The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

Tracing the Tissue of Origin in Cancers of Unknown Primary Using Cell-Free DNA (cfDNA)

Tuna Önel (author)[1]

[1]Cambridge Centre for International Research, Istanbul, Turkey

Abstract

Cancers of unknown primary (CUP) are metastatic cancers in which the origin cannot be identified despite thorough diagnostic evaluations. This lack of knowledge delays the institution of specific treatment protocols and is linked with unfavorable clinical outcomes. One of the solutions to this problem lies in the application of liquid biopsy, which is a minimally invasive process that analyzes circulating tumor DNA fragments present in the blood (Hum & Lee, 2025). The principal purpose of liquid biopsy is to identify the origin of cancer by recognizing specific DNA methylation signatures. The systematic review provides a summary of 56 research articles that have used DNA methylation profiling of cell-free DNA (cfDNA) for identifying the tissue of origin of cancers classified as tumors of unknown primary. Across these studies, researchers employed targeted methylation panels, whole-genome bisulfite sequencing (WGBS), reduced representation bisulfite sequencing (RRBS), and Illumina EPIC arrays. The researchers applied logistic regression, support vector machines (SVM), random forests, and deep learning, trained

on cancers with known primaries and tested on CUP samples (Jia et al., 2021). Most of the classifiers performed well, with areas under the curves (AUC) measures exceeding 0.90 (Li et al., 2025). The CUPiD study reported an AUC of 0.99 and predicted the probable tumor origin in almost 80% of patients with CUP (Stanley et al., 2024). Collectively, the results indicate that cfDNA methylation profiling is a potentially useful and less invasive strategy to enhance the precision of diagnosis and treatment planning for patients with CUP (Liu, 2022).

Keywords: cancer of unknown primary (CUP), cell-free DNA (cfDNA), multi-omics, liquid biopsy, methylation profiling.

Introduction

CUP is a subset of metastatic cancers for which the location of the primary cancer is not identified even after thorough diagnostic investigations. CUP accounts for variable incidence, constituting approximately 3–5% of all cancers, and is a significant challenge in the field of oncology. Being unable to identify the primary site does not allow oncologists to institute site-specific, targeted therapeutic sessions, treatment strategies often associated with better prognostic outcomes (Fece de la Cruz & Corcoran, 2018; Sharma et al., 2022; Spurgeon et al., 2025; Zarkavelis et al., 2024).

So, CUP patients are typically treated with empirical and non-specific combination chemotherapy that is low in response rate and does not result in survival advantage.

The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

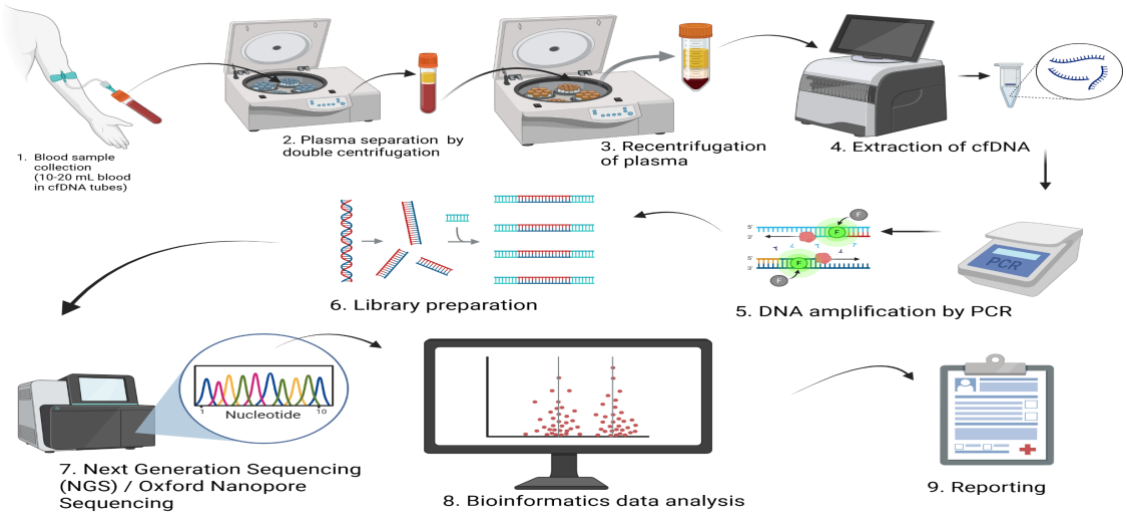
The primary organ has traditionally been determined using an amalgamation of radiologic, histologic, and immunohistochemical tests. These tests, nonetheless, may render ambiguous outcomes, especially when the neoplasm is inadequately differentiated or the metastatic characteristics do not possess unique anatomical characteristics (Helzer et al., 2023; Liu et al., 2020; Raman, 2021). It is for this reason that researchers have sought other avenues, such as molecular profiling.

Liquid biopsies were created as a noninvasive approach, with real-time access to the tumor biology by utilizing easily obtainable blood or fluid samples. These techniques provide the isolation of tumor-derived material -predominantly cfDNA- from a peripheral blood sample or from other body fluids, such as pleural or ascitic fluid (Duckett et al., 2025; Hum & Lee, 2025; Luo et al., 2021). Of the molecular features of cfDNA, DNA methylation is emphasized given its tissue type specificity and stability. Methylation patterns can be extremely variably expressed in various tissues, creating tissue-specific signatures that enable identification of a tumor's origin (Conway et al., 2024; Ma, Wu, et al., 2024; Zhou et al., 2022). Through this capacity, cfDNA methylation profiling has emerged as an essential element in addressing the diagnostic dilemma of CUP. In this review, we take a close look at various studies to examine how well cfDNA methylation profiling, often paired with advanced computational methods, performs in predicting the origin of CUP. This review summarizes the experimental strategies, assesses the performance of various models, discusses the technological improvements and limitations, and addresses future prospects for clinical application.

The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

Figure 1.

Overview of Liquid Biopsy



Created in BioRender. Önel, T. (2026) https://BioRender.com/l6hg6d8

Methodology

Studies were deemed eligible if they met the following criteria: original research articles, published from 2012 to 2025, describing cfDNA or ctDNA methylation profiling from liquid biopsies, and tackling tissue-of-origin prediction in CUP or cancers in which the approach was amenable to be used in CUP scenarios. An extensive search of databases, including PubMed, Web of Science, Nature, ScienceDirect, and open-access preprint servers like BioRxiv and medRxiv was conducted.

The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

Methylation Profiling Techniques

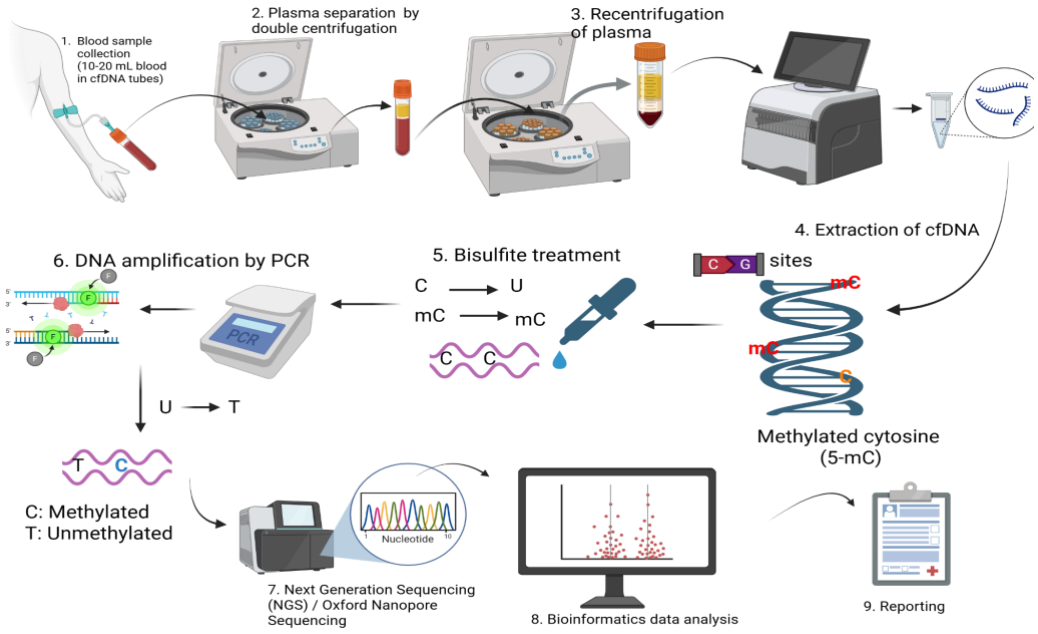
The 56 studies gathered in this paper utilize diverse approaches to establish methylation signatures in cell-free DNA (cfDNA). Whole-genome bisulfite sequencing (WGBS) was most frequently employed in research with the priority of whole-genome methylome representation; however, its excessive DNA input needs and expense restricted its more widespread application (Stark et al., 2024). As opposed to this, reduced representation bisulfite sequencing (RRBS) was less expensive and provided an equilibrium between analytical resolution and input needs (Mathur et al., 2025).

Targeted methylation panels have also been extensively used for their clinical scalability and sensitivity. The panels were specifically designed to target differentially methylated regions (DMRs) that are highly tissue-specific (Lee et al., 2025). One such example is the panel used in the CUPiD study, which was built from over 6,000 tumor samples across 29 cancer types (Duckett et al., 2025). While Illumina's EPIC arrays were initially designed for solid tissue analysis, some studies modified them to be able to include liquid biopsy samples, thus providing an equivalent methodology for methylation profiling (Nguyen et al., 2023; Passemiers et al., 2024).

The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

Figure 2.

Overview of cfDNA Methylation Profiling



Created in BioRender. Önel, T. (2026) https://BioRender.com/2y1mk3i

Machine Learning Classifiers

In the majority of the reports, supervised learning models were employed to contrast cfDNA methylation profiles against a reference database of various known cancer types. Classifiers that were utilized include logistic regression, support vector machines (SVM), random forests, and more recently, neural networks and graph-based models. Deep learning approaches -primarily convolutional neural networks and attention models- were applied to high-

throughput data sets, where pattern discrimination power might be greater than that of linear approaches (Li et al., 2025; Ma, Wu, et al., 2024; Sun et al., 2024; Zhang et al., 2024).

All classifiers required a reference dataset that was usually derived from The Cancer Genome Atlas (TCGA), the International Cancer Genome Consortium (ICGC), or from biobanks specific to individual institutions. In several projects, publicly available datasets were merged with in-house sequencing datasets to improve tissue representation (Taryma-Leśniak et al., 2024).

Table 1.

Machine Learning Classifiers Used for Tissue-of-Origin Prediction

Study	Classifier Type	Input Features	Reference Dataset	Validation Method
CUPiD	SVM	Methylation	TCGA	External
MetDetect	Random Forest	Methylation	ICGC + In-house	10-fold CV
Liu et al. (2022)	Logistic Regression	Methylation + Fragmentomics	TCGA	Retrospective
Wang et al.	Graph Neural Net	Methylation	In-house	External

(2024)				
Zhang et al. (2023)	CNN + Fragmentomics	Methylation + Fragmentomics	TCGA + GEO	Split-sample

Reference Datasets and Validation Protocols

Validation approaches were also a key point differentiating the methodological soundness of each study. Some studies employed internal validation methods, k-fold cross-validation being the prominent one, particularly in cases of small-sized datasets. Other studies used large-scale, independent validation cohorts to test for real-world applicability and feasibility. In CUP-specific testing, models were tested on retrospective clinical samples where the final tumor origin was confirmed by delayed diagnosis or post-mortem autopsy (Lau et al., 2023; Mattox et al., 2023; Nguyen et al., 2024).

Some studies establish clinical concordance as a performance measure -model predictions and clinical history compared with histopathological reports, immunohistochemistry (IHC), and imaging reports. Some also followed patients longitudinally, observing response to treatment when therapies were aligned with predicted tumor origin. In such studies, methylation predictions were utilized to guide site-specific therapy, and patient outcomes were accepted as indirect validation.

Enhancing classifier interpretability was also a focus. Deep learning experiments routinely involved SHAP (Shapley Additive Explanations) or feature importance scores to determine the important methylation sites driving classification (Mahmood et al., 2023). This allowed for greater insight into model decisions and improved transparency, particularly where findings were not in line with clinical expectations.

In addition, reproducibility of methylation markers across sequencing platforms has been dealt with in a subset of studies. Such studies highlighted the importance of platform-independent markers and reproducible bioinformatic pipelines to reduce technical noise. Standard pre-processing, including correction of bisulfite conversion efficiency, filtering of CpG sites, and correction of batch effects, was shown to significantly improve comparability across studies. Together, these validation studies highlight the point that while cfDNA methylation classifiers are promising, their clinical integration will depend on external validation, replicability, and clarity in the reporting of how predictions are derived (Conway et al., 2024; Lu et al., 2021; Ma, Guo, et al., 2024; Óscar Rapado-González et al., 2023).

Results

Overall Classifier Performance

In the 56 studies included, classification performance was generally high. More than 75% of the models had an area under the receiver operating characteristic curve (AUC) of >0.90. The CUPiD study was particularly notable, with an AUC of 0.99 and an estimated prediction success

rate of approximately 80% in determining the tissue of origin in cancer of unknown primary (CUP) patients (Duckett et al., 2025; Lau et al., 2023).

Research employing deep learning models, i.e., convolutional neural networks, tended to demonstrate higher sensitivity and specificity, especially for challenging tissue classes such as pancreatic, biliary, or upper gastrointestinal cancer.

Table 2.

Performances of Classifier Across Studies

Study	Tumor Types	Accuracy (%)	AUC	Sample Type	Notes
CUPiD	29	93	0.99	Plasma	Best performance
MetDetect	18	89	0.94	Pleural	Low ctDNA
Liu et al. (2022)	11	87	0.91	Plasma	Mid-range
Wang et al. (2024)	20	91	0.95	Ascitic	Rare cancers

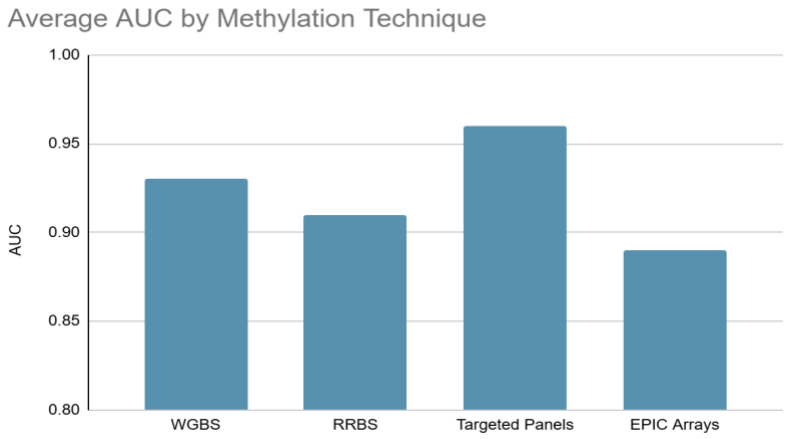
Zhang et al. (2023)	22	90	0.96	Plasma	Integrated fragmentomics

Comparative Evaluation of Techniques

Studies using targeted methylation panels generally had higher clinical utility in the form of lower sequencing costs, ease of sample processing, and shorter turnaround times. These, however, often had limited resolution for rare or histologically similar tumor types. WGBS and RRBS, conversely, provided broader methylome coverage and more comprehensive epigenetic data but required higher cfDNA input and sequencing depth. EPIC arrays were often used in studies that highlighted the reproducibility of platforms and proved effective in cross-validation settings (Bixby et al., 2024; Passemiers et al., 2024; Stanley et al., 2024).

Figure 3.

Performances of Methylation Profiling Methods Across Studies



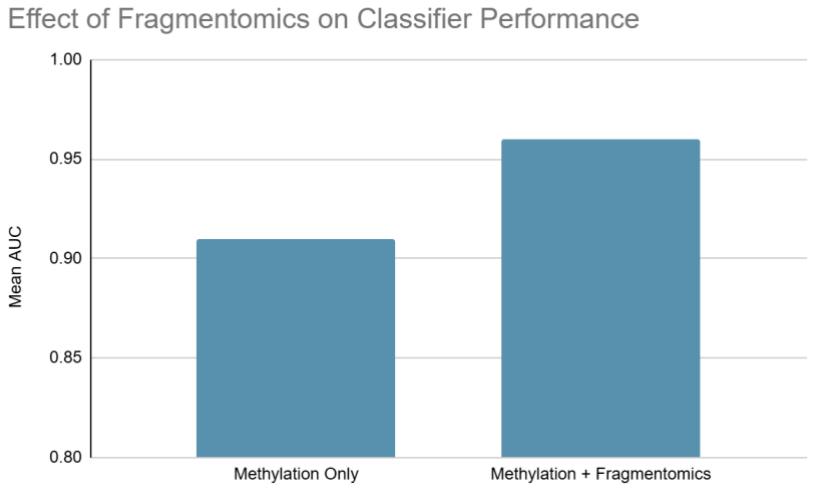
The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

Role of Fragmentomics

Approximately 20% of the studies integrated fragmentomics with methylation profiling. Fragment length patterns, transcriptional start site (TSS) coverage, and nucleosome positioning were used to augment classifier features (Zhang, Liu, et al., 2023). Fragmentomics achieved 5–10% accuracy improvements, particularly for samples with low ctDNA fraction (Jia et al., 2021). Fragmentomics was also useful in distinguishing between blood contamination and actual tumor-derived signals in pleural or ascitic fluid samples.

Figure 4.

Impact of Fragmentomics Integration on Classifier Performance



Sample Type and Yield

Several studies explored alternative fluid sources beyond blood plasma. Pleural effusion and ascitic fluid were found to contain higher concentrations of tumor-derived cfDNA in certain

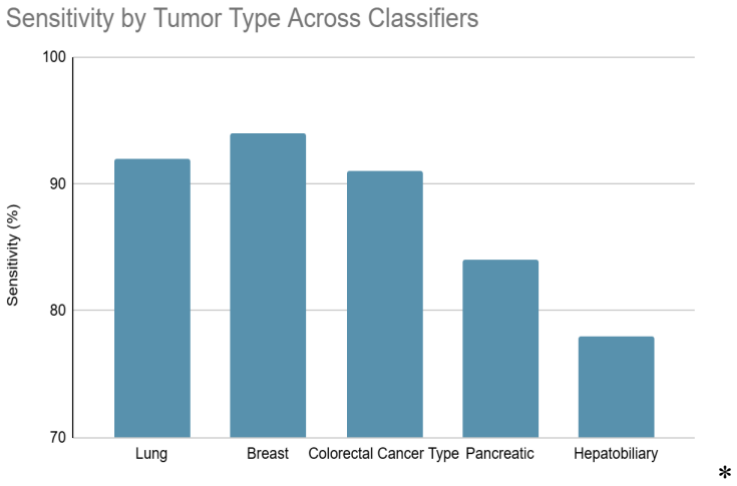
cancers such as ovarian, gastrointestinal, and lung adenocarcinoma. In these studies, methylation profiling yielded better prediction accuracy than plasma-based approaches, likely due to increased tumor DNA abundance and reduced background noise (Mahmood et al., 2023; Richardson et al., 2025).

Tumor Type-Specific Observations

Classifiers were found to be very proficient in the determination of the origin of some cancer types. For instance, methylation panels trained on cohorts covering lung, breast, and colon cancers had sensitivities of over 90% consistently (Zhou et al., 2022). Conversely, those cancers with less discrete methylation profiles -such as neuroendocrine tumors or hepatobiliary cancers- continued to be challenging to classify (Möhrmann et al., 2022). A few attempts circumvented this issue by employing probabilistic modeling or confidence scoring for ambiguous cases (Spurgeon et al., 2025).

Figure 5.

Clinical Impact of Methylation-Based Tissue-of-Origin Prediction in CUP



*

Impact on Treatment Decisions and Outcomes

In studies with clinical follow-up, models of prediction based on methylation were at times utilized to guide treatment decisions. For instance, patients with altered treatment recommendations based on the output of classifiers experienced limited benefit for both progression-free survival (PFS) and overall survival (OS). Such studies were, nevertheless, small in number and mostly had a retrospective nature, pointing to the need for prospective trials (Helzer et al., 2023; Stark et al., 2024).

Table 3.

Outcomes of CUP Treatments Across Studies

Study	Sample Type	Modified Treatment	Reported Outcome	Notes
CUPiD	Plasma	Yes	PFS 1	Classifier-guided
Zhang et al. (2023)	Plasma	Yes	PFS, OS 1	Survival benefit
Wang et al. (2024)	Ascitic	No	No change	Exploratory
MetDetect	Pleural	Yes	Improved response	Retrospective cohort
Liu et al. (2022)	Plasma	No	Unclear	Pilot

Technical and Analytical Challenges

Despite strong outcomes, several technical challenges remain. They included cfDNA degradation, incomplete bisulfite conversion, and clonal hematopoiesis noise. Model overfitting was also observed in some deep learning uses with small training sample sizes (Nguyen et al., 2023; Noravesh & Palizban, 2022; Shuihui Wang Lorkowski et al., 2023). Research tackling

those challenges routinely suggested the imposition of quality control measures and clearly articulated model evaluation paradigms to avert performance metric corruption.

Together, these findings validate the hypothesis that methylation profiling of cfDNA – especially when augmented by fragmentomics and trained on appropriate fluid types– is an accurate approach to predicting tissue of origin in CUP patients. However, it is still critical to maintain consistency in assay design, classifier training, and validation procedures to move forward towards clinical deployment, as shown in Table 4.

Table 4.

Challenges of Methylation Profiling of cfDNA

Challenge	Number of Studies Reporting	Impact	Proposed Solution
Bisulfite conversion loss	18	Signal loss	Optimized protocols
Clonal hematopoiesis	12	False positives	CHIP filtering
Low ctDNA concentration	20	Low sensitivity	Fragmentomics
Batch effects	15	Reduced comparability	Standard pipelines

Overfitting (DL models)	10	Inflated metrics	External validation

Discussion

The overall evidence from the 56 studies examined strongly substantiates the application of cfDNA methylation profiling as a less invasive, valid approach to CUP identification of the tissue of origin. The consistently good classification performance -frequently in excess of 90% sensitivity or AUC- demonstrates that this approach has evolved beyond a hypothetical theoretical concept and now manifests actual diagnostic promise for everyday clinical practice. A number of technical trends and considerable methodological patterns can be identified (Luo et al., 2021; Nguyen et al., 2024; Richardson et al., 2025; Stark et al., 2024; Zhang, He, et al., 2023).

The primary advantage of targeted methylation panels is that they are cost-effective and flexible for everyday clinical application. A few of the panels mentioned in the literature covered were developed by identifying unique tissue-specific differentially methylated regions (DMRs). The targeted methods have worked well in clinical contexts where diagnostic outputs that are quick and cheap are necessary, but may not perform well in discriminating epigenetically similar tissues (Lee et al., 2025; Ma, Guo, et al., 2024; Passemiers et al., 2024; Stanley et al., 2024; Wang et al., 2021).

By comparison, genome-wide approaches like WGBS and RRBS provide more extensive epigenetic resolution, enabling the detection of rare or unclear tissue signals. Nevertheless, they demand deeper sequencing, greater cfDNA input, and advanced bioinformatic analysis, which hinders their scalability for everyday use. Research that coupled such extensive strategies with powerful machine learning pipelines –especially those involving ensemble learning or deep neural networks– produced the best performance, particularly for malignancies with low signal-to-noise ratios in cfDNA. Another theme that repeatedly emerged was the judicious use of fragmentomics. Integration of fragment length characteristics and nucleosome occupancy data with methylation improved model discrimination in samples with low ctDNA content. This integrated approach was especially helpful for tumor types like pancreatic or ovarian, where methylation alone was sometimes not quite discriminatory enough (Jia et al., 2021; Ma, Wu, et al., 2024; Mattox et al., 2023; Óscar Rapado-González et al., 2023).

The heterogeneity of the sample source -ranging from plasma to pleural effusion and ascitic fluid- also impacted diagnostic accuracy. cfDNA concentrations in fluids proximal to the tumor site were higher and increased methylation-based classifier predictive performance. This finding holds promise for the use of site-specific biofluids in future CUP diagnostic assays (Li & Sun, 2024; Rendek et al., 2024; Spurgeon et al., 2025).

Importantly, several studies demonstrated that the integration of methylation-based predictions into clinical practice could potentially guide treatment and improve patient outcomes. Even modest gains in progression-free survival or response to targeted treatment provide a basis for designing future clinical trials to establish the utility of cfDNA methylation profiling in real-

time CUP patient decision-making (Hum & Lee, 2025; Mahmood et al., 2023; Zhang et al., 2024).

Another direction in recent research is the application of multi-omics methods - approaches that combine multiple types of biological data, such as DNA methylation, gene expression, proteomics, and fragmentomics, into a single analysis. A different window into tumor biology is offered by each omic layer: methylation is representative of stable, tissue-specific patterns; RNA expression yields real-time cellular activity; and fragmentomics yields structural features of cfDNA. Taken together, the layers provide a more informative and dynamic portrait of cancer activity. In CUP, this is perhaps particularly valuable. A number of studies demonstrated that multi-omics models could outperform methylation-only classifiers, especially for difficult-to-discriminate tumors such as gastrointestinal or neuroendocrine cancers. Synergy between the omics layers enabled better sensitivity in low-ctDNA samples and added biological context for equivocal results. Some models leveraged transcriptomic signals to validate tissue predictions, while others utilized fragment length patterns to sharpen classifier boundaries. Computationally more intensive and less standardized if they may be, such approaches represent a step toward more comprehensive diagnostics. They don't merely detect cancer; they start to explain it. As such, multi-omics is an exciting avenue to improve tissue-of-origin predictions, diminish uncertainty, and bring precision oncology another step closer to CUP patients. Future studies should address their clinical scalability and regulatory pathways more concretely (Lu et al., 2021; Michaël Noë et al., 2024; Noravesh & Palizban, 2022; Sun et al., 2024).

Yet, there are key clinical adoption hurdles. These include heterogeneity in bisulfite conversion efficiency, cfDNA fragmentation bias, and lack of universal benchmarking datasets. Deep learning models, while powerful, lack interpretability that could potentially make regulatory approval challenging. Additionally, the prevalence of clonal hematopoiesis and technical batch effects continues to be a problem unless controlled severely (Liu et al., 2020; Taryma-Leśniak et al., 2024; Zhou et al., 2022).

In spite of these difficulties, the literature reviewed here presents a strong argument for ongoing development and validation of cfDNA methylation profiling as a key component in the diagnosis of CUP. Future research should be aimed at standardizing methodologies, conducting multicenter validations, and combining other omics-based strategies -like proteomics or transcriptomics- to develop even stronger classifiers (Chemi et al., 2022; Helzer et al., 2023; Nguyen et al., 2023; Shuihui Wang Lorkowski et al., 2023).

The convergence of liquid biopsy technology, artificial intelligence-driven analytics, and high-resolution epigenomics has the potential to transform how we diagnose and manage CUP in the years to come (Alharbi et al., 2024; Fece de la Cruz & Corcoran, 2018; Seval Turna & Semra Demokan, 2024; Sirajee et al., 2024).

Limitations

While the prospects of cfDNA methylation profiling for identifying tissue of origin in CUP are encouraging, there are some significant limitations that cannot be ignored. The majority of the studies we reviewed were conducted on retrospective cohorts, many of which were from

single institutions. This limits generalizability and raises potential sampling bias with regards to sample processing and handling and cohort demographics (De Wilde et al., 2024).

Another significant challenge is the absence of standardized protocols among studies. Variation in DNA extraction protocols, bisulfite conversion efficiency, and sequencing platforms introduces variability that can affect classifier performance. While some studies performed cross-platform validations, the lack of consistency in pre-analytical and analytical workflows is a key limiting factor to large-scale clinical translation (Krawczyk et al., 2022).

Sample size variation is also a problem. Most of the deep learning models were trained on small or imbalanced datasets, which predisposes to overfitting and compromises robustness when applied to external populations. In addition, CUP is itself a heterogeneous condition, and some subtypes with unique methylation patterns can be underrepresented in training data, reducing model performance in such cases (Barefoot et al., 2021).

Fragmentomics, while with enhanced performance, is still not widely supported by clinical-grade software and requires additional processing steps that can delay result turnaround times. Its extension to routine diagnostics will also require validation within regulatory frameworks, which are still evolving for multi-modal liquid biopsy assays (Hwang et al., 2024).

Moreover, there have been few studies that have dealt with the influence of biological confounding variables, including inflammation, comorbidities, and clonal hematopoiesis, which can all influence cfDNA methylation patterns and thus their interpretation. Failure to account for these introduces the risk of misclassification (van et al., 2023). Lastly, although numerous

investigations have demonstrated enhanced survival with classifier-directed therapy, none have conducted prospective clinical validation. In the absence of randomized studies, the clinical usefulness of cfDNA methylation-based tissue-of-origin classification to enhance treatment selection and outcome is promising but as yet untested (Laprovitera et al., 2021).

Conclusion

The collective evidence from 56 studies indicates that cfDNA methylation profiling is a promising non-invasive method for the determination of the tissue of origin in cancers of unknown primary. As sequencing technology and machine learning advance, classifiers trained on tissue-specific methylation patterns have demonstrated high performance in a broad variety of cancer types and biological fluids (Bie et al., 2023). The integration of fragmentomics, fluid-specific sampling, and improved model transparency have all collectively contributed towards more robust diagnostic outcomes. Methodological heterogeneity, small sample sizes, and insufficient prospective validation are issues that plague greater clinical uptake at present, though. Future research needs to prioritize multicenter prospective studies, pipeline standardization, and harmonization with other omic layers for increased diagnostic strength (Esfahani et al., 2022). With bioinformatics software and regulatory frameworks becoming more transparent, cfDNA methylation profiling will be at the forefront of CUP diagnostics in the near future, guiding therapy in a personalized but exact manner (BIXBY et al., 2024).

The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

References

Alharbi, F., Vakanski, A., Elbashir, M. K., & Mohammed, M. (2024). LASSO–MOGAT: a multi-omics graph attention framework for cancer classification. *Academia Biology*, *2*(3). https://doi.org/10.20935/acadbiol7325

Barefoot, M. E., Loyfer, N., Kiliti, A. J., McDeed, A. P., Kaplan, T., & Wellstein, A. (2021). Detection of Cell Types Contributing to Cancer From Circulating, Cell-Free Methylated DNA. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/fgene.2021.671057

Bie, F., Wang, Z., Li, Y., Guo, W., Hong, Y., Han, T., Lv, F., Yang, S., Li, S., Li, X., Nie, P., Xu, S., Zang, R., Zhang, M., Song, P., Feng, F., Duan, J., Bai, G., Li, Y., & Huai, Q. (2023). Multimodal analysis of cell-free DNA whole-methylome sequencing for cancer detection and localization. *Nature Communications*, *14*(1), 6042. https://doi.org/10.1038/s41467-023-41774-w

BIXBY, B. A., BUSSEY, K., KUMAR, S., OSHIRO, M., & NELSON, M. (2024). A DISEASE CLASSIFIER FOR CANCER DETECTION IN PLEURAL FLUID USING LIQUID BIOPSY AND DNA METHYLATION ANALYSIS. *Chest*, *166*(4), A3754–A3755. https://doi.org/10.1016/j.chest.2024.06.2250

Bixby, B., Vrba, L., Jyoti Lenka, Oshiro, M. M., Watts, G. S., Hughes, T., Erickson, H., Chopra, M., Knepler, J. L., Knox, K. S., Jarnagin, L., Raed Alalawi, Kala, M., Bernert, R., Routh, J., Roe, D. J., Garland, L. L., Futscher, B. W., & Nelson, M. A. (2024). Cell-free DNA methylation analysis as a marker of malignancy in pleural fluid. *Scientific Reports*, *14*(1).

https://doi.org/10.1038/s41598-024-53132-x

Chemi, F., Pearce, S. P., Clipson, A., Hill, S. M., Conway, A.-M., Richardson, S. A., Katarzyna Kamieniecka, Caeser, R., White, D. J., Mohan, S., Foy, V., Simpson, K. L., Galvin, M., Frese, K. K., Priest, L., Egger, J., Kerr, A., Massion, P. P., Poirier, J. T., & Brady, G. (2022). cfDNA methylome profiling for detection and subtyping of small cell lung cancers. *Nature Cancer*, *3*(10), 1260–1270. https://doi.org/10.1038/s43018-022-00415-9

Conway, A.-M., Pearce, S. P., Clipson, A., Hill, S. M., Chemi, F., Slane-Tan, D., Ferdous, S., Hossain, M., Katarzyna Kamieniecka, White, D. J., Mitchell, C., Kerr, A., Krebs, M. G., Brady, G., Dive, C., Cook, N., & Rothwell, D. G. (2024). A cfDNA methylation-based tissue-of-origin classifier for cancers of unknown primary. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-47195-7

De Wilde, J., Van Paemel, R., De Koker, A., Roelandt, S., Van de Velde, S., Callewaert, N., Van Dorpe, J., Creytens, D., De Wilde, B., & De Preter, K. (2024). A Fast, Affordable, and Minimally Invasive Diagnostic Test for Cancer of Unknown Primary Using DNA Methylation Profiling. *Laboratory Investigation*, *104*(8), 102091. https://doi.org/10.1016/j.labinv.2024.102091

Duckett, D., Vormittag-Nocito, E. R., Jamshidi, P., Madina Sukhanova, Parker, S., Brat, D. J., Jennings, L. J., & Santana-Santos, L. (2025). Accurate identification of primary site in tumors of unknown origin (TUO) using DNA methylation. *Npj Precision Oncology*, *9*(1). https://doi.org/10.1038/s41698-025-00805-z

Esfahani, M. S., Hamilton, E. G., Mehrmohamadi, M., Nabet, B. Y., Alig, S. K., King, D. A., Steen, C. B., Macaulay, C. W., Schultz, A., Nesselbush, M. C., Soo, J., Schroers-Martin, J. G., Chen, B., Binkley, M. S., Stehr, H., Chabon, J. J., Sworder, B. J., Hui, A. B.-Y. ., Frank, M. J., & Moding, E. J. (2022). Inferring gene expression from cell-free DNA fragmentation profiles. *Nature Biotechnology*, *40*(4), 585–597. https://doi.org/10.1038/s41587-022-01222-4

Fece de la Cruz, F., & Corcoran, R. B. (2018). Methylation in cell-free DNA for early cancer detection. *Annals of Oncology*, *29*(6), 1351–1353. https://doi.org/10.1093/annonc/mdy134

Helzer, K. T., Sharifi, M. N., Sperger, J. M., Shi, Y., Annala, M., Bootsma, M. L., Reese, S. R., Taylor, A., Kaufmann, K. R., Krause, H. K., Schehr, J. L., N. Sethakorn, Kosoff, D., Kyriakopoulos, C., Burkard, M. E., Rydzewski, N. R., Yu, M., Harari, P. M., Bassetti, M., & Blitzer, G. (2023). Fragmentomic analysis of circulating tumor DNA-targeted cancer panels. *Annals of Oncology*, *34*(9), 813–825. https://doi.org/10.1016/j.annonc.2023.06.001

Hum, M., & Lee, A. S. G. (2025). DNA methylation in breast cancer: early detection and biomarker discovery through current and emerging approaches. *Journal of Translational Medicine*, *23*(1). https://doi.org/10.1186/s12967-025-06495-2

Hwang, J., Lee, Y., Yoo, S.-K., & Kim, J.-I. (2024). Image-based deep learning model using DNA methylation data predicts the origin of cancer of unknown primary. *Neoplasia*, *55*, 101021–101021. https://doi.org/10.1016/j.neo.2024.101021

Jia, S., Zhang, Y., Mao, Y., Gao, J., Chen, Y., Jiang, Y., Luo, H., Lv, K., & Su, J. (2021). *A new parsimonious method for classifying Cancer Tissue-of-Origin Based on DNA Methylation 450K data*. ArXiv.org. https://arxiv.org/abs/2101.00570

Krawczyk, P., Jassem, J., Wojas-Krawczyk, K., Krzakowski, M., Dziadziuszko, R., & Olszewski, W. (2022). New Genetic Technologies in Diagnosis and Treatment of Cancer of Unknown Primary. *Cancers*, *14*(14), 3429. https://doi.org/10.3390/cancers14143429

Laprovitera, N., Salamon, I., Gelsomino, F., Porcellini, E., Mattia Riefolo, Garonzi, M., Tononi, P., Valente, S., Sabbioni, S., Fontana, F., Nicolò Manaresi, D'Errico, A., Pantaleo, M. A., Ardizzoni, A., & Ferracin, M. (2021). Genetic Characterization of Cancer of Unknown Primary Using Liquid Biopsy Approaches. *Frontiers in Cell and Developmental Biology*, *9*. https://doi.org/10.3389/fcell.2021.666156

Lau, B. T., Almeda, A., Schauer, M. C., McNamara, M., Bai, X., Meng, Q., Partha, M., Grimes, S. M., Lee, H., Heestand, G. M., & Ji, H. P. (2023). Single-molecule methylation profiles of cell-free DNA in cancer with nanopore sequencing. *Genome Medicine*, *15*(1). https://doi.org/10.1186/s13073-023-01178-3

Lee, K.-J., Sampath, D., & Mavrommatis, K. (2025). *A DNA Methylation Classification Model Predicts Organ and Disease Site*. ArXiv.org. https://arxiv.org/abs/2506.00146

Li, L., & Sun, Y. (2024). Circulating tumor DNA methylation detection as biomarker and its application in tumor liquid biopsy: advances and challenges. *MedComm*, *5*(11). https://doi.org/10.1002/mco2.766

Li, N., Song, K., Chen, H., & Dai, M. (2025). Advance and challenge of DNA methylation as cancer biomarkers for risk stratification, screening and early detection. *Journal of the National Cancer Center*, *5*(2). https://doi.org/10.1016/j.jncc.2024.12.007

Liu, M. C., Oxnard, G. R., Klein, E. A., Swanton, C., Seiden, M. V., Cummings, S. R., Absalan, F., Alexander, G., Allen, B., Amini, H., Aravanis, A. M., Bagaria, S., Bazargan, L., Beausang, J. F., Berman, J., Betts, C., Blocker, A., Bredno, J., Calef, R., & Cann, G. (2020). Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Annals of Oncology*, *31*(6). https://doi.org/10.1016/j.annonc.2020.02.011

Liu, P. (2022). Pan-Cancer DNA Methylation Analysis and Tumor Origin Identification of Carcinoma of Unknown Primary Site Based on Multi-Omics. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/fgene.2021.798748

Lu, M. Y., Chen, T. Y., Williamson, D. F. K., Zhao, M., Shady, M., Lipkova, J., & Mahmood, F. (2021). AI-based pathology predicts origins for cancers of unknown primary. *Nature*. https://doi.org/10.1038/s41586-021-03512-4

Luo, H., Wei, W., Ye, Z., Zheng, J., & Xu, R. (2021). Liquid Biopsy of Methylation Biomarkers in Cell-Free DNA. *Trends in Molecular Medicine*, *27*(5), 482–500. https://doi.org/10.1016/j.molmed.2020.12.011

Ma, L., Guo, H., Zhao, Y., Liu, Z., Wang, C., Bu, J., Sun, T., & Wei, J. (2024). Liquid biopsy in cancer current: status, challenges and future prospects. *Signal Transduction and Targeted*

Therapy, *9*(1). https://doi.org/10.1038/s41392-024-02021-w

Ma, W., Wu, H., Chen, Y., Xu, H., Jiang, J., Du, B., Wan, M., Ma, X., Chen, X., Lin, L., Su, X., Bao, X., Shen, Y., Xu, N., Ruan, J., Jiang, H., & Ding, Y. (2024). New techniques to identify the tissue of origin for cancer of unknown primary in the era of precision medicine: progress and challenges. *Briefings in Bioinformatics*, *25*(2). https://doi.org/10.1093/bib/bbae028

Mahmood, K., Jampani, P., Clarke, J. M., Wolf, S., Wang, X., Wahidi, M. M., Giovacchini, C. X., Dorry, M., Shofer, S. L., Shier, J., Jones, G., Antonia, S. J., & Nixon, A. B. (2023). High Yield of Pleural Cell-Free DNA for Diagnosis of Oncogenic Mutations in Lung Adenocarcinoma. *CHEST*, *164*(1), 252–261. https://doi.org/10.1016/j.chest.2023.01.019

Mathur, K., Jain, S., Bajiya, N., Kumar, N., & Raghava, Gajendra P. S. (2025). *In silico tool for identification of colorectal cancer from cell-free DNA biomarkers*. ArXiv.org. https://arxiv.org/abs/2505.11041

Mattox, A. K., Douville, C., Wang, Y., Popoli, M., Ptak, J., Silliman, N., Dobbyn, L., Schaefer, J., Lu, S., Pearlman, A. H., Cohen, J. D., Tie, J., Gibbs, P., Kamel Lahouel, Chetan Bettegowda, Hruban, R. H., Tomasetti, C., Jiang, P., K.C. Allen Chan, & Ming, Y. (2023). The Origin of Highly Elevated Cell-Free DNA in Healthy Individuals and Patients with Pancreatic, Colorectal, Lung, or Ovarian Cancer. *Cancer Discovery*, *13*(10), 2166–2179. https://doi.org/10.1158/2159-8290.cd-21-1252

Michaël Noë, Dimitrios Mathios, Annapragada, A. V., Koul, S., Foda, Z. H., Medina, J. E.,

Cristiano, S., Cherry, C., Bruhm, D. C., Noushin Niknafs, Vilmos Adleff, Ferreira, L., Hari Easwaran, Baylin, S., Phallen, J., Scharpf, R. B., & Velculescu, V. E. (2024). DNA methylation and gene expression as determinants of genome-wide cell-free DNA fragmentation. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-50850-8

Möhrmann, L., Werner, M., Oleś, M., Mock, A., Uhrig, S., Jahn, A., Kreutzfeldt, S., Fröhlich, M., Hutter, B., Paramasivam, N., Richter, D., Beck, K., Winter, U., Pfütze, K., Heilig, C. E., Teleanu, V., Lipka, D. B., Zapatka, M., Hanf, D., & List, C. (2022). Comprehensive genomic and epigenomic analysis in cancer of unknown primary guides molecularly-informed therapies despite heterogeneity. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-31866-4

Nguyen, T. H., Tan, N., Tran, T. H., Anh, L., Doan, P. L., Hue, T., Thien, V., Huong, T., Nguyen, H.-N., Giang, H., Tran, L. S., & Phan, M. D. (2024). Tissue of origin detection for cancer tumor using low-depth cfDNA samples through combination of tumor-specific methylation atlas and genome-wide methylation density in graph convolutional neural networks. *Journal of Translational Medicine*, *22*(1). https://doi.org/10.1186/s12967-024-05416-z

Nguyen, V. T. C., Nguyen, T. H., Doan, N. N. T., Pham, T. M. Q., Nguyen, G. T. H., Nguyen, T. D., Tran, T. T. T., Vo, D. L., Phan, T. H., Jasmine, T. X., Nguyen, V. C., Nguyen, H. T., Nguyen, T. V., Nguyen, T. H. H., Huynh, L. A. K., Tran, T. H., Dang, Q. T., Doan, T. N., Tran, A. M., & Nguyen, V. H. (2023). Multimodal analysis of methylomics and

fragmentomics in plasma cell-free DNA for multi-cancer early detection and localization. *ELife*, *12*, RP89083. https://doi.org/10.7554/eLife.89083

Noravesh, F., & Palizban, F. (2022). *Identifying OCRs in cfDNA WGS Data by Correlation Clustering*. ArXiv.org. https://arxiv.org/abs/2202.09618

Óscar Rapado-González, Ana María Rodríguez-Ces, López-López, R., & María Mercedes Suárez-Cunqueiro. (2023). Liquid biopsies based on cell-free DNA as a potential biomarker in head and neck cancer. *Japanese Dental Science Review*, *59*(3), 289–302. https://doi.org/10.1016/j.jdsr.2023.08.004

Passemiers, A., Tuveri, S., Dhanya Sudhakaran, Tatjana Jatsenko, Laga, T., Punie, K., Hatse, S., Tejpar, S., Coosemans, A., Els Van Nieuwenhuysen, Timmerman, D., Floris, G., Anne-Sophie Van Rompuy, Sagaert, X., Testa, A., Ficherova, D., Raimondi, D., Amant, F., Liesbeth Lenaerts, & Moreau, Y. (2024). MetDecode: Methylation-based deconvolution of cell-free DNA for non-invasive multi-cancer typing. *Bioinformatics*, *40*(9). https://doi.org/10.1093/bioinformatics/btae522

Raman, L. (2021, January 1). *Copy number and methylation profiling of cell-free DNA for tumor and prenatal diagnostics - a multi-cohort research effort*. Ghent University's Academic Bibliography. https://www.researchgate.net/publication/352706167_Copy_number_and_methylation_profiling_of_cell-free_DNA_for_tumor_and_prenatal_diagnostics_-_a_multi-cohort_research_effort

Rendek, T., Pos, O., Terezia Duranova, Saade, R., Jaroslav Budis, Repiska, V., & Szemes, T. (2024). Current Challenges of Methylation-Based Liquid Biopsies in Cancer Diagnostics. *Cancers*, *16*(11), 2001–2001. https://doi.org/10.3390/cancers16112001

Richardson, S. A., Gulati, G., Harker, P. J. B., Redfern, A., Pearce, S. P., Karimpour, M., Hill, S. M., Makeev, V. J., Brennan, B., Lee, A. T., Clipson, A., Dive, C., Rothwell, D. G., McCabe, M. G., & Mouliere, F. (2025). Cell-free DNA methylome and fragmentome analysis for disease relapse monitoring in patients with Ewing Sarcoma. *MedRxiv*. https://doi.org/10.1101/2025.04.25.25325254

Seval Turna, & Semra Demokan. (2024). *Methylation Biomarkers and Liquid Biopsy Studies in Cancer*. 84–129. https://doi.org/10.26650/b/lsb28lsb48lsb56.2024.019.004

Sharma, M., Verma, R. K., Kumar, S., & Kumar, V. (2022). Computational challenges in detection of cancer using cell-free DNA methylation. *Computational and Structural Biotechnology Journal*, *20*(2), 26–39. https://doi.org/10.1016/j.csbj.2021.12.001

Shuihui Wang Lorkowski, Dermawan, J. K., & Rubin, B. P. (2023). The practical utility of AI-assisted molecular profiling in the diagnosis and management of cancer of unknown primary: an updated review. *Virchows Archiv*, *484*(2), 369–375. https://doi.org/10.1007/s00428-023-03708-1

Sirajee, A. S., Kabiraj, D., & De, S. (2024). Cell-free nucleic acid fragmentomics: A non-invasive window into cellular epigenomes. *Translational Oncology*, *49*(22), 102085. https://doi.org/10.1016/j.tranon.2024.102085

Spurgeon, L., Mitchell, C., Cook, N., & Conway, A.-M. (2025). Cancer of unknown primary: the hunt for its elusive tissue-of-origin – is it time to call off the search? *British Journal of Cancer*, *132*. https://doi.org/10.1038/s41416-025-03073-7

Stanley, K. E., Tatjana Jatsenko, Tuveri, S., Dhanya Sudhakaran, Lannoo, L., Kristel Van Calsteren, Marie de Borre, Ilse Van Parijs, Leen Van Coillie, Van, K., De, R., Liesbeth Lenaerts, Tejpar, S., Punie, K., Rengifo, L. Y., Vandenberghe, P., Thienpont, B., & Joris Robert Vermeesch. (2024). Cell type signatures in cell-free DNA fragmentation profiles reveal disease biology. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-46435-0

Stark, L., Atsuko Kasajima, Stögbauer, F., Schmidl, B., Jakob Rinecker, Holzmann, K., Färber, S., Pfarr, N., Steiger, K., Wollenberg, B., Ruland, J., Winter, C., & Wirth, M. (2024). Head and neck cancer of unknown primary: unveiling primary tumor sites through machine learning on DNA methylation profiles. *Clinical Epigenetics*, *16*(1). https://doi.org/10.1186/s13148-024-01657-3

Sun, M., Xu, B., Chen, C., Zhu, Y., Li, X., & Chen, K. (2024). Tissue of origin prediction for cancer of unknown primary using a targeted methylation sequencing panel. *Clinical Epigenetics*, *16*(1). https://doi.org/10.1186/s13148-024-01638-6

Taryma-Leśniak, O., Buczkowski, K., Klepinowski, T., Jankowski, M., Retfiński, S., Sagan, B., Sagan, L., Szutowicz, E., Winiarski, K., Masztalewicz, M., Nowacki, P., Sobalska-Kwapis, M., Strapagiel, D., Kloc, W., Lipska-Ziętkiewicz, B., Iżycka-Świeszewska, E., & Wojdacz, T. K. (2024). Classification of gliomas based on genome-wide DNA

methylation profiles and immunohistochemical analysis: a short study. *Folia Neuropathologica*, *62*(4), 362–375. https://doi.org/10.5114/fn.2024.147538

van, Normastuti Adhini Tantyo, Nils Evander, Hentschel, A. E., Wever, B. M. M., Ramaker, J., Bootsma, S., Fransen, M. F., Lenos, K., Vermeulen, L., Schneiders, F. L., Idris Bahce, Nieuwenhuijzen, J. A., Renske D.M. Steenbergen, D. Michiel Pegtel, Moldován, N., & Florent Moulière. (2023). Real-time analysis of the cancer genome and fragmentome from plasma and urine cell-free DNA using nanopore sequencing. *EMBO Molecular Medicine*, *15*(12). https://doi.org/10.15252/emmm.202217282

Wang, J., Chen, L., Zhang, X., Tong, Y., & Zheng, T. (2021). OCRDetector: Accurately Detecting Open Chromatin Regions via Plasma Cell-Free DNA Sequencing Data. *International Journal of Molecular Sciences*, *22*(11), 5802–5802. https://doi.org/10.3390/ijms22115802

Zarkavelis, G., Yerolatsite, M., Nanteznta Torounidou, Anna-Lea Amylidi, Pinelopi Gogou, Athanasia Karavasili, Eleftherios Kampletsas, & Mauri, D. (2024). The Use of Circulating Tumor DNA in Cancers of Unknown Primary: A Review of the Literature. *European Journal of Clinical Oncology*, *6*(2), 1–4. https://www.iomcworld.org/articles/the-use-of-circulating-tumor-dna-in-cancers-of-unknown-primary-a-review-of-the-literature-110391.html

Zhang, N., Liu, Z., Li, K., Xing, X., Long, C., Liu, F., She, B., & Che, N. (2023). DNA Methylation Analysis of the SHOX2 and RASSF1A Panel Using Cell-Free DNA in the Diagnosis of Malignant Pleural Effusion. *Journal of Oncology*, *2023*, 1–9.

https://doi.org/10.1155/2023/5888844

Zhang, S., He, S., Zhu, X., Wang, Y., Xie, Q., Song, X., Xu, C., Wang, W., Xing, L., Xia, C., Wang, Q., Li, W., Zhang, X., Yu, J., Ma, S., Shi, J., & Gu, H. (2023). DNA methylation profiling to determine the primary sites of metastatic cancers using formalin-fixed paraffin-embedded tissues. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-41015-0

Zhang, X., Li, J., Zhuang, Z., Wang, J., Bu, Z., & Lan, X. (2024). Challenges and prospects of cell-free DNA in precision oncology. *Medicine Plus*, *1*(4), 100059–100059. https://doi.org/10.1016/j.medp.2024.100059

Zhou, X., Cheng, Z., Dong, M., Liu, Q., Yang, W., Liu, M., Tian, J., & Cheng, W. (2022). Tumor fractions deciphered from circulating cell-free DNA methylation for cancer early diagnosis. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-35320-3