**SUMMARY**

In the US, approximately 600,000 lung nodules are detected each year. While over 75% are benign, distinguishing benign vs. malignant nodules is difficult and over $52 billion is spent each year on additional testing. The goal of this study was to develop a blood-based test to distinguish benign vs. malignant lung nodules, thereby reducing both the risk associated with invasive procedures such as biopsies and surgery, and the cost.

Identification of cancer biomarkers in blood is a difficult task due to the wide dynamic range of blood proteins and the low concentration of tumor-derived proteins in the circulation. To maximize the opportunity to detect tumor-derived proteins, we used a two-step discovery approach.

The initial discovery study was conducted on malignant and matched normal tissue. Tissue samples were processed to enrich for secreted proteins and for epithelial and endothelial cell surface proteins that are likely to be released into the circulation, and the expression of these protein populations in malignant vs. normal tissue was compared.

MRM mass spectrometry was selected for validation and for the clinical assay due to its high specificity and sensitivity, multiplexing capacity, significant dynamic range, and rapid and reliable development and deployment. A highly-multiplexed MRM-based assay was developed to screen for 217 differentially expressed proteins from the discovery study plus an additional 171 candidate biomarkers from the literature. Of these 371 target proteins, a total of 190 (51%) were detected in plasma.

The MRM panel was used to analyze plasma samples from multiple clinical centers and an initial classifier was generated. This classifier was then further refined and verified in independent clinical samples. The final, validated MRM assay, which is now being offered as an LDT in a CLIA certified lab, includes 5 proteins which classify benign vs. malignant as well as 6 proteins used for normalization of protein concentration in plasma. This test is being used to rule out those likely to have a benign nodule, along with consideration of other risk factors, including age, nodule size and pack-years of smoking.

This study illustrates a robust proteomics approach for discovery and development of cancer biomarkers. By conducting the initial discovery study on secreted and membrane proteins from the tumor and then screening for these in the blood, we were able to identify biologically and clinically relevant markers to address an unmet need.

**BIOMARKER DISCOVERY IN LUNG TISSUE**

**Methods**

- To enhance the opportunity to discover biologically and clinically relevant biomarkers, discovery was first performed in lung tissue and then translated to plasma
- Enriched for tumor-associated proteins likely to be detected in plasma (Panel 1A), including:
  - Lung tumor-specific plasma membrane proteins from epithelial and endothelial cells
  - Lung tumor-specific secreted proteins
- Secreted and membrane fractions were then analyzed using a shotgun proteomics workflow (Panel 1B)

**Results**

- A total of over 3000 proteins were detected, including:
  - Over 1200 secreted proteins
  - Over 2000 membrane proteins (epithelial and endothelial)
- Good separation was observed between tumor and normal samples using Principal Component Analysis (PCA). An example for epithelial proteins is shown in Panel 1C.
- 217 proteins that were upregulated in the cancer tissue were selected for further evaluation in plasma. The origins of the upregulated proteins are shown in Panel 1D.

**BIOMARKER DISCOVERY IN PLASMA**

**Methods**

- Selected 217 upregulated proteins from tissue-based discovery
- Added another 171 targets from the literature
- Developed an MRM assay for 388 proteins (Panel 2A)
- Built a classification model to distinguish benign and malignant lung nodules using the logistic regression classification method. Classifiers were selected based on their clinical applicability on the classifier rather than on individual performance.

**Results**

- Successfully developed MRM assay for 371 proteins, with an average of 4 peptides per protein
- Detected 190 of the 371 proteins (51%) in plasma, with equal success for targets from the tissue-based discovery study and from the literature
- Developed a 13 protein logistic regression model to classify malignant vs benign nodules (Panel 2B)
- Tested the 13 protein model on plasma from an additional cohort of patients (52 malignant and 52 benign)
- Classifier score showed no association with patient age, nodule size, or smoking pack years (Panel 2C), suggesting the classifier provides independent and complementary information
- Pathway analysis (IPA) showed that all but one protein (SL9) could be linked to 4 transcription factors (FOS, AHR, NFE2L2, MYC) that are associated with lung cancer, lung inflammation, and oxidative stress (Panel 2D)

**CLASSIFIER REFINEMENT AND CLINICAL UTILITY**

- Panel 3A illustrates the process for selecting the diagnostic proteins in the final model from over 3200 proteins detected in the initial discovery study.
- The 13 classifier model was further refined to a final model which included 5 diagnostic proteins and 6 additional proteins used for normalization (Panel 3B).
- The classifier was validated in a retrospective study of 141 subjects (78 cancer and 63 benign).
- The classifier demonstrated 90% negative predictive value (NPV) and 26% positive predictive value (PPV) assuming a 23% prevalence of disease and a classifier cutoff of 0.36 (Panel 3C).

**CONCLUSIONS**

- This study demonstrates a successful translation of a panel of protein biomarkers from discovery to the clinic.
- Enrichment of secreted and membrane proteins from tissue provided tumor-associated markers that were then assessed in plasma.
- Over 50% of the proteins that were upregulated in malignant lung tissue were detected in plasma.
- The final assay includes 5 classifier proteins (with 6 normalization proteins) and demonstrated a NPV of 90%.
- The test is now in clinical use for identification of benign lung nodules in the 8 - 30 mm range.