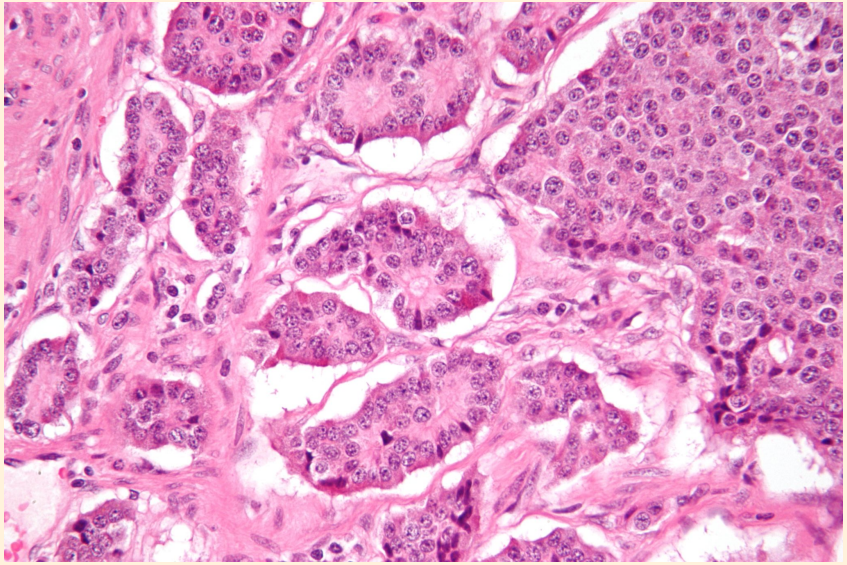
Neuroendocrine Neoplasm DNA Methylation Biomarker Study

Eric Eleam, Computer Science, Temple University
Lee Lab, Fox Chase Cancer Center

Neuroendocrine Neoplasms (NENs)

NENs are tumors that originate from neuroendocrine cells, which possess characteristics of both nerve and hormone-producing endocrine cells. These tumors can be challenging to diagnose early, as they often grow silently and remain asymptomatic until advanced stages. NENs vary widely in their behavior and aggressiveness, and they are graded based on their growth patterns and cellular differentiation. Low-grade (Grade 3) NENs are well-differentiated and tend to grow slowly, whereas high-grade (Grade 1) carcinomas are poorly differentiated, more aggressive, and exhibit rapid growth. Understanding these distinctions is crucial for patient care and treatment strategies.

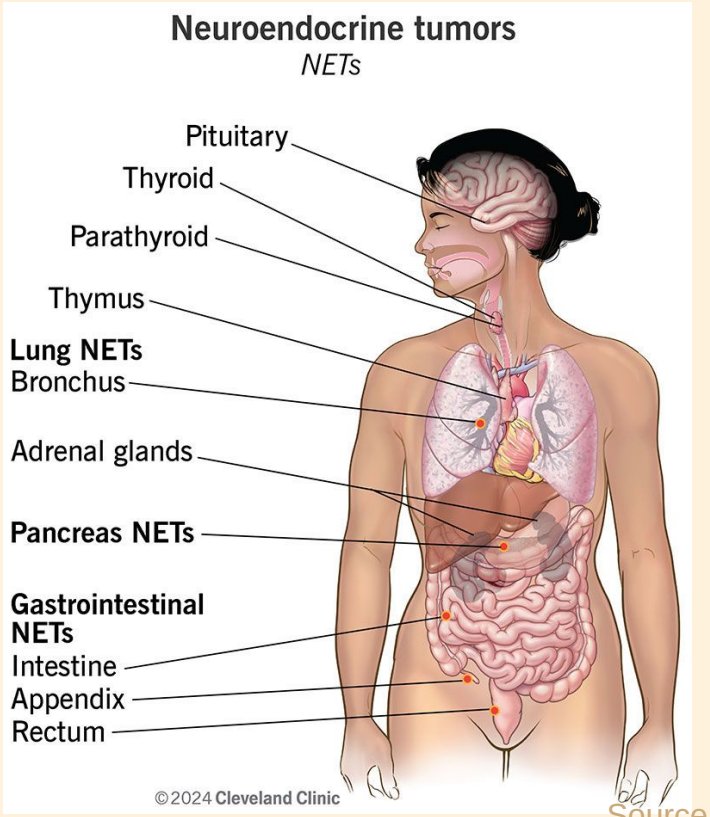


Challenges

- A major clinical challenge with NENs lies in both their initial detection and accurate subtype classification. These tumors often remain asymptomatic in early stages, leading to delayed diagnosis leading the disease to continuously advance.
- Even after detection, distinguishing between low-grade, well-differentiated tumors (neuroendocrine tumors, NETs) and high-grade, poorly differentiated carcinomas (neuroendocrine carcinomas, NECs) is challenging. This distinction is critical for guiding treatment decisions and predicting patient outcomes, but it can be difficult to accurately identify subtypes only using histopathological methods.

Primary sites of NENs

NENs most commonly occur in the gastropancreatic tract and the lungs. Within the gastropancreatic system, the small intestine, rectum, and pancreas are frequent sites of origin with the small intestine being the most dominant site for well differentiated cells. Other sites for NENs outside of the gastropancreatic and lung systems are the thymus, skin, adrenal gland and more.



Source

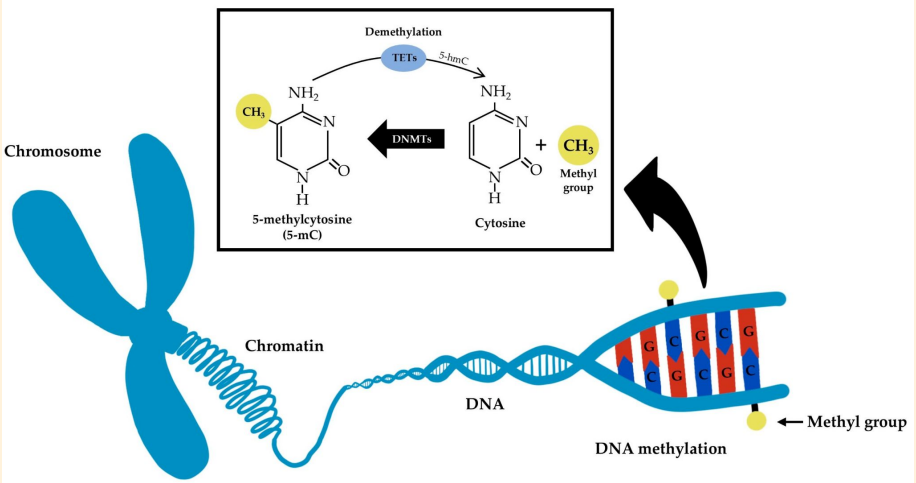
Lung NEN

Lung NENs are a subset of neuroendocrine neoplasms found in the lungs, they range from slow growing NETs to fast acting NECs. Accurate identification of Lung NEN subtypes is difficult but is needed for effective patient care.

Source

DNA Methylation

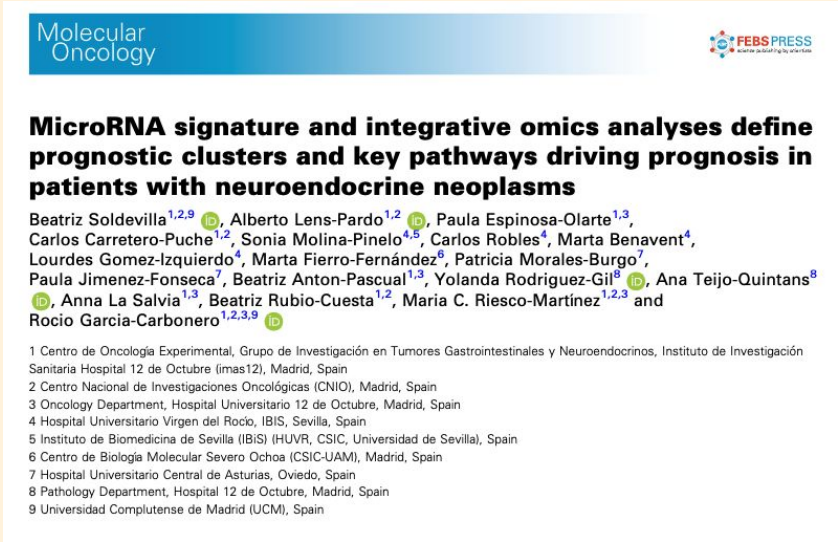
DNA methylation is an epigenetic modification that regulates gene expression by adding methyl groups to DNA, often affecting how genes are turned on or off. In cancer, abnormal methylation patterns are commonly observed and can provide valuable biomarkers for diagnosis and classification. In the context of NENs, DNA methylation offers a promising approach to overcome the limitations of histological analysis. By combining DNA methylation data with machine learning, we aim to develop clinically actionable and cost effective tools for early detection and accurate classification of NEN subtypes.



Source

Related Paper

This study investigated the role of microRNAs in neuroendocrine neoplasms from both lung and gastroenteropancreatic origins. The researchers identified 8-miRNA signature able to predict survival rate of patients with NENs into three prognostic groups (5 year survival of 80%, 66%, and 36%). These miRNAs were linked to 71 target genes involved in key cancer pathways, with 28 genes associated with patient survival. Importantly, five CpG sites were found to epigenetically regulate the expression of these miRNAs. It closely relates to my research, which also focuses on using epigenetic tools like DNA methylation to classify lung NENs and improve diagnostic accuracy.



Molecular Oncology

FEBS PRESS

MicroRNA signature and integrative omics analyses define prognostic clusters and key pathways driving prognosis in patients with neuroendocrine neoplasms

Beatriz Soldevilla[1,2,9], Alberto Lens-Pardo[1,2], Paula Espinosa-Olarte[1,3], Carlos Carretero-Puche[1,2], Sonia Molina-Pinelo[4,5], Carlos Robles[4], Marta Benavent[4], Lourdes Gomez-Izquierdo[4], Marta Fierro-Fernández[6], Patricia Morales-Burgo[7], Paula Jimenez-Fonseca[7], Beatriz Anton-Pascual[1,3], Yolanda Rodriguez-Gil[8], Ana Teijo-Quintans[8], Anna La Salvia[1,3], Beatriz Rubio-Cuesta[1,2], Maria C. Riesco-Martínez[1,2,3] and Rocio Garcia-Carbonero[1,2,3,9]

1 Centro de Oncología Experimental, Grupo de Investigación en Tumores Gastrointestinales y Neuroendocrinos, Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Spain
2 Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid, Spain
3 Oncology Department, Hospital Universitario 12 de Octubre, Madrid, Spain
4 Hospital Universitario Virgen del Rocío, IBIS, Sevilla, Spain
5 Instituto de Biomedicina de Sevilla (IBiS) (HUVR, CSIC, Universidad de Sevilla), Spain
6 Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Madrid, Spain
7 Hospital Universitario Central de Asturias, Oviedo, Spain
8 Pathology Department, Hospital 12 de Octubre, Madrid, Spain
9 Universidad Complutense de Madrid (UCM), Spain

Pancreatic NEN

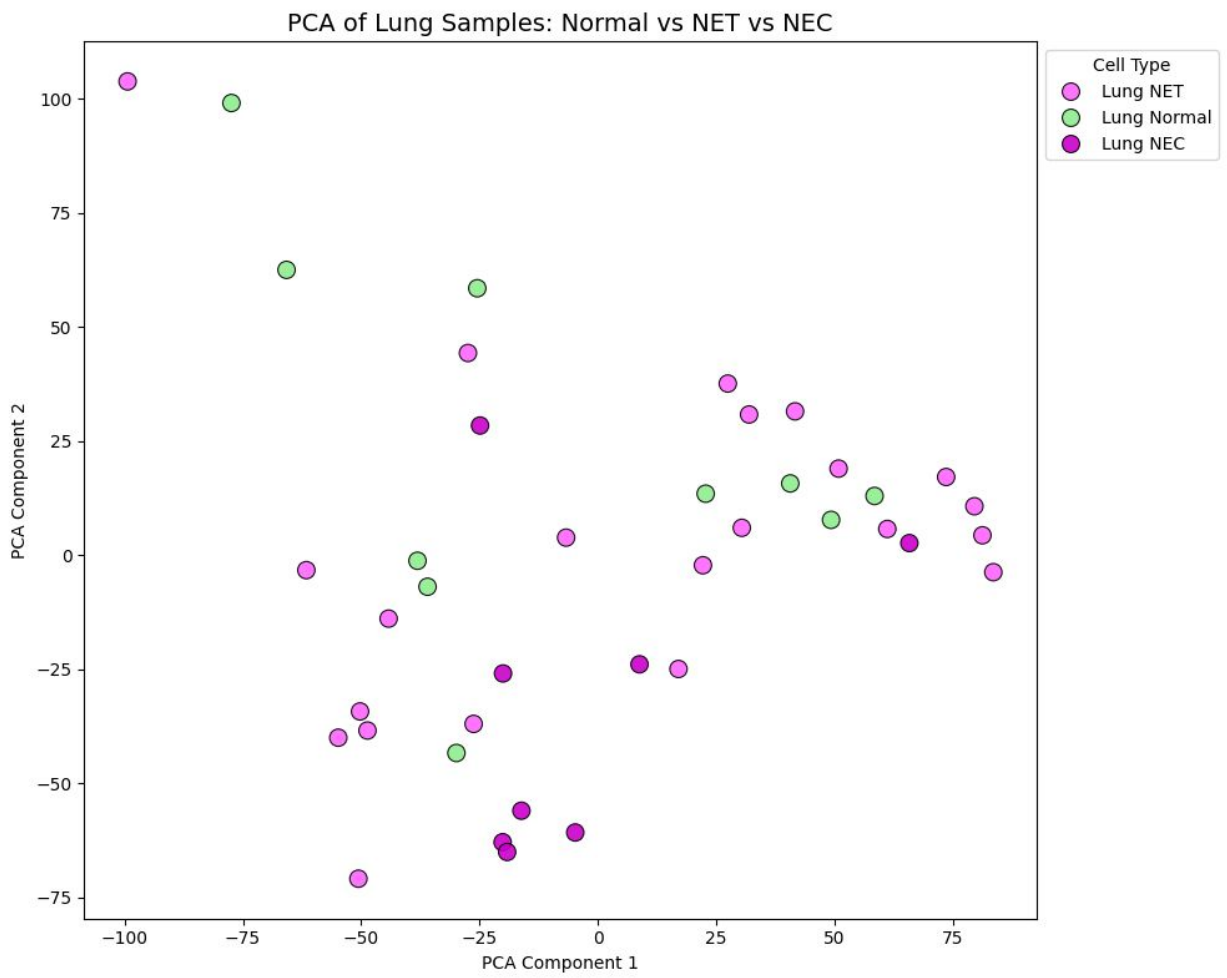
Steve Jobs, co founder of Apple, was diagnosed in 2003 with a pancreatic cancer known as pancreatic neuroendocrine tumor (PNET). PNETs are so uncommon there is no standard of care or standard treatment for these type of tumors. Jobs lived for 8 years after his diagnosis until succumbing to his condition in 2011.

Data

Cell Types	Number of Lung Samples
Normal	8
NET	17
NEC	8

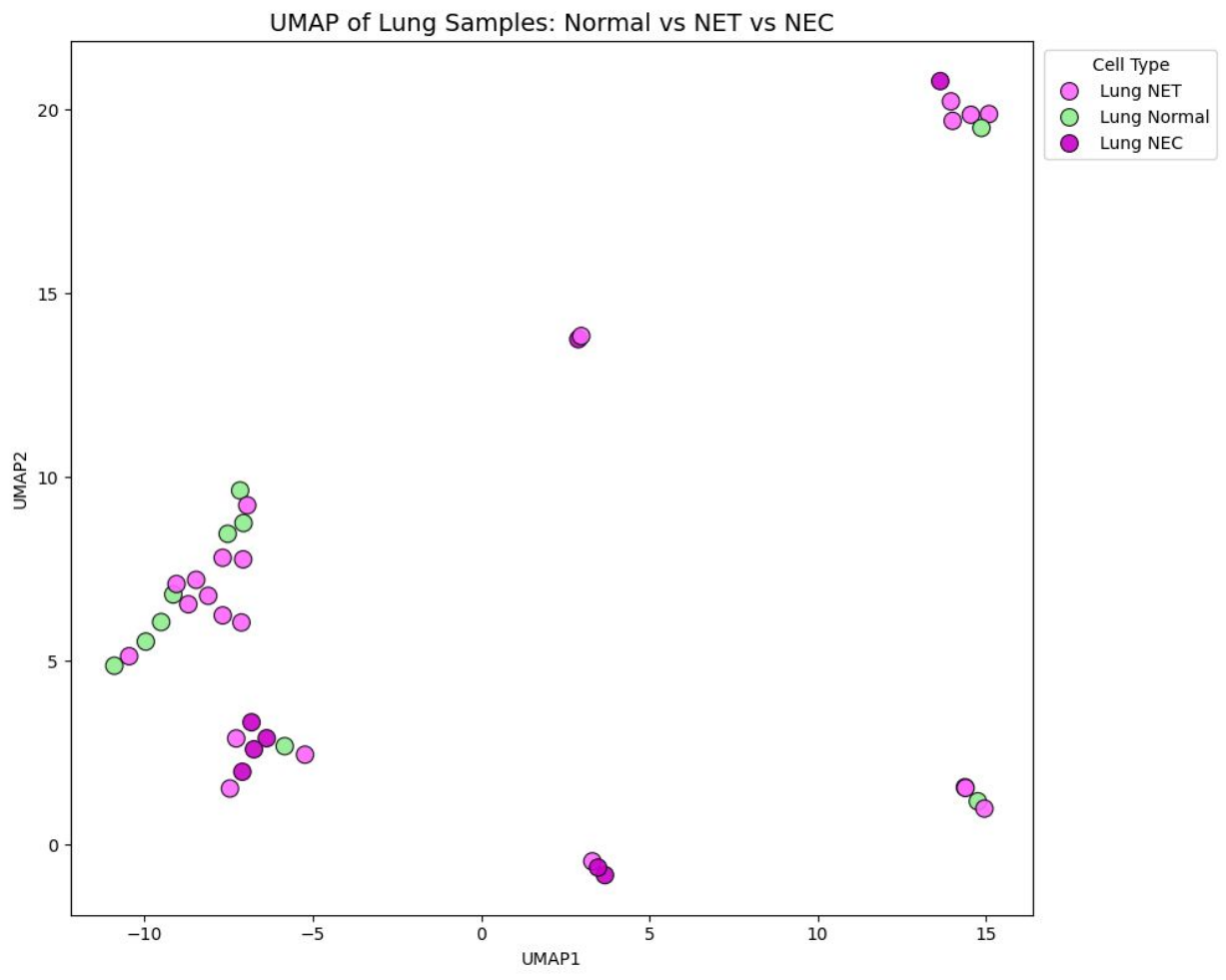
Total of 33 Lung Samples

PCA

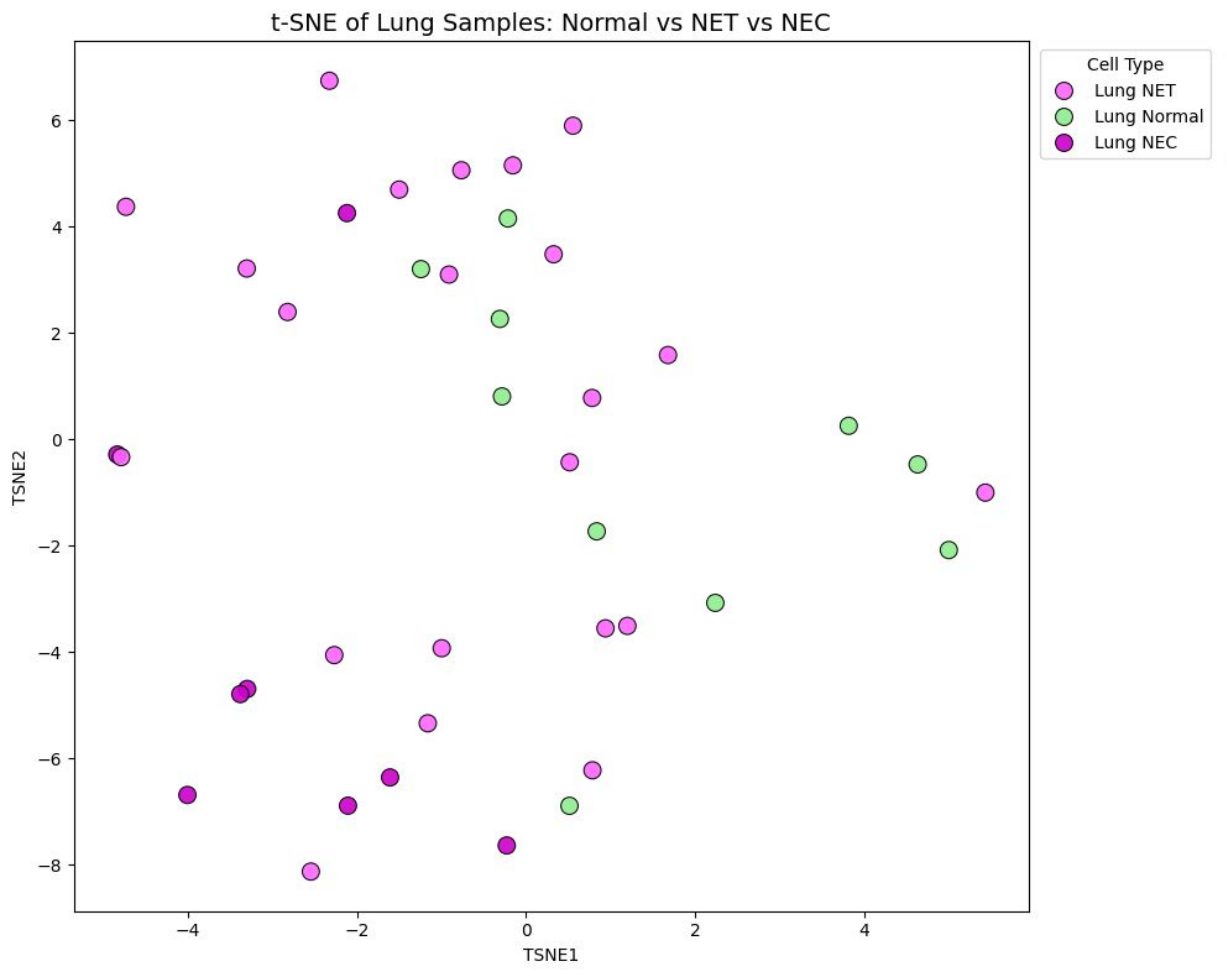


They’re not separable in linear space <- hard problem.

UMAP


UMAP of Lung Samples: Normal vs NET vs NEC
Cell Type
Lung NET
Lung Normal
Lung NEC
UMAP1
UMAP2
20
15
10
5
0
−10
−5
0
5
10
15

T-SNE


t-SNE of Lung Samples: Normal vs NET vs NEC
Cell Type
Lung NET
Lung Normal
Lung NEC
TSNE1
TSNE2
6
4
2
0
−2
−4
−6
−8
−4
−2
0
2
4

Kolmogorov Smirnov Test (KS Test)

The KS Test is a test that compares two samples to determine whether they are drawn from the same distribution. It is especially useful when you don't want to assume a specific distribution type like normality. The test calculates the maximum difference between the cumulative distribution functions (CDFs) of the two samples. In Jupyter Notebook I utilized scipy.stats and the ks_2samp() to implement the test, the function returns a test statistic value and a p value. A low p value suggests the two samples likely come from different distributions, while a high p value indicates similarity in distribution.

False Discovery Rate (FDR)

When performing multiple statistical tests (KS Test), the chance of false positives increases. It ensures that only a controlled proportion of significant results are expected to be false discoveries. The function false_discovery_control() applies the Benjamini Hochberg to find these false positives. In my research I used false_discovery_control() on all the p values I received from the KS Test.

Analysis pipeline

Subset Data

From the original dataset, create three separate data frames containing only the samples from each cell type group: Normal, NET, and NEC.

Calculate Group Means

For each subset, compute the mean value for every cgID.

Compute Group Differences

Make three new group means, NET - Normal, NEC - NET, NEC - Normal. Append these difference values as new columns to the original dataframe.

Statistical Testing

Apply the KS Test on the difference columns to get the p values for each comparison group. Then perform the FDR Test on the KS Test p values to control for false positives across comparisons

Rank Differences

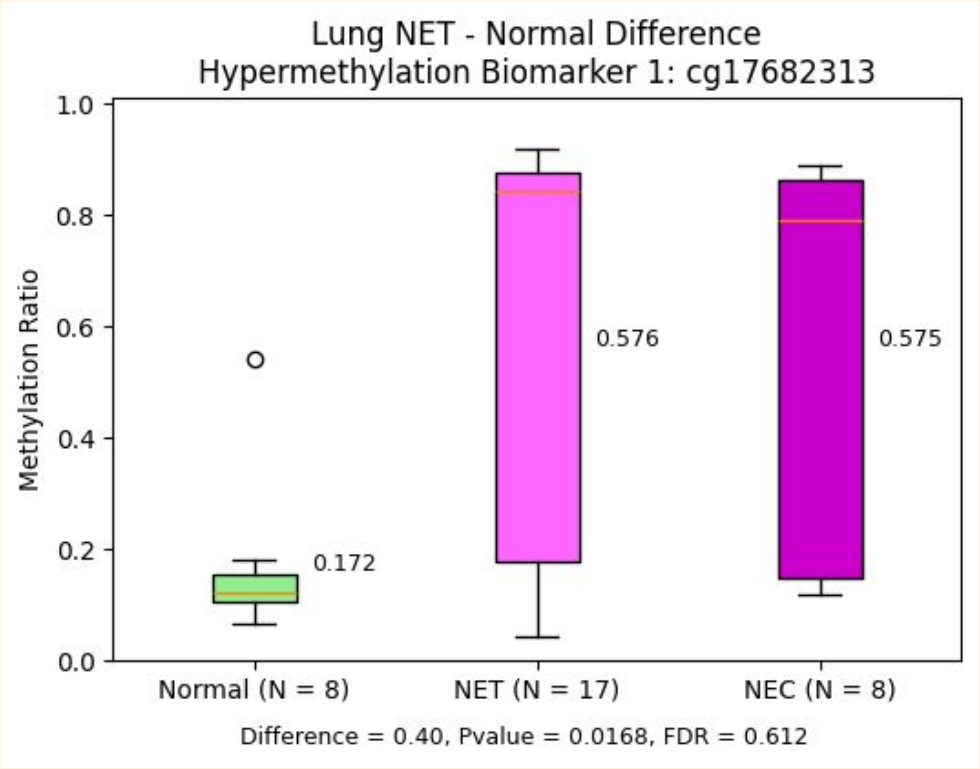
Sort the difference comparisons (eg. NET - Normal) and find the top 5 and bottom 5 cgIDs for each of the three groups and create box plots.

Data Visualization

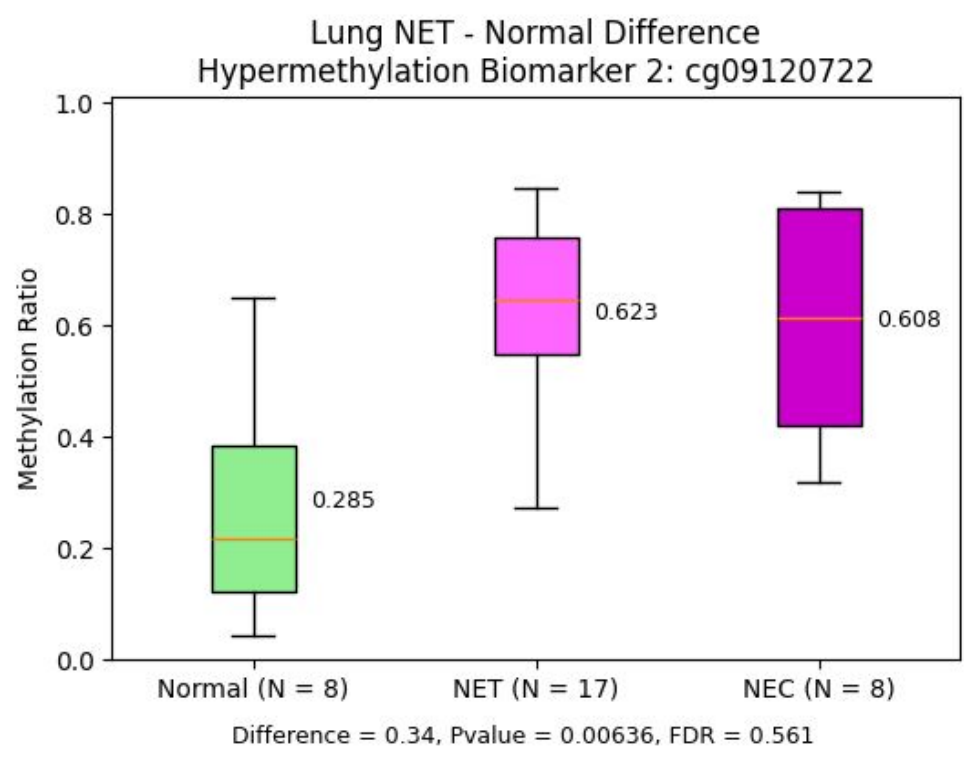
Use the original data frame to generate data visualization graphs using PCA, UMAP and t-SNE to analyze sample clustering and separability.

Hypermethylated Biomarker Candidates NET - Normal

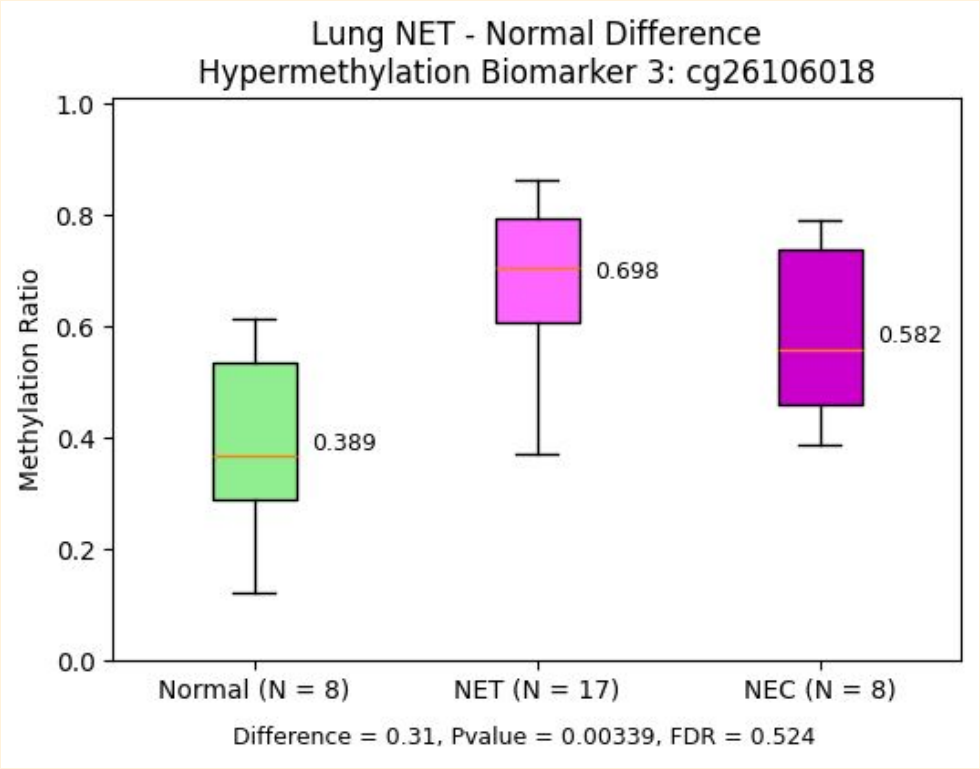
DNA Methylation Biomarker Candidate 1
NET - Normal



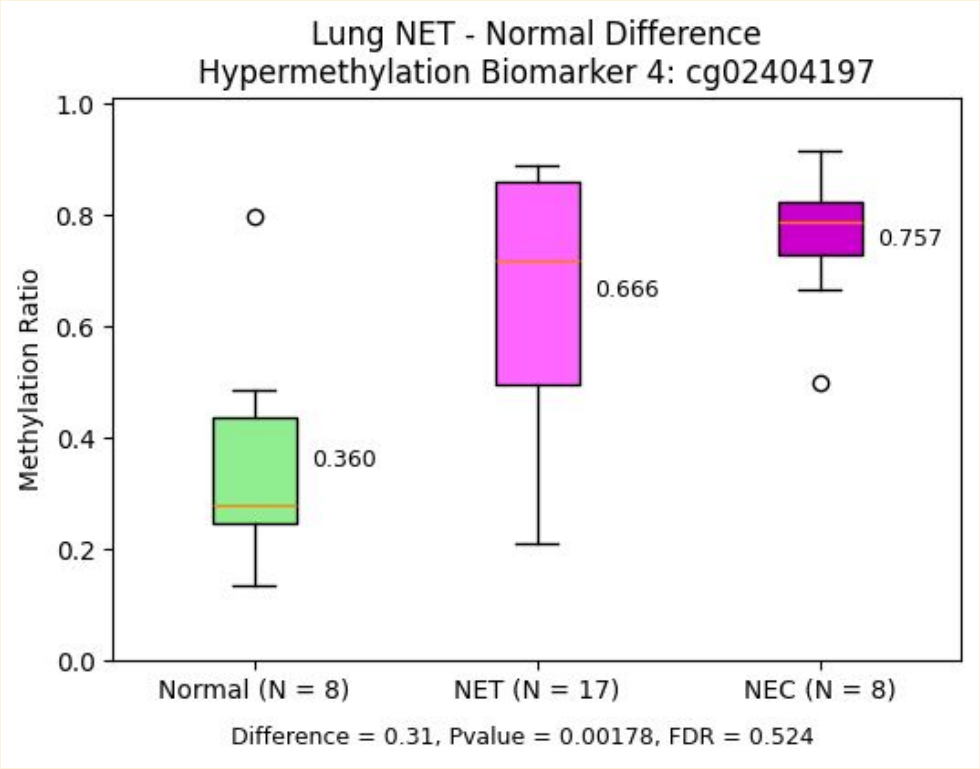
DNA Methylation Biomarker Candidate 2 NET - Normal



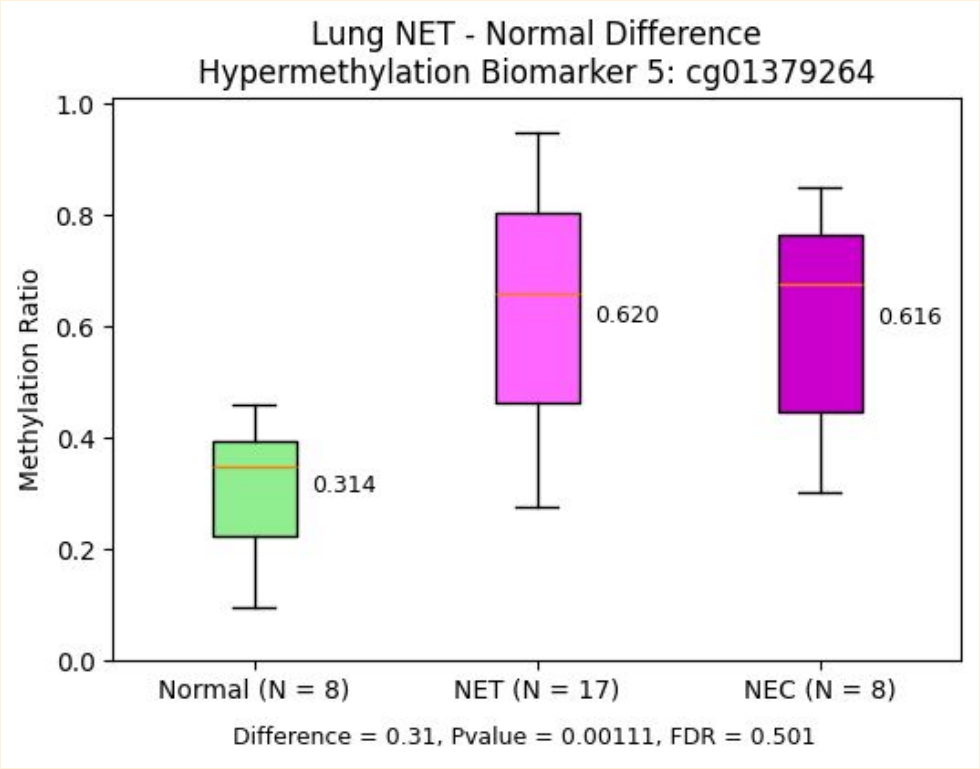
DNA Methylation Biomarker Candidate 3 NET - Normal



DNA Methylation Biomarker Candidate 4 NET - Normal

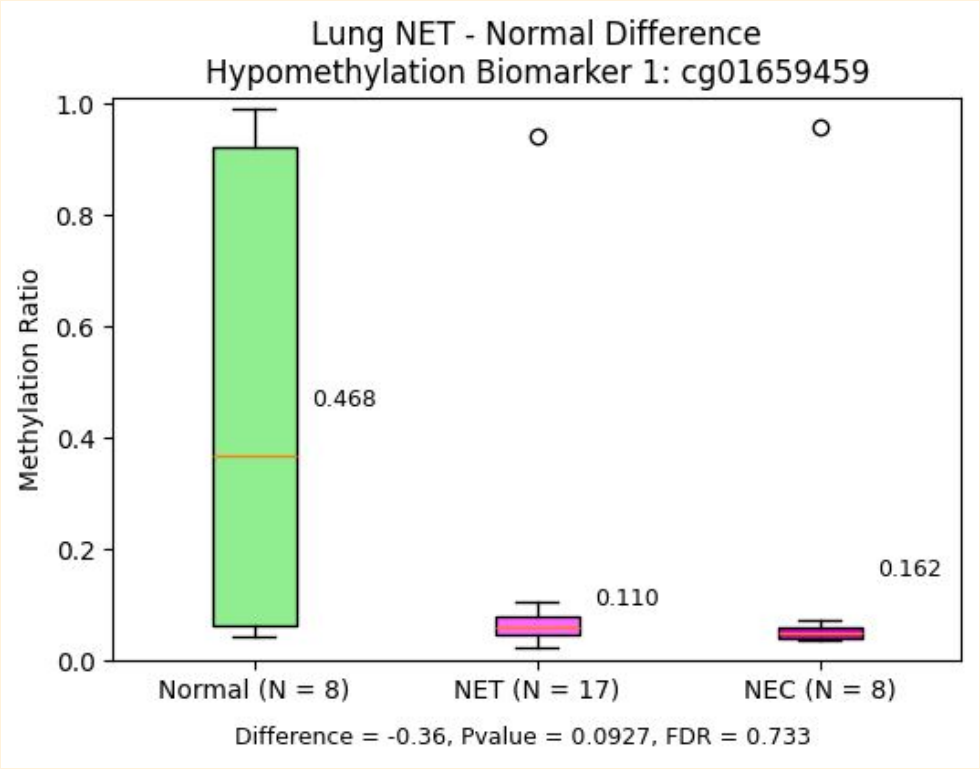


DNA Methylation Biomarker Candidate 5 NET - Normal

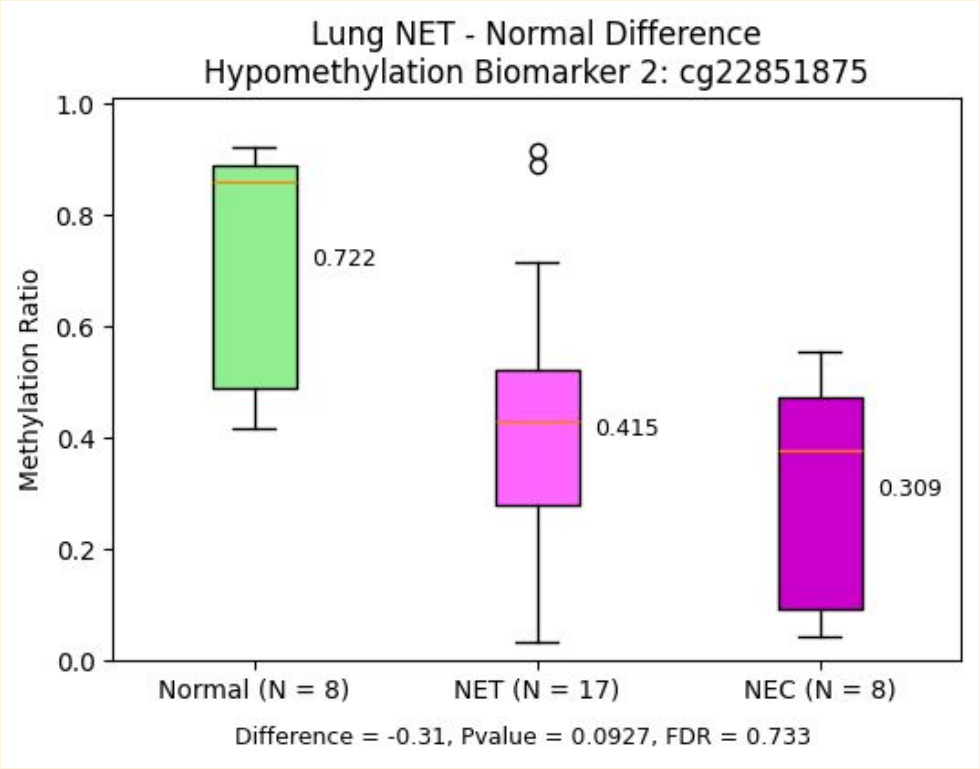


Hypomethylated Biomarker Candidates NET - Normal

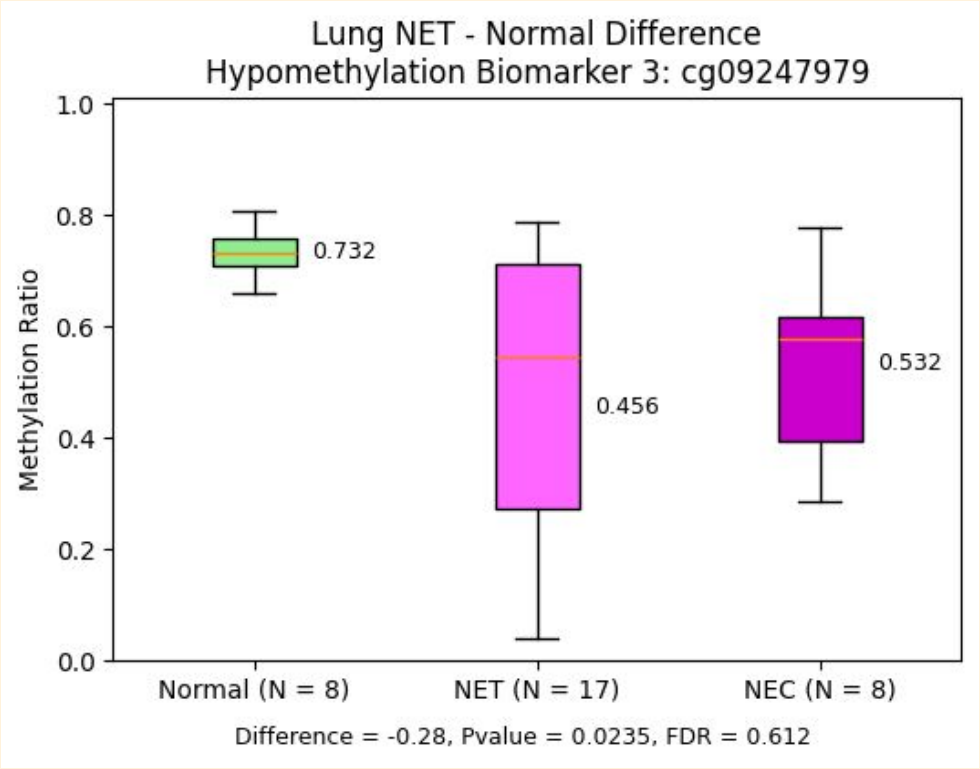
DNA Methylation Biomarker Candidate 1
NET - Normal



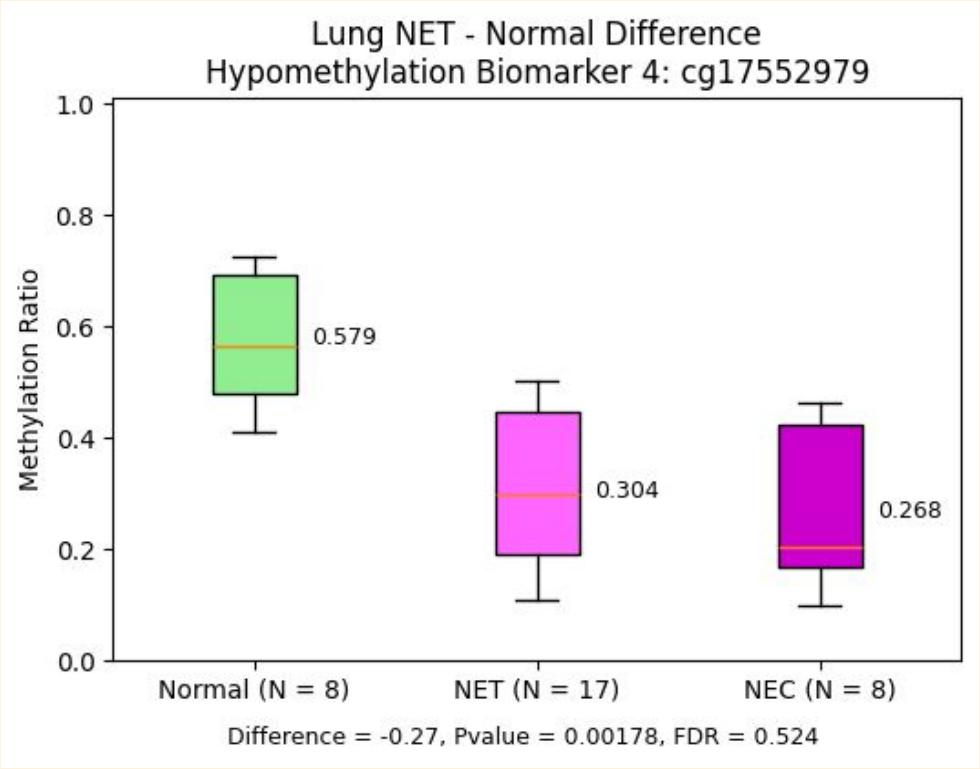
DNA Methylation Biomarker Candidate 2 NET - Normal



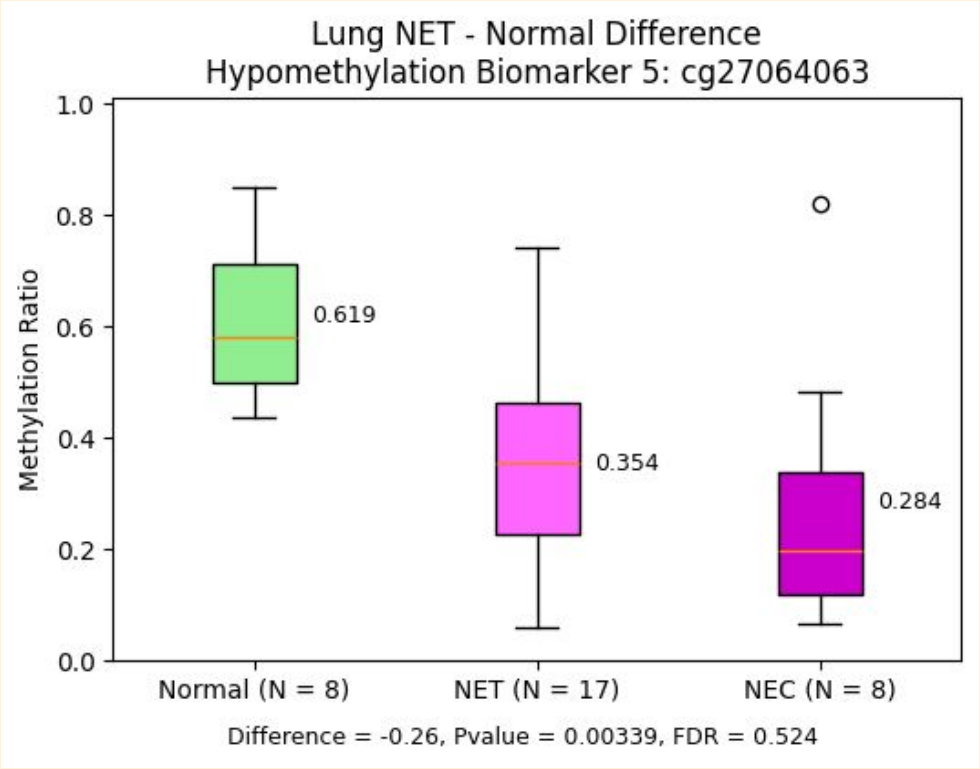
DNA Methylation Biomarker Candidate 3 NET - Normal



DNA Methylation Biomarker Candidate 4 NET - Normal

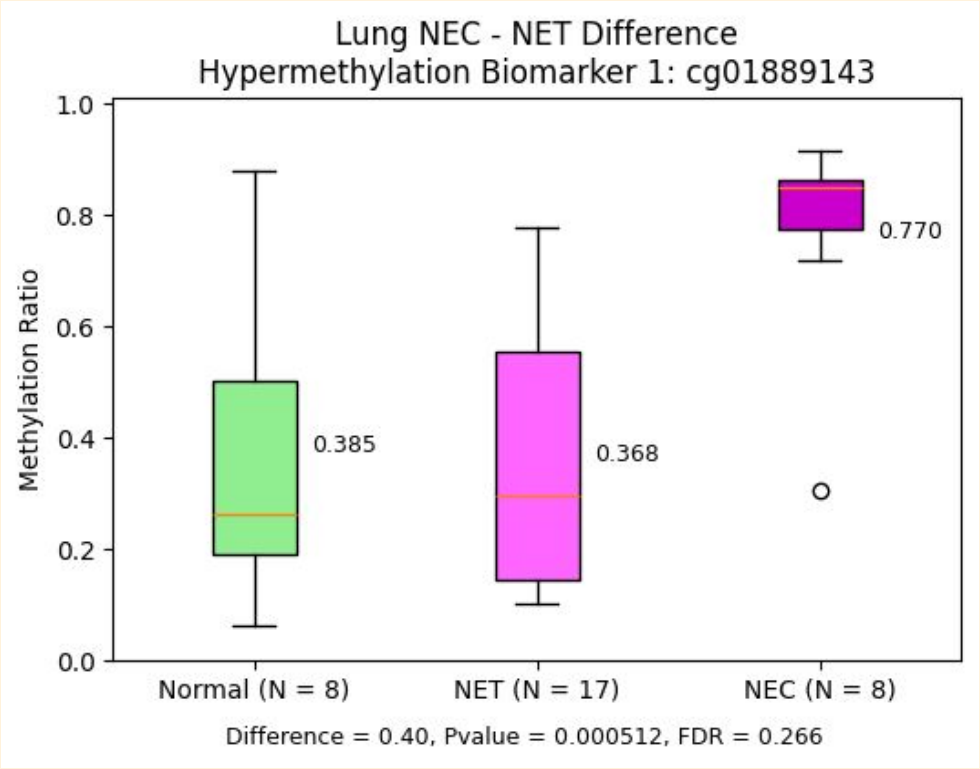


DNA Methylation Biomarker Candidate 5 NET - Normal

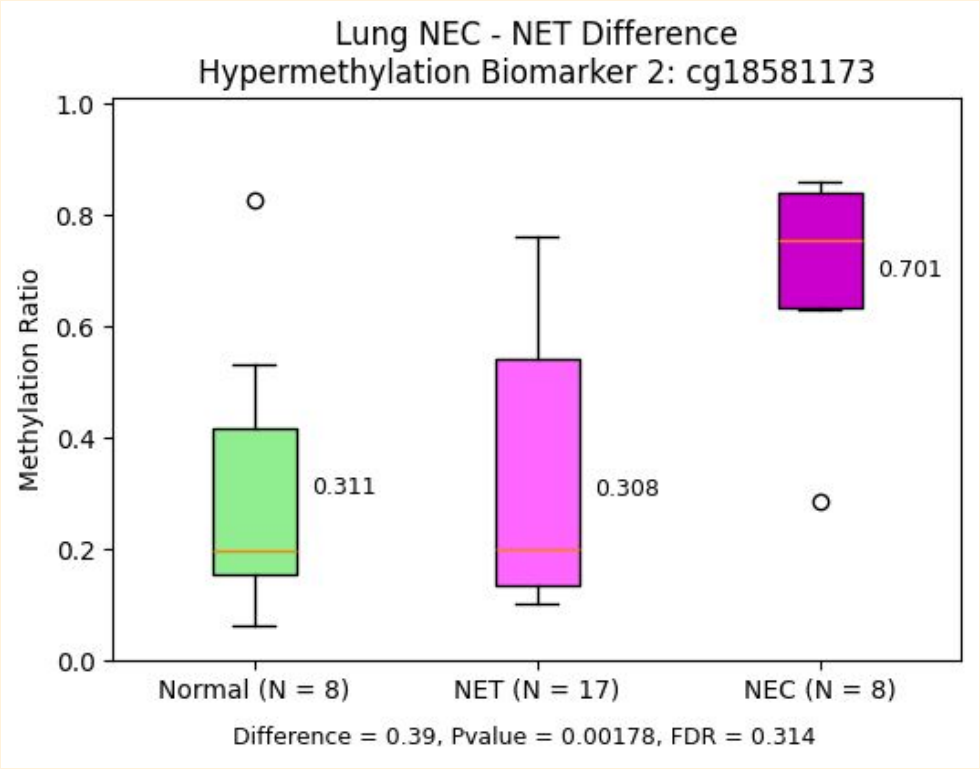


Hypermethylated Biomarker Candidates NEC - NET

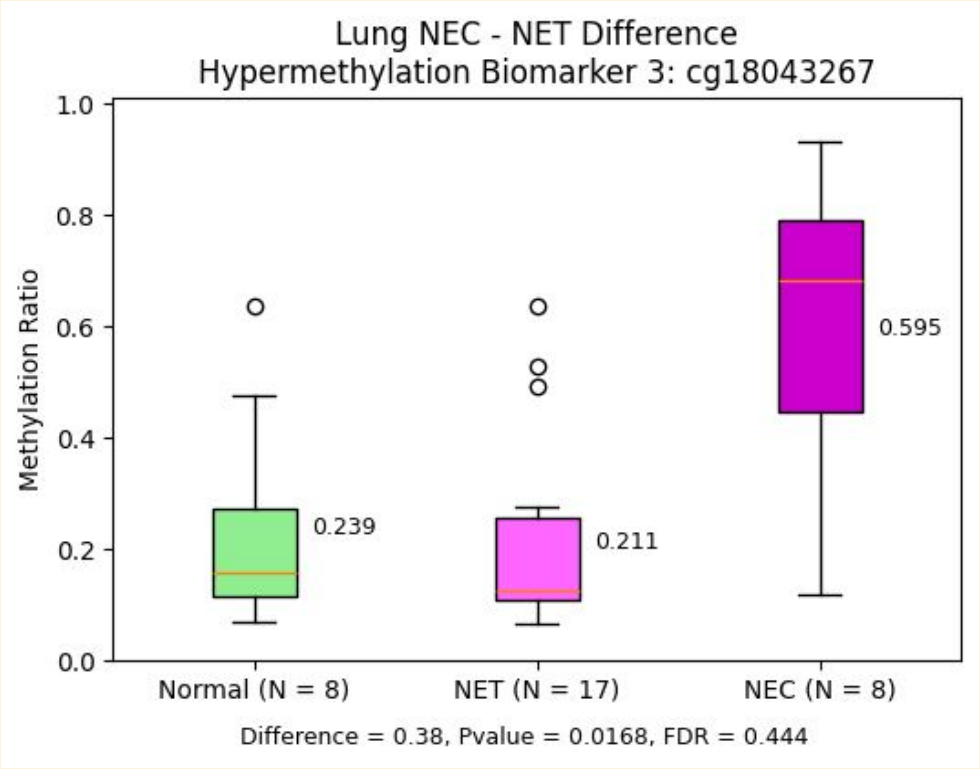
DNA Methylation Biomarker Candidate 1 NEC - NET



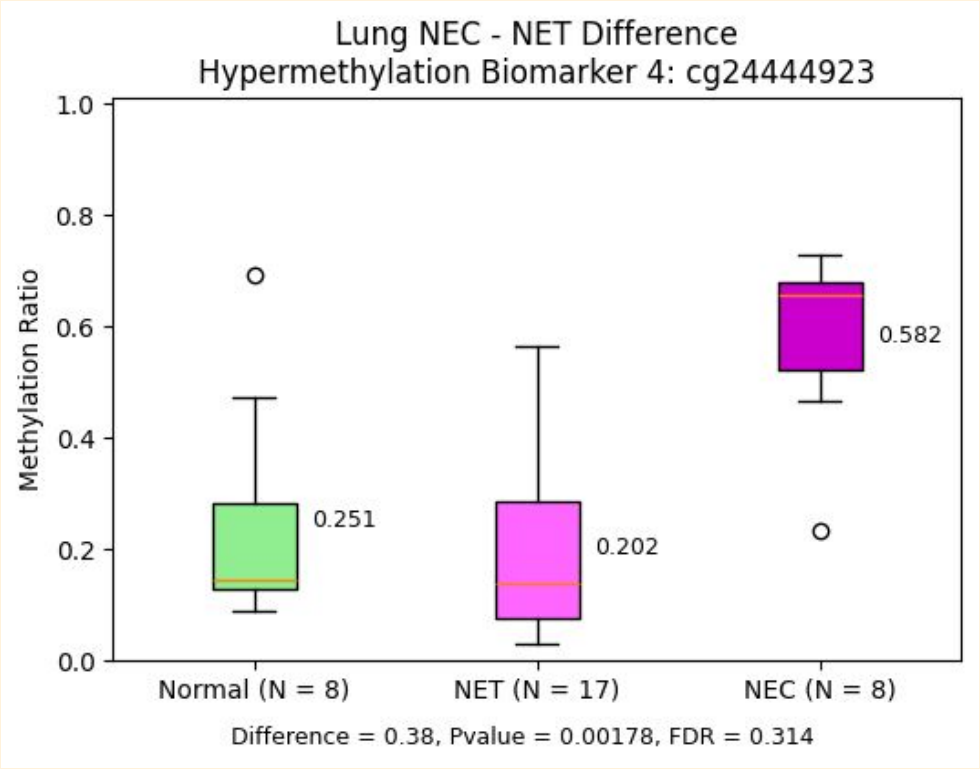
DNA Methylation Biomarker Candidate 2
NEC - NET



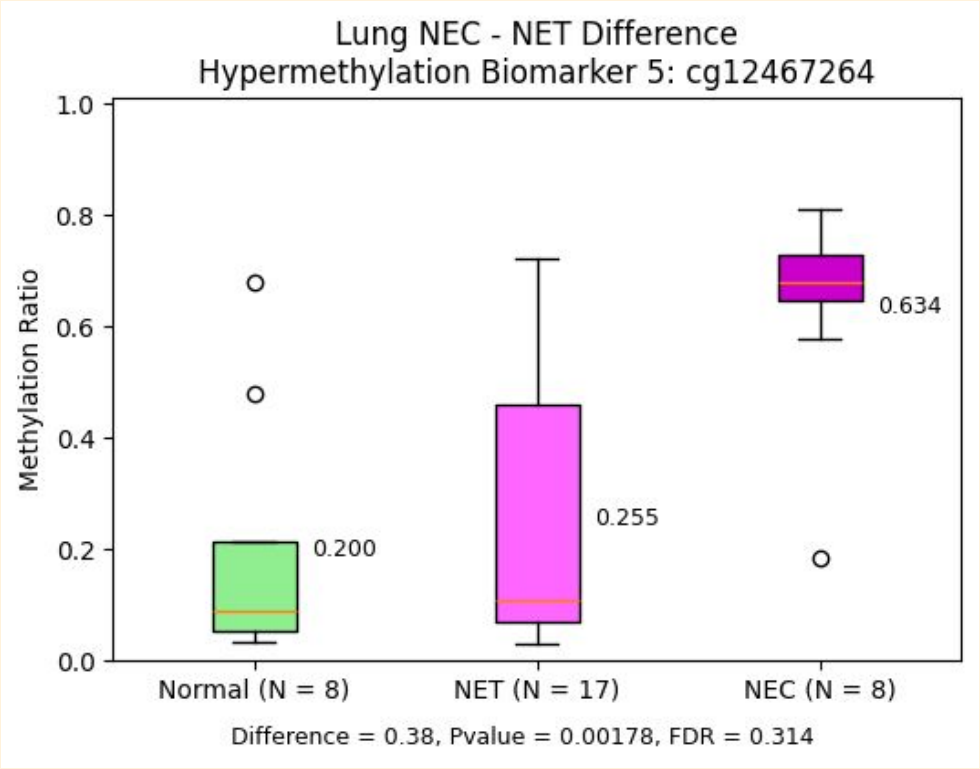
DNA Methylation Biomarker Candidate 3 NEC - NET



DNA Methylation Biomarker Candidate 4
NEC - NET

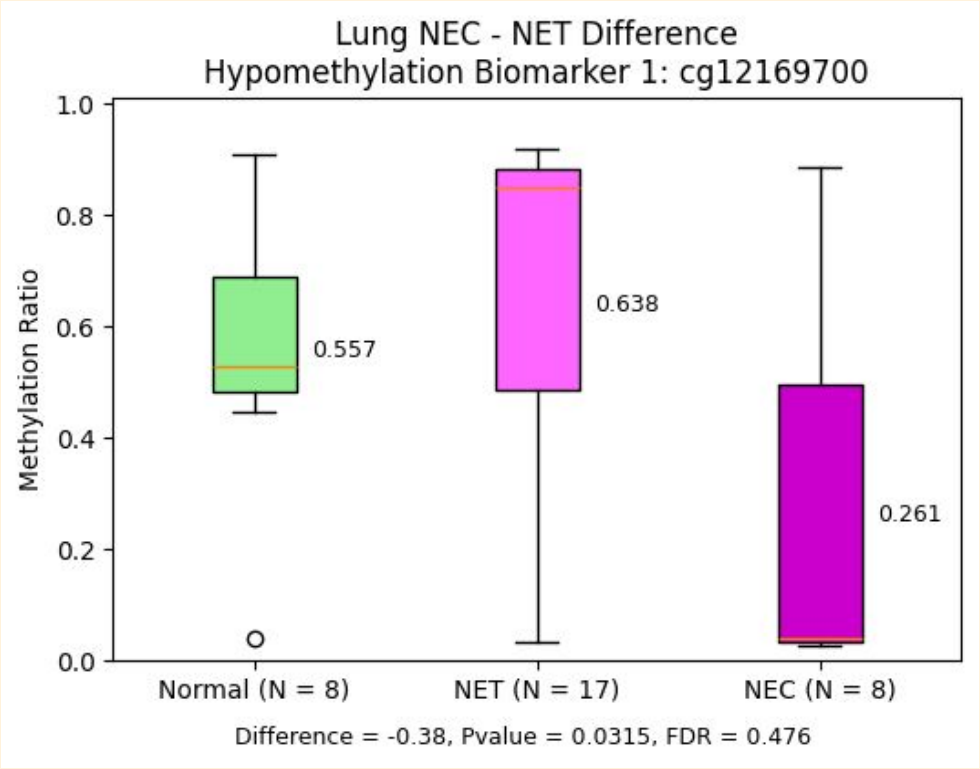


DNA Methylation Biomarker Candidate 5
NEC - NET

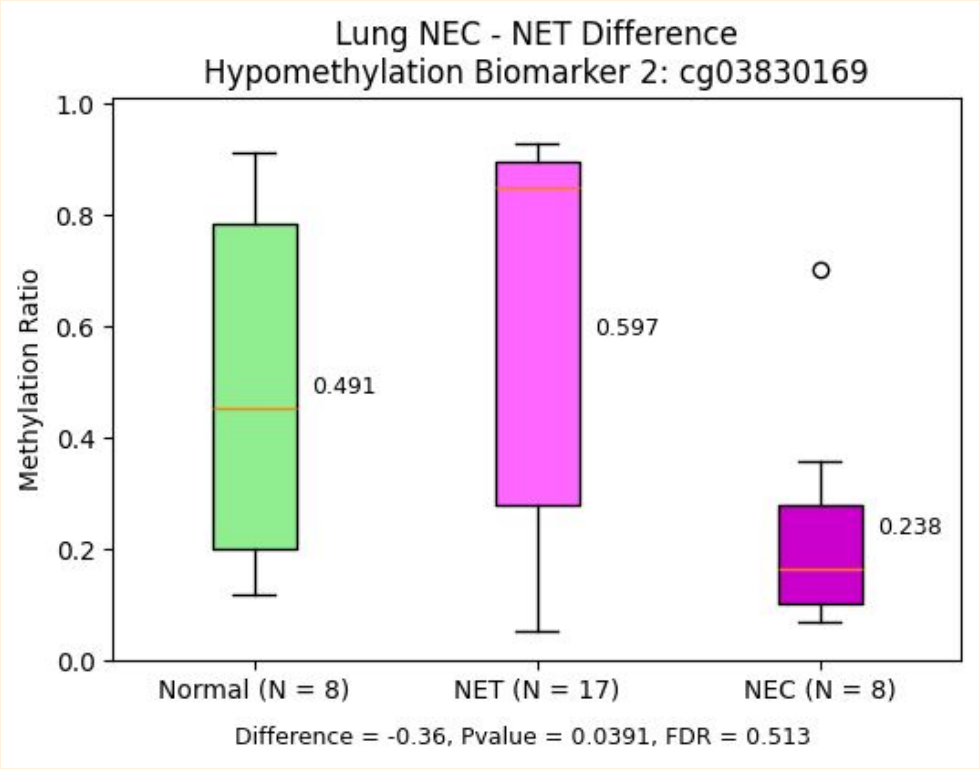


Hypomethylated Biomarker Candidates NEC - NET

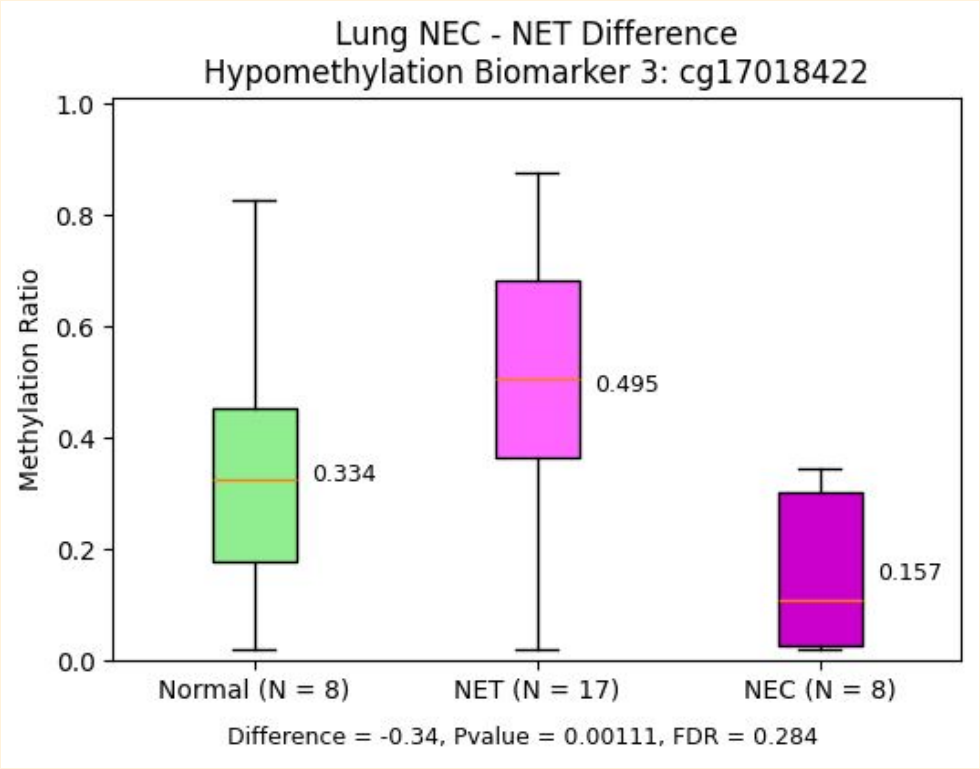
DNA Methylation Biomarker Candidate 1
NEC - NET



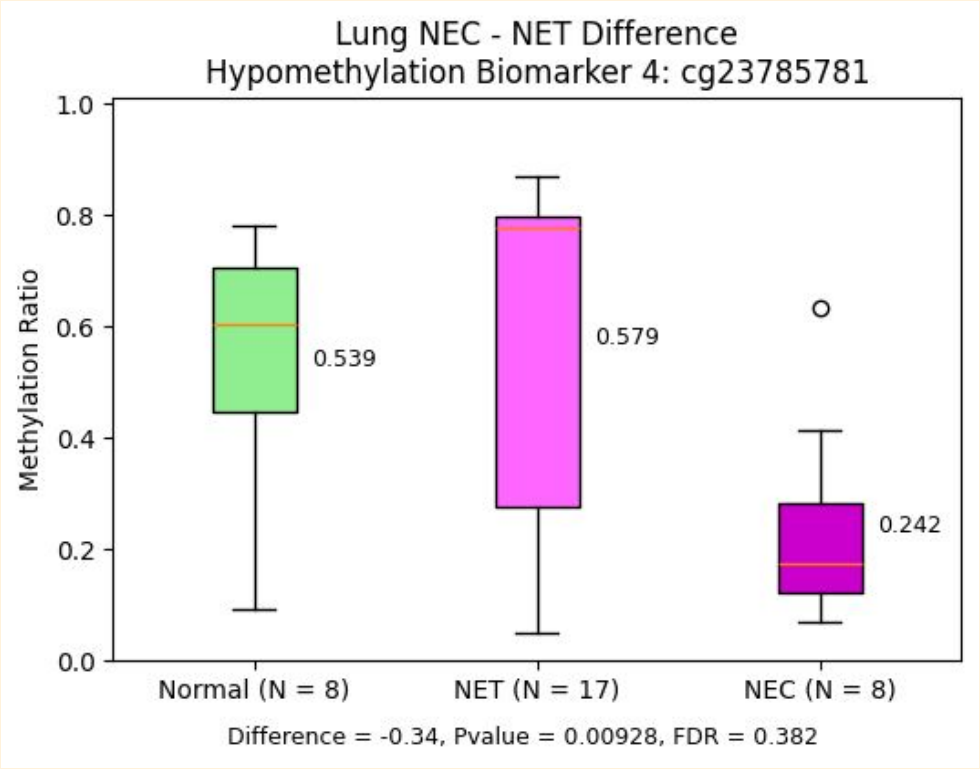
DNA Methylation Biomarker Candidate 2
NEC - NET



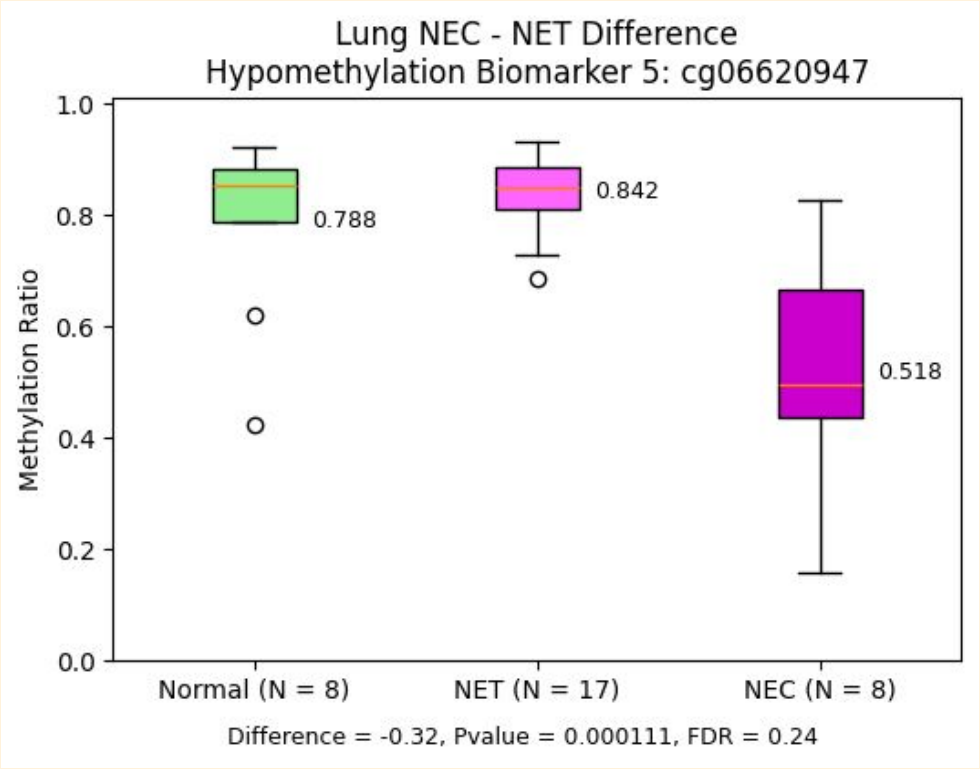
DNA Methylation Biomarker Candidate 3 NEC - NET



DNA Methylation Biomarker Candidate 4 NEC - NET

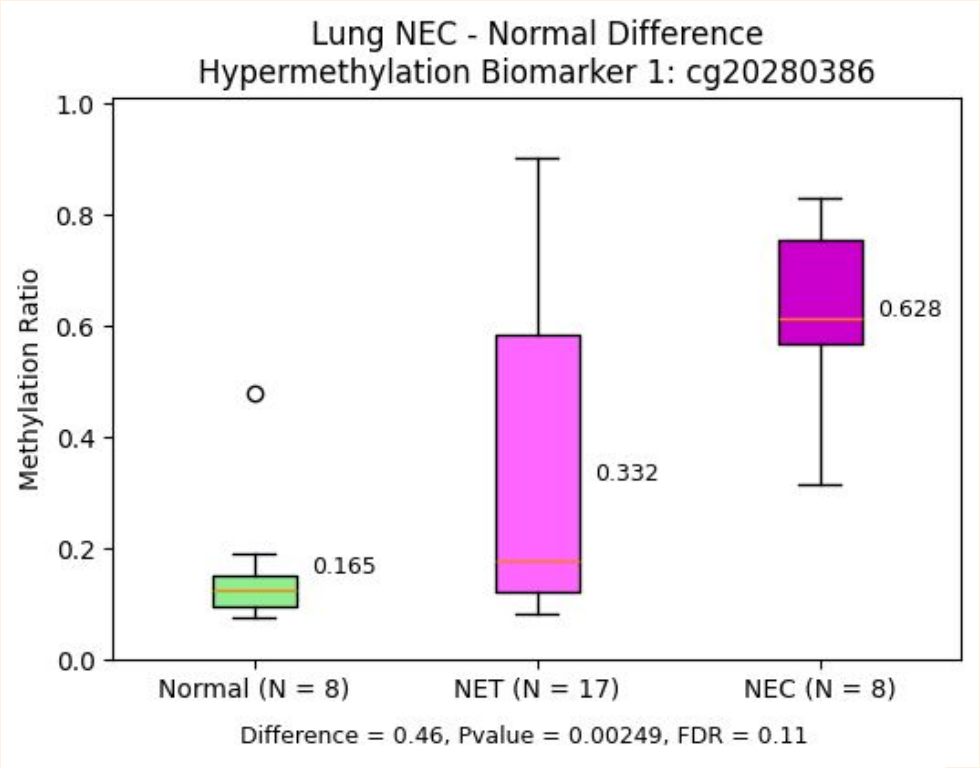


DNA Methylation Biomarker Candidate 5
NEC - NET

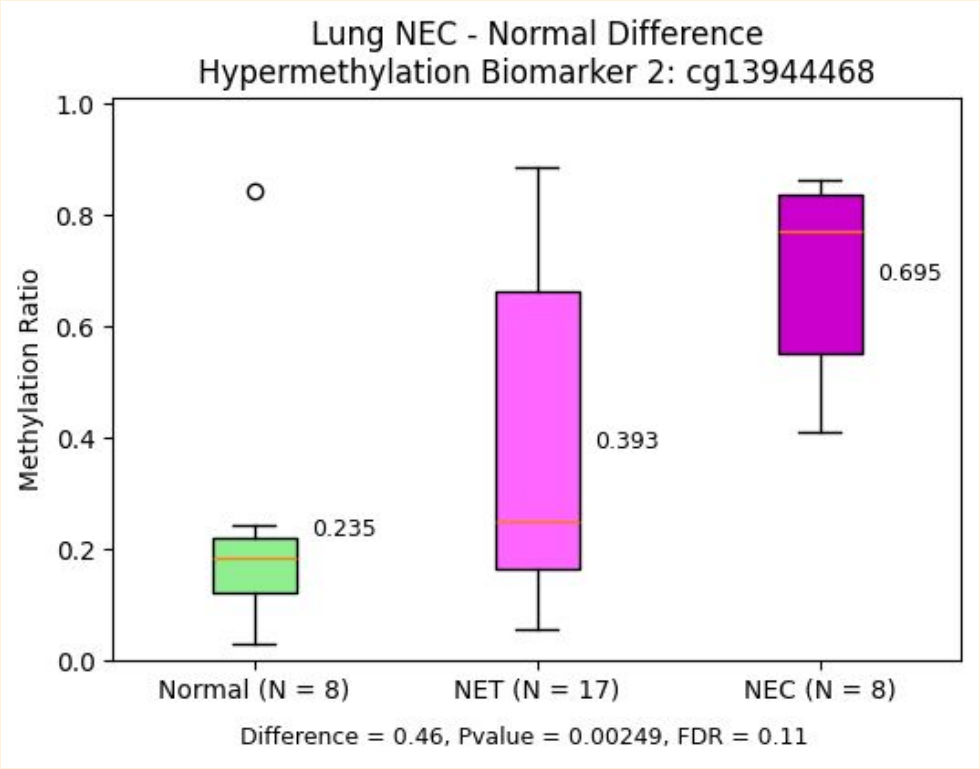


Hypermethylated Biomarker Candidates NEC - Normal

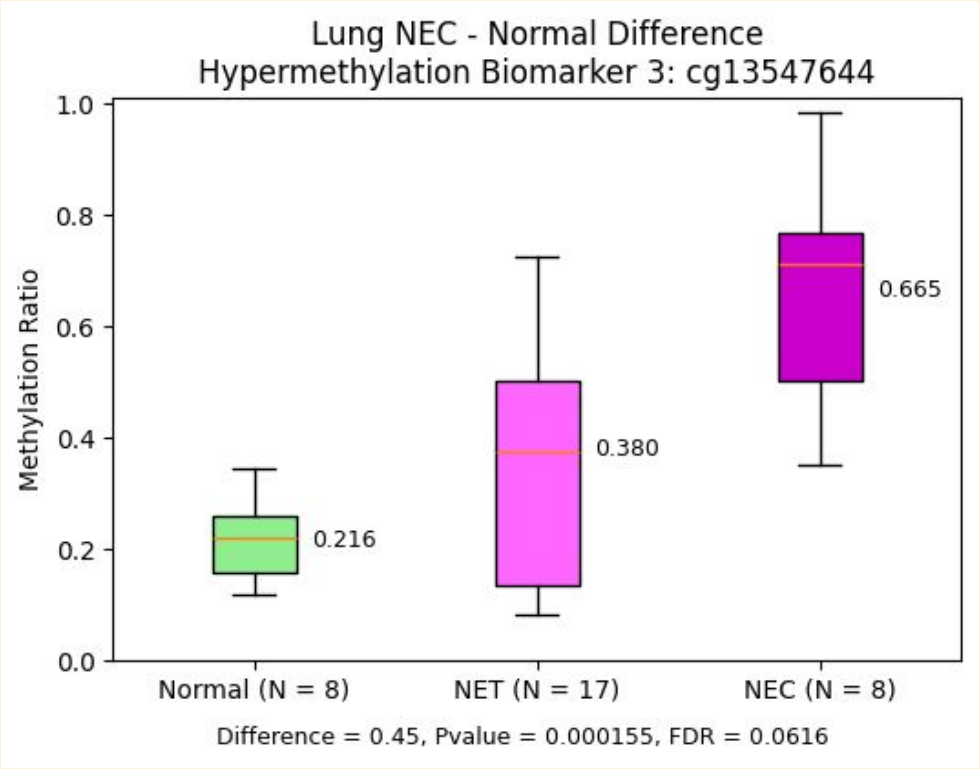
DNA Methylation Biomarker Candidate 1 NEC - Normal



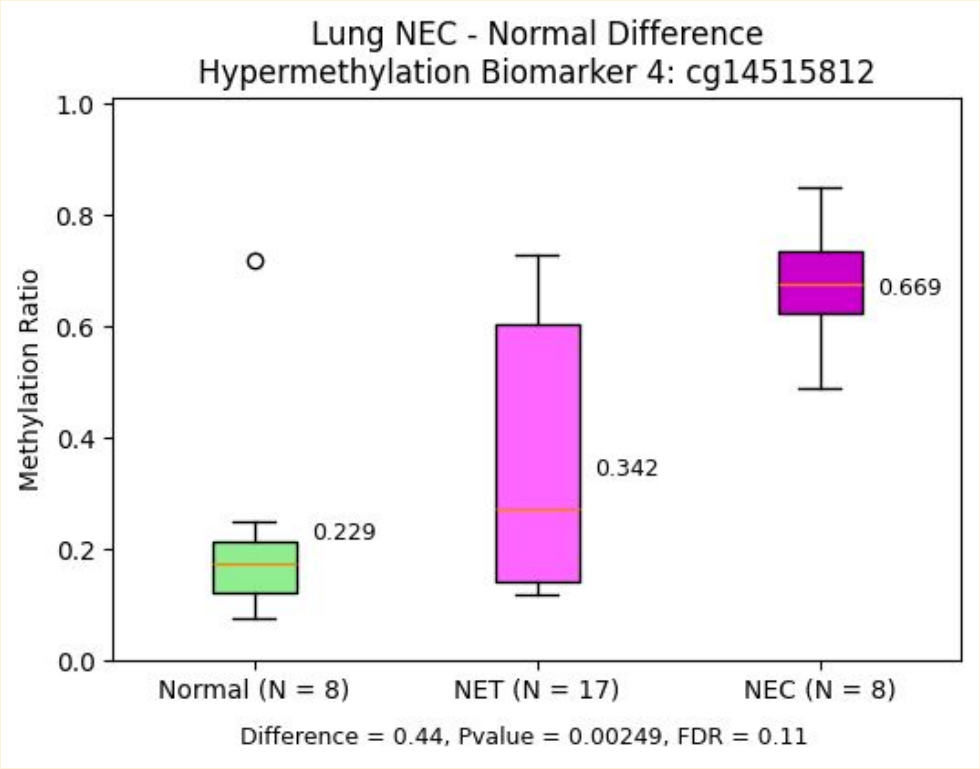
DNA Methylation Biomarker Candidate 2 NEC - Normal



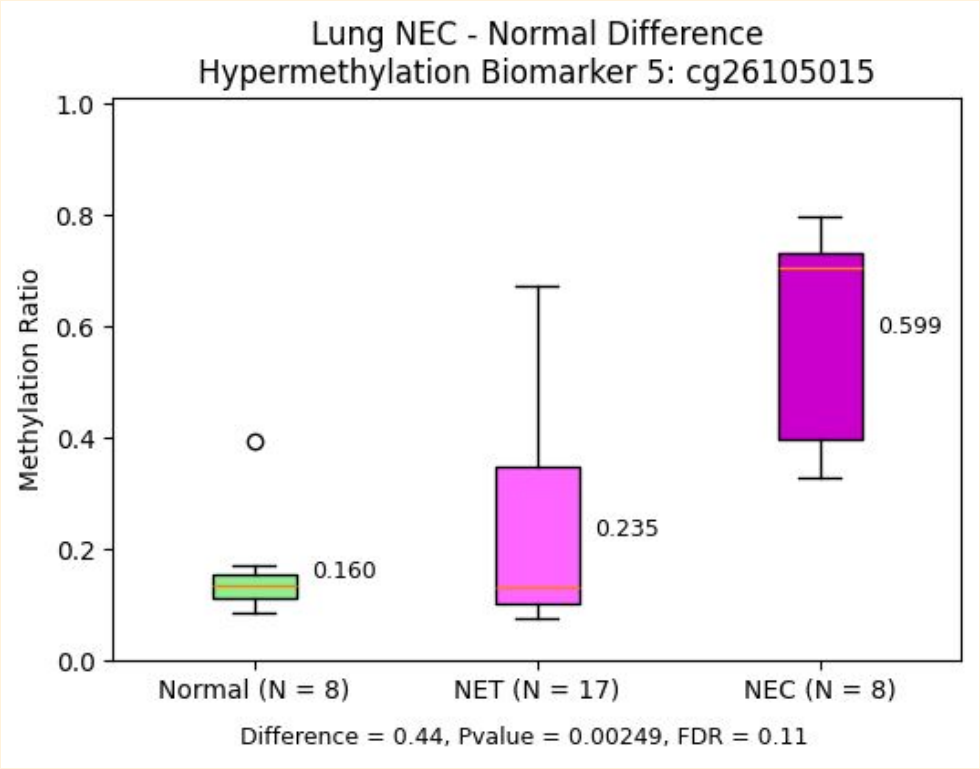
DNA Methylation Biomarker Candidate 3 NEC - Normal



DNA Methylation Biomarker Candidate 4 NEC - Normal

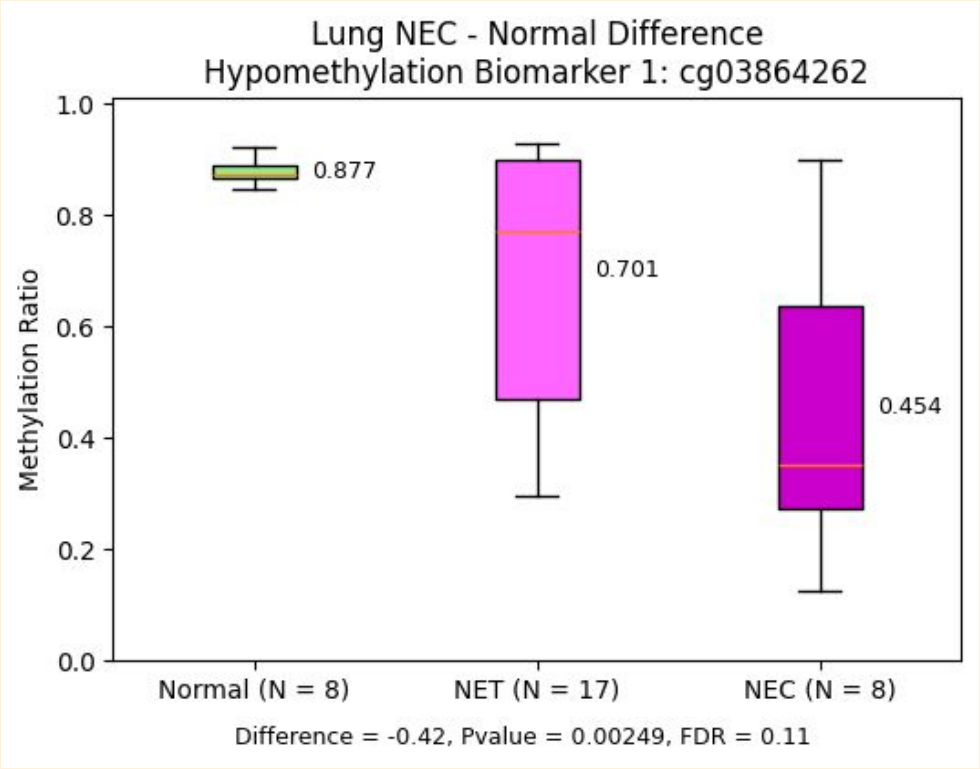


DNA Methylation Biomarker Candidate 5 NEC - Normal

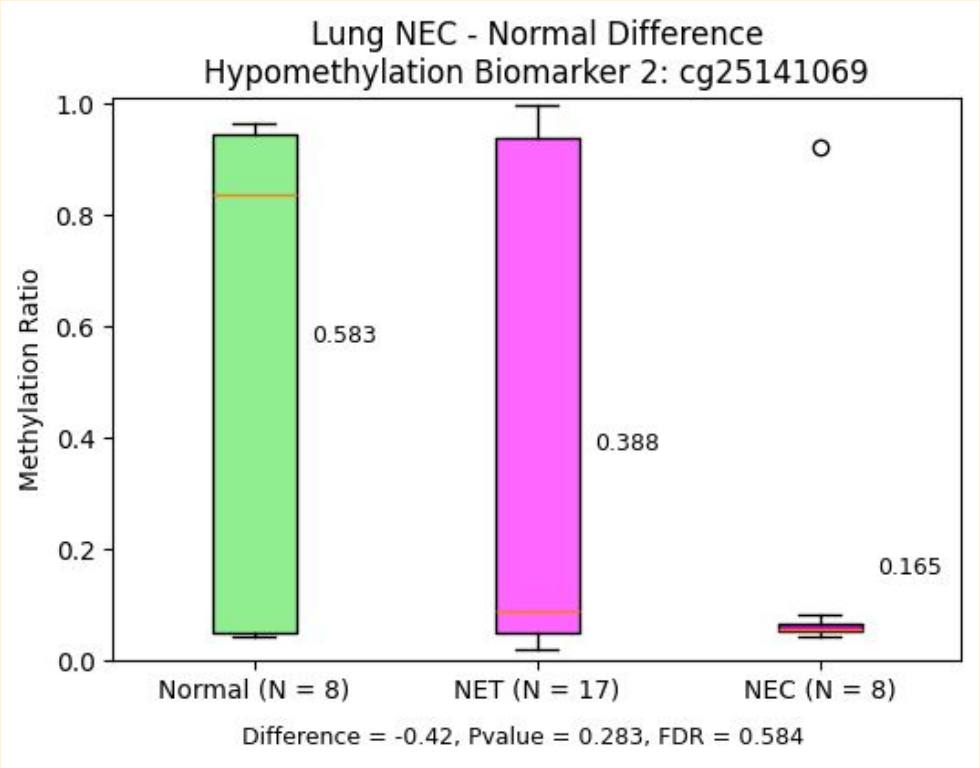


Hypomethylated Biomarker Candidates NEC - Normal

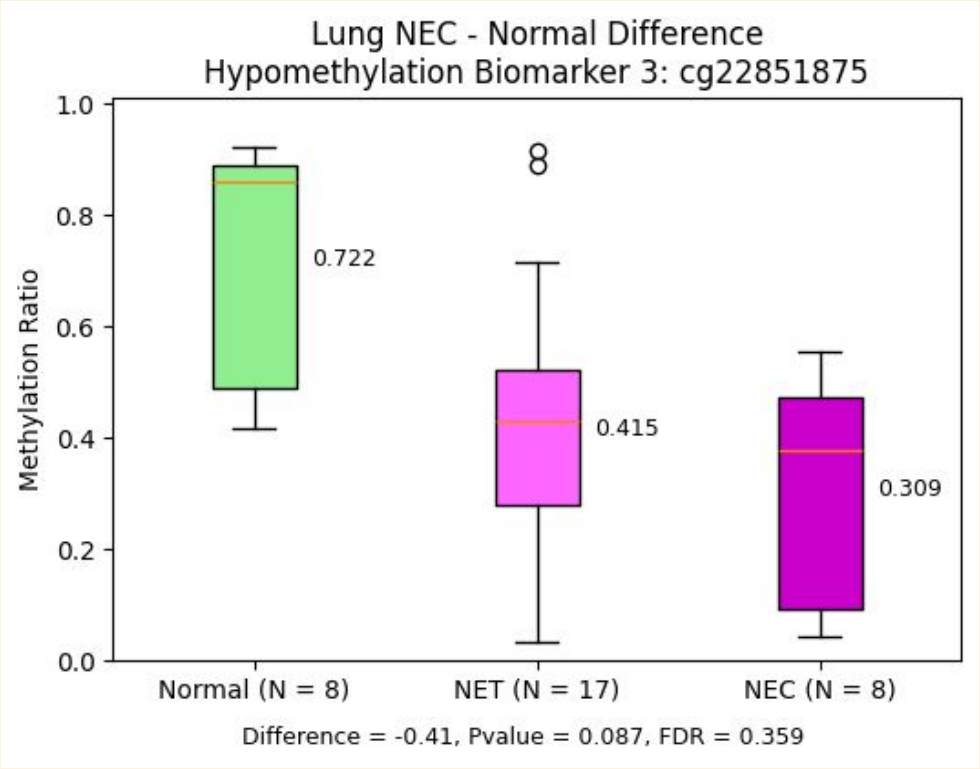
DNA Methylation Biomarker Candidate 1
NEC - Normal



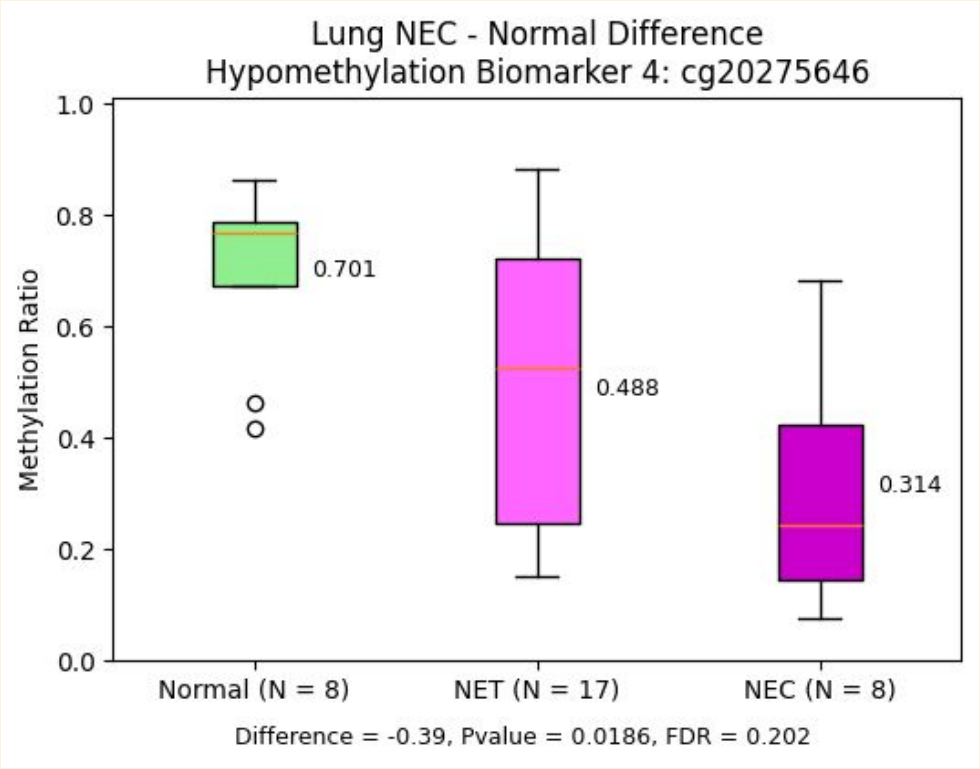
DNA Methylation Biomarker Candidate 2 NEC - Normal



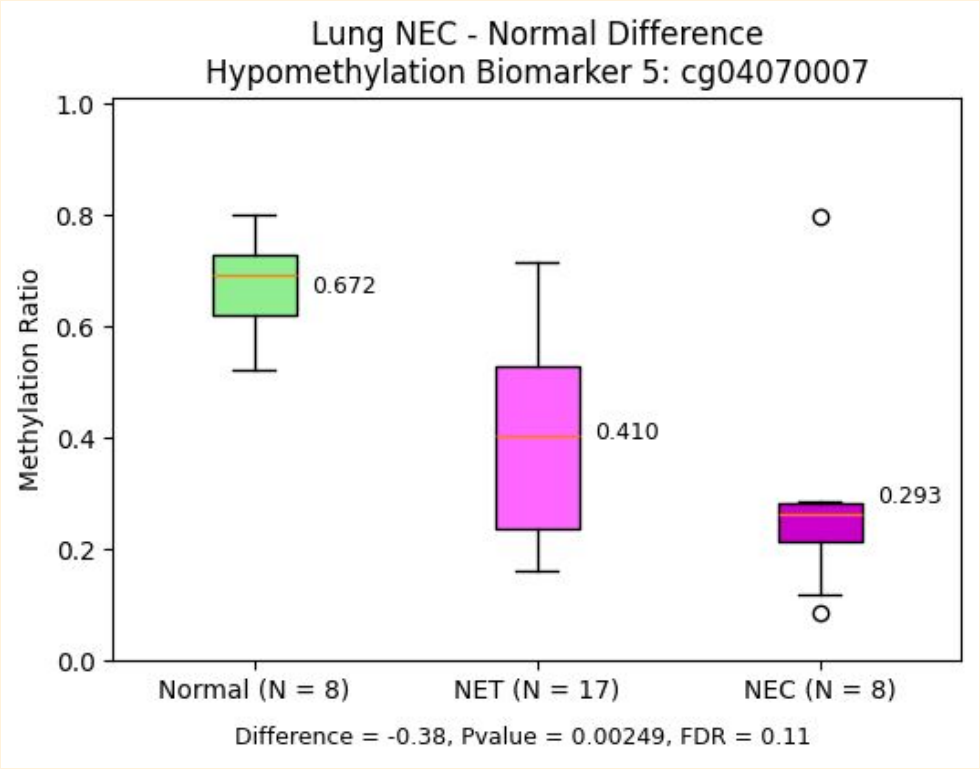
DNA Methylation Biomarker Candidate 3 NEC - Normal



DNA Methylation Biomarker Candidate 4 NEC - Normal



DNA Methylation Biomarker Candidate 5 NEC - Normal



Conclusions

- In summary, neuroendocrine neoplasms (NENs), present significant diagnostic challenges in its initial detection and subtype classification.
- Accurate classification between low grade well differentiated neuroendocrine tumors (NETs) and high grade poorly differentiated neuroendocrine carcinomas (NECs) is critical for patient treatment but often remains difficult using Histopathological methods. My research focuses on addressing this gap by applying DNA Methylation to lung NEN samples. As an epigenetic marker, DNA methylation reflects underlying tumor biology and holds potential as a reliable biomarker for tumor classification. By integrating methylation data with machine learning models, we aim to create a cost effective and clinically useful tool to improve identification of NENs and its subtypes.

Acknowledgments

Lee Lab,
Fox Chase Cancer Center

Timothy Kumphan
Hayan Lee

Temple University,
College of Science & Technology,
Research Scholars Program

Dominic Letarte

Q&A