsky AM, Breakefield XO. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. Nat Cell Biol. 2008;10(12):1470–6. https://doi.org/10.1038/ncb1800.
- Yáňez-Mó M, Siljander PR-M, Andreu Z, Bedina Zavec A, Borràs FE, Buzas El, Buzas K, Casal E, Cappello F, Carvalho J. Biological properties of extracellular vesicles and their physiological functions. J Extracell Vesicles. 2015;4(1):27066. https://doi.org/10.3402/jev.v4.27066.
- Rosa-Fernandes L, Rocha VB, Carregari VC, Urbani A. Palmisano, G. A perspective on extracellular vesicles proteomics. Front Chem. 2017;5:102. https://doi. org/10.3389/fchem.2017.00102.
- Ratajczak J, Wysoczynski M, Hayek F, Janowska-Wieczorek A, Ratajczak M. Membrane-derived microvesicles: important and underappreciated mediators of cell-to-cell communication. Leukemia. 2006;20(9):1487–95. https://doi. org/10.1038/sj.leu.2404296.
- Sommella E, Capaci V, Aloisio M, Salviati E, Campiglia P, Molinario G, Licastro D, Di Lorenzo G, Romano F, Ricci G et al. A Label-Free Proteomic Approach for the Identification of Biomarkers in the Exosome of Endometrial Cancer Serum. Cancers. 2022;14(24):6262. https://doi.org/10.3390/cancers14246262.
- Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, et al. Mutational landscape and significance across 12 major cancer types. Nature. 2013;502(7471):333–9. https://doi.org/1 0.1038/nature12634.
- Blitzer GC, Zhao SG, Bradley KA, Hartenbach EM. The role of ctDNA in endometrial cancer: a tool for risk stratification and disease monitoring. Gynecol Oncol. 2023;178:170–1. https://doi.org/10.1016/j.ygyno.2023.08.008.
- Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM et al. Detection of circulating tumor DNA in earlyand late-stage human Malignancies. Sci Transl Med. 2014;6(224):224ra224. htt ps://doi.org/10.1126/scitranslmed.3007094
- Casamassimi A, Federico A, Rienzo M, Esposito S, Ciccodicola A. Transcriptome profiling in Human diseases: new advances and Perspectives. Int J Mol Sci. 2017;18(8):1652. https://doi.org/10.3390/ijms18081652.
- Agopiantz M, Forgez P, Casse JM, Lacomme S, Charra-Brunaud C, Clerc-Urmès I, Morel O, Bonnet C, Guéant JL, Vignaud JM et al. Expression of neurotensin receptor 1 in endometrial adenocarcinoma is correlated with histological grade and clinical outcome. Virchows Archiv: Int J Pathol 2017, 471 (4), 521–30. https://doi.org/10.1007/s00428-017-2215-y
- 55. Yan H, Bu P. Non-coding RNA in cancer. Essays Biochem. 2021;65(4):625–39. h ttps://doi.org/10.1042/EBC20200032.
- 56. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Nature. 2012;482(7385):339–46. https://doi.org/10.1038/nature10887.

- Sindhoo A, Sipy S, Khan A, Selvaraj G, Alshammari A, Casida ME, Wei DQ. ESOMIR: a curated database of biomarker genes and miRNAs associated with esophageal cancer. Database J Biol Databases Curat. 2023. https://doi.org/10. 1093/database/baad063
- Liu J, Eckert MA, Harada BT, Liu S-M, Lu Z, Yu K, Tienda SM, Chryplewicz A, Zhu AC, Yang Y et al. m6A mRNA methylation regulates AKT activity to promote the proliferation and tumorigenicity of endometrial cancer. Nat Cell Biol. 2018;20(9):1074–83. https://doi.org/10.1038/s41556-018-0174-4.
- Dong P, Xiong Y, Yue J, Hanley JB, Kobayashi S, Todo N, Watari Y. H. Exploring IncRNA-Mediated Regulatory Networks in Endometrial Cancer Cells and the Tumor Microenvironment: Advances and Challenges. Cancers. 2019;11(2):234. https://doi.org/10.3390/cancers11020234.
- Dou Y, Kawaler EA, Cui Zhou D, Gritsenko MA, Huang C, Blumenberg L, Karpova A, Petyuk VA, Savage SR, Satpathy S et al. Proteogenomic characterization of endometrial carcinoma. Cell. 2020;180(4):729–e748726. https://doi.org /10.1016/j.cell.2020.01.026.
- 61. Lu TX, Rothenberg ME. MicroRNA. J Allergy Clin Immunol. 2018;141(4):1202– 7. https://doi.org/10.1016/j.jaci.2017.08.034
- Vos PD, Leedman PJ, Filipovska A, Rackham O. Modulation of miRNA function by natural and synthetic RNA-binding proteins in cancer. Cell Mol Life Sci. 2019;76(19):3745–52. https://doi.org/10.1007/s00018-019-03163-9
- Chen C, Zhang Q, Kong B. miRNA-576-5p promotes endometrial cancer cell growth and metastasis by targeting ZBTB4. Clin Transl Oncol. 2022;25(3):706– 20. https://doi.org/10.1007/s12094-022-02976-8.
- Lu J, Liang J, Xu M, Wu Z, Cheng W, Wu J. Identification of an eleven-miRNA signature to predict the prognosis of endometrial cancer. Bioengineered. 2021;12(1):4201–16. https://doi.org/10.1080/21655979.2021.1952051.
- Wang Q, Xu K, Tong Y, Dai X, Xu T, He D, Ying J. Novel miRNA markers for the diagnosis and prognosis of endometrial cancer. J Cell Mol Med. 2020;24(8):4533–46. https://doi.org/10.1111/jcmm.15111.
- Zhou H, Chen L, Qin M, Lei Y, Li T, Li H, Cheng X. An miRNA signature associated with tumor mutation burden in endometrial cancer. Biosci Rep. 2020;40(11):BSR20203398. https://doi.org/10.1042/BSR20203398.
- 67. Huarte M. The emerging role of IncRNAs in cancer. Nat Med. 2015;21(11):1253–61. https://doi.org/10.1038/nm.3981.
- Jiang Y, Malouf GG, Zhang J, Zheng X, Chen Y, Thompson EJ, Weinstein JN, Yuan Y, Spano J-P, Broaddus R et al. Long non-coding RNA profiling links subgroup classification of endometrioid endometrial carcinomas with trithorax and polycomb complex aberrations. Oncotarget. 2015;6(37):39865–76. https: //doi.org/10.18632/oncotarget.5399.
- Guo J, Tong J, Zheng J. Circular RNAs: a Promising Biomarker for Endometrial Cancer. Cancer Manag Res. 2021;13:1651–65. https://doi.org/10.2147/CMAR.S 290975.
- Shi Y, He R, Yang Y, He Y, Shao K, Zhan L, Wei B. Circular RNAs: novel biomarkers for cervical, ovarian and endometrial cancer (review). Oncol Rep. 2020;44(5):1787–98. https://doi.org/10.3892/or.2020.7780.
- Mittal P, Klingler-Hoffmann M, Arentz G, Zhang C, Kaur G, Oehler MK, Hoffmann P. Proteomics of endometrial cancer diagnosis, treatment, and prognosis. Proteomics Clin Appl. 2016;10(3):217–29. https://doi.org/10.1002/ prca.201500055.
- Bauer M, Ahrné E, Baron AP, Glatter T, Fava LL, Santamaria A, Nigg EA, Schmidt A. Evaluation of data-dependent and -independent mass spectrometric workflows for sensitive quantification of proteins and phosphorylation sites. J Proteome Res. 2014;13(12):5973–88. https://doi.org/10.1021/pr500860c.
- Wang Y-s, Cao R, Jin H, Huang Y-p, Zhang X-y, Cong Q, He Y-f, Xu C. -j. Altered protein expression in serum from endometrial hyperplasia and carcinoma patients. J Hematol Oncol. 2011;4:15. https://doi.org/10.1186/1756-8722-4-1
- Jamaluddin MFB, Ko Y-A, Ghosh A, Syed SM, lus Y, O'Sullivan R, Netherton JK, Baker MA, Nahar P, Jaaback K et al. Proteomic and functional characterization of intra-tumor heterogeneity in human endometrial cancer. Cell Rep Med. 2022;3(9):100738. https://doi.org/10.1016/j.xcrm.2022.100738.
- Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet. 2011;12(8):529–41. https://doi.o rg/10.1038/nrg3000.
- Jones A, Teschendorff AE, Li Q, Hayward JD, Kannan A, Mould T, West J, Zikan M, Cibula D, Fiegl H et al. Role of DNA methylation and epigenetic silencing of HAND2 in Endometrial cancer development. PLoS Med. 2013;10(11):e1001551. https://doi.org/10.1371/journal.pmed.1001551.
- 77. van den Helder R, Wever BMM, van Trommel NE, van Splunter AP, Mom CH, Kasius JC, Bleeker MCG, Steenbergen RDM. Non-invasive detection of

endometrial cancer by DNA methylation analysis in urine. Clin Epigenetics. 2020;12(1):165. https://doi.org/10.1186/s13148-020-00958-7.

- Strepkos D, Markouli M, Klonou A, Papavassiliou AG, Piperi C. Histone methyltransferase SETDB1: a common denominator of Tumorigenesis with therapeutic potential. Cancer Res. 2021;81(3):525–34. https://doi.org/10.1158 /0008-5472.CAN-20-2906.
- Romano A, Rižner TL, Werner HMJ, Semczuk A, Lowy C, Schröder C, Griesbeck A, Adamski J, Fishman D, Tokarz J. Endometrial cancer diagnostic and prognostic algorithms based on proteomics, metabolomics, and clinical data: a systematic review. Front Oncol. 2023;13:1120178. https://doi.org/10.3389/fon c.2023.1120178.
- Fu C, Zhao H, Wang Y, Cai H, Xiao Y, Zeng Y, Chen H. Tumor-associated antigens: Tn antigen, sTn antigen, and T antigen. HLA. 2016;88(6):275–86. https:// doi.org/10.1111/tan.12900.
- Meyer KD, Jaffrey SR. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. Nat Rev Mol Cell Biol. 2014;15(5):313–26. https:/ /doi.org/10.1038/nrm3785.
- He L, Li H, Wu A, Peng Y, Shu G, Yin G. Functions of N6-methyladenosine and its role in cancer. Mol Cancer. 2019;18(1):176. https://doi.org/10.1186/s1294 3-019-1109-9.
- Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell. 2012;150(1):12–27. https://doi.org/10.1016/j.cell.2012.06.013.
- Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. Nucleic Acids Res. 2018;46(D1):D956– 63. https://doi.org/10.1093/nar/gkx1090.
- Tomczak K, Czerwińska P, Wiznerowicz M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemporary Oncology/ Współczesna Onkologia. 2015;2015(1):68–77. https://doi.org/10.5114/wo.201 4.47136.
- Miao Y, Konno Y, Wang B, Zhu L, Zhai T, Ihira K, Kobayashi N, Watari H, Jin X, Yue J et al. Integrated multi-omics analyses and functional validation reveal TTK as a novel EMT activator for endometrial cancer. J Translational Med. 2023;21(1):151. https://doi.org/10.1186/s12967-023-03998-8.
- Wu J, Zhou X, Ren J, Zhang Z, Ju H, Diao X, Jiang S, Zhang J. Glycosyltransferase-related prognostic and diagnostic biomarkers of uterine corpus endometrial carcinoma. Comput Biol Med. 2023;163:107164. https://doi.org/ 10.1016/j.compbiomed.2023.107164.
- Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, Chambers MC, Zimmerman LJ, Shaddox KF, Kim S, et al. Proteogenomic characterization of human colon and rectal cancer. Nature. 2014;513(7518):382–7. https://doi.org/10.1038/nat ure13438.
- Zhang Q, Xia T, Qi C, Du J, Ye C. High expression of S100A2 predicts poor prognosis in patients with endometrial carcinoma. BMC Cancer. 2022;22(1):77. https://doi.org/10.1186/s12885-022-09180-5.
- Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419. https://doi.o rg/10.1126/science.1260419.
- Zhu X, Yuan Z, Cheng S, Wang H, Liao Y, Zhou D, Wu Z. TIMM8A is associated with dysfunction of immune cell in BRCA and UCEC for predicting anti-PD-L1 therapy efficacy. World J Surg Oncol. 2022;20(1):336. https://doi.org/10.1186/ s12957-022-02736-6.
- Mischak H, Allmaier G, Apweiler R, Attwood T, Baumann M, Benigni A, Bennett SE, Bischoff R, Bongcam-Rudloff E, Capasso G, et al. Recommendations for biomarker identification and qualification in clinical proteomics. Sci Transl Med. 2010;2(46):46ps42. https://doi.org/10.1126/scitranslmed.3001249.
- Ling D, Liu A, Sun J, Wang Y, Wang L, Song X, Zhao X. Integration of IDPC Clustering Analysis and Interpretable Machine Learning for Survival Risk Prediction of patients with ESCC. Interdisciplinary Sciences: Comput Life Sci. 2023;15(3):480–98. https://doi.org/10.1007/s12539-023-00569-9.
- Xu R, Wang J, Zhu Q, Zou C, Wei Z, Wang H, Ding Z, Meng M, Wei H, Xia S, et al. Integrated models of blood protein and metabolite enhance the diagnostic accuracy for Non-small Cell Lung Cancer. Biomark Res. 2023;11(1):71. https ://doi.org/10.1186/s40364-023-00497-2.
- Maćkiewicz A, Ratajczak W. Principal components analysis (PCA). Comput Geosci. 1993;19(3):303–42. https://doi.org/10.1016/0098-3004(93)90090-R.
- 96. Fisher RA. The use of multiple measurements in taxonomic problems. Annals Eugenics. 1936;7(2):179–88. https://doi.org/10.1111/j.1469-1809.1936.tb0213 7.x.
- 97. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res 2008;9(11).

- McInnes L, Healy J, Melville J, Umap. Uniform manifold approximation and projection for dimension reduction. 2018. https://doi.org/10.48550/arXiv.180 2.03426.
- Li T, Ruan Z, Song C, Yin F, Zhang T, Shi L, Zuo M, Lu L, An Y, Wang R, et al. Integrative Analysis of Multi-omic Data for the characteristics of Endometrial Cancer. ACS Omega. 2024. https://doi.org/10.1021/acsomega.4c00375.
- 100. Ho TK. Random decision forests. In Proceedings of 3rd international conference on document analysis and recognition, 1995, Vol. 1, pp 278-282-278–282.
- 101. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Softw. 2010;36:1–13. https://doi.org/10.18637/jss.v036.i11.
- Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer classification using support Vector machines. Mach Learn. 2002;46(1):389–422. https://doi. org/10.1023/A:1012487302797.
- Tibshirani R. Regression shrinkage and Selection Via the Lasso. J Roy Stat Soc: Ser B (Methodol). 1996;58(1):267–88. https://doi.org/10.1111/j.2517-6161.199 6.tb02080.x.
- Johnson SC. Hierarchical clustering schemes. Psychometrika. 1967;32(3):241– 54. https://doi.org/10.1007/bf02289588.
- MacQueen J. Some methods for classification and analysis of multivariate observations. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1967; Oakland, CA, USA: Vol. 1, pp 281–297.
- 106. Guo G, Wang H, Bell D, Bi Y, Greer K. KNN Model-Based Approach in Classification. In On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, Berlin, Heidelberg, 2003; Meersman, R., Tari, Z., Schmidt, D. C., Eds.; Springer Berlin Heidelberg: pp 986–996.
- Dong C, Zhao L, Liu X, Dang L, Zhang X. Single-cell analysis reveals landscape of endometrial cancer response to estrogen and identification of early diagnostic markers. PLoS ONE. 2024;19(3):e0301128. https://doi.org/10.1371/jour nal.pone.0301128.
- Cristianini N, Ricci E. Support Vector machines. In: Kao M-Y, editor. Encyclopedia of algorithms. Ed.; 2008. pp. 928–32. Springer US.
- Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:1–21. https:/ /doi.org/10.1186/s13059-014-0550-8.
- Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40. https://doi.org/10.1093/bioinformatics/btp616.
- 111. McCullagh P. Generalized linear models. Routledge; 2019. https://doi.org/10.1 201/9780203753736.
- 112. Robbins H. The empirical Bayes approach to statistical decision problems. Ann Math Stat. 1964;35(1):1–20. https://doi.org/10.1214/aoms/1177703729.
- 113. Robbins HE. An empirical Bayes approach to statistics. Breakthroughs in statistics: foundations and basic theory. Springer; 1992. pp. 388–94.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth G. K. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47–47. https://doi.org/10.1093/nar/gkv007.
- 115. Cox DR. Regression models and life-tables. J Roy Stat Soc: Ser B (Methodol). 1972;34(2):187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x.
- 116. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53(282):457–81. https://doi.org/10.2307/2281868.
- 117. Kralik P, Ricchi MA. Basic Guide to Real Time PCR in Microbial Diagnostics: definitions, parameters, and everything. Front Microbiol. 2017;8:108. https://d oi.org/10.3389/fmicb.2017.00108.
- Chen Y-C, Tsao C-M, Kuo C-C, Yu M-H, Lin Y-W, Yang C-Y, Li H-J, Yan M-D, Wang T-J, Chou Y-C, et al. Quantitative DNA methylation analysis of selected genes in endometrial carcinogenesis. Taiwan J Obstet Gynecol. 2015;54(5):572–9. ht tps://doi.org/10.1016/j.tjog.2015.08.010.
- 119. Li J, Ramzan F, Zhong G. Investigating novel biomarkers in uterine corpus endometrial carcinoma: in silico analysis and clinical specimens validation via RT-qPCR and immunohistochemistry. Am J Cancer Res. 2023;13(9):4376–400.
- 120. Su P, Mao X, Ma J, Huang L, Yu L, Tang S, Zhuang M, Lu Z, Osafo KS, Ren Y et al. ERRa promotes glycolytic metabolism and targets the NLRP3/caspase-1/ GSDMD pathway to regulate pyroptosis in endometrial cancer. J Exp Clin Cancer Res. 2023;42(1):274. https://doi.org/10.1186/s13046-023-02834-7.
- 121. Ura B, Monasta L, Arrigoni G, Franchin C, Radillo O, Peterlunger I, Ricci G, Scrimin F. A proteomic approach for the identification of biomarkers in endometrial cancer uterine aspirate. Oncotarget. 2017;8(65):109536–45. https://do i.org/10.18632/oncotarget.22725.
- 122. Walentowicz-Sadlecka M, Dziobek K, Grabiec M, Sadlecki P, Walentowicz P, Mak P, Szymankiewicz M, Kwinta P, Dutsch-Wicherek M. The analysis of human leukocyte antigen-G level in patients with endometrial cancer by

Western blot technique. Am J Reprod Immunol. 2019;81(1):e13070. https://d oi.org/10.1111/aji.13070.

- 123. Romano A, Rižner TL, Werner HMJ, Semczuk A, Lowy C, Schröder C, Griesbeck A, Adamski J, Fishman D, Tokarz J. Endometrial cancer diagnostic and prognostic algorithms based on proteomics, metabolomics, and clinical data: a systematic review. Front Oncol. 2023;13:1120178. https://doi.org/10.3389/fon c.2023.1120178.
- 124. Shen L, Zhang C, Cui K, Liang X, Zhu G, Hong L. Leptin secreted by adipocytes promotes EMT transition and endometrial cancer progression via the JAK2/ STAT3 signaling pathway. Adipocyte. 2024;13(1):2293273. https://doi.org/10.1 080/21623945.2023.2293273.
- 125. Martinez-Garcia E, Lesur A, Devis L, Campos A, Cabrera S, van Oostrum J, Matias-Guiu X, Gil-Moreno A, Reventos J, Colas E, et al. Development of a sequential workflow based on LC-PRM for the verification of endometrial cancer protein biomarkers in uterine aspirate samples. Oncotarget. 2016;7(33):53102–15. https://doi.org/10.18632/oncotarget.10632.
- 126. Martinez-Garcia E, Lesur A, Devis L, Campos A, Cabrera S, van Oostrum J, Matias-Guiu X, Gil-Moreno A, Reventos J, Colas E et al. Development of

a sequential workflow based on LC-PRM for the verification of endometrial cancer protein biomarkers in uterine aspirate samples. Oncotarget. 2016;7(33):53102–15. https://doi.org/10.18632/oncotarget.10632.

- 127. Walsh CS, Hacker KE, Secord AA, DeLair DF, McCourt C, Urban R. Molecular testing for endometrial cancer: an SGO clinical practice statement. Gynecol Oncol. 2023;168:48–55. https://doi.org/10.1016/j.ygyno.2022.10.024.
- Boroń D, Zmarzły N, Wierzbik-Strońska M, Rosińczuk J, Mieszczański P, Grabarek BO. Recent multiomics approaches in Endometrial Cancer. Int J Mol Sci. 2022;23(3):1237. https://doi.org/10.3390/ijms23031237.
- 129. Gao L, Nie X, Zheng M, Li X, Guo Q, Liu J, Liu Q, Hao Y, Lin B. TMEFF2 is a novel prognosis signature and target for endometrial carcinoma. Life Sci. 2020;243:116910. https://doi.org/10.1016/j.lfs.2019.116910.

Publisher's note

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