UNIVERSITY of MARYLAND
SCHOOL OF MEDICINE

University of Maryland Baltimore
Biomarker Reference Laboratory

Non-coding RNAs (ncRNAs) as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

Lung and Upper Aerodigestive Cancer Collaborative Group

June 30, 2020

Sanford A. Stass, MD
Feng Jiang, MD, PhD
Søren M. Bentzen, MSc, PhD, DMSc

University of Maryland Baltimore
Biomarker Reference Laboratory (UMB-BRL)

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

- NCI-NIST found a 20% reduction in lung cancer mortality for screening with low-dose CT (LDCT), compared with chest radiography, and is used for early detection of lung cancer.
- LDCT increases the number of indeterminate pulmonary nodules (IPNs) detected in asymptomatic individuals, whereas only a small fraction of IPNs may be lung cancer.
- Sputum is noninvasive and accessible.
- Molecular studies of sputum may provide a noninvasive approach for a preoperative diagnosis of malignant IPNs.

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

- Targeted NGS
  - High throughput
  - Wide detectable expression ranges
  - Analyzing expression of all annotated miRNAs
  - Detection of novel miRNAs

- Previously NGS analyzed lung tumor tissue and paired non-cancerous lung tissue using Illumina Genome Analyzer II, a high performance NGS platform.
  - 102 miRNAs displayed dysregulation with fold change ≥2.0 in NSCLC tissue vs. normal lung tissue.

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

Aim 1: Test and validate 102 miRNAs defined from whole-genome NGS in NSCLC and noncancerous tissues to determine if changes of the genes can be validated using targeted NGS in a CLIA/CAP laboratory.

TABLE 1: 102 mRNAs have a ≥2.0-fold change in NSCLC vs. normal lung tissues

miRNAs	Fold-Chng	miRNAs	Fold-Chng	miRNAs	Fold-Chng	miRNAs	Fold-Chng	miRNAs	Fold-Chng	miRNAs	Fold-Chng
miR-375	-7.55	miR-22	2.97	miR-34a	2.54	miR-125b	-2.65	miR-1254	3.33	miR-668	4.22
miR-135a-5p	-6.29	mR-451a	3.07	miR-412	2.56	miR-451	-2.52	miR-25	3.40	miR-147b	4.28
miR-184	-6.25	miR-192-5p	3.09	miR-149-5p	2.61	Let-7e	-2.48	miR-577	3.55	miR-92	4.32
miR-486-5p	-5.92	miR-410	3.13	miR-18a	2.64	miR-101	-2.46	miR-93	3.471	miR-210	4.39
miR-30a-3p	-5.88	miR-521	3.13	miR-1910	2.66	miR-99a	-2.45	miR-20a	3.47	miR-323a-3p	4.46
miR-138-1-3p	-4.64	miR-380-3p	3.17	miR-28	2.67	miR-196	-2.45	miR-196a-5p	3.53	miR-431-5p	5.08
miR-30c-2-3p	-4.30	miR-487a	3.15	miR-873-5p	2.69	Let-7a	-2.43	miR-128	3.55	miR-1291	5.82
miR-29b-1-5p	-4.10	miR-214	3.19	miR-490-3p	2.73	miR-17-5p	-2.33	miR-4435	3.57	miR-548a-3p	6.07
miR-223-5p	-3.87	miR-1197	3.14	miR-156	2.73	miR-126	-2.31	miR-520	3.57	miR-758-3p	2.10
miR-3614-5p	-3.76	miR-3609	3.23	miR-3676-5p	2.85	miR-100	-2.25	miR-200a	3.57	miR-409-5p	2.49
miR-449c-5p	-3.73	miR-21-5p	3.24	miR-31-5p	2.94	miR-26b	-2.21	miR-574-5p	3.57	miR-376b-3p	4.02
miR-145	-3.37	miR-146	3.35	miR-708-5p	2.95	miR-652	-2.11	miR-222	3.62	miR-141	4.03
miR-133	-2.23	miR-200b	3.25	miR-9-5p	4.07	miR-203	-2.10	miR-146b	3.69	miR-152	-2.11
miR-181a	-2.78	miR-335-3p	3.26	miR-944	4.10	miR-886-3p	-2.16	miR-221	3.69	miR-301b	2.07
miR-212	-2.72	miR-24	3.35	miR-1269b	4.16	miR-140-3p	-2.09	miR-106a	3.79	miR-155	3.99
miR-143	-2.66	miR-200C	3.35	miR-660	4.17	miR-224-5P	-2.16P	miR-320	-2.44	miR-504	2.68
miR-10b	-2.51	miR-191	4.35	miR-183-3p	4.13	miR-197	-2.89	miR142c	2.35	miR-1185-1	3.25

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

All data for miRNA targets were preprocessed by the bioinformatics group at the Institute for Genomic Sciences (IGS) and analyzed by Dr. Søren Bentzen, Director of the UMGCCC Biostatistics Shared Service at University of Maryland.

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

Patient Samples

- UMB-BRL has 553 lung tumor tissues and matched non-cancerous lung tissue. The 553 samples are represented by 236 (43%) adenocarcinoma, 231 (42%) squamous cell carcinoma, 86 (16%) other types of lung cancer, including large cell carcinoma.

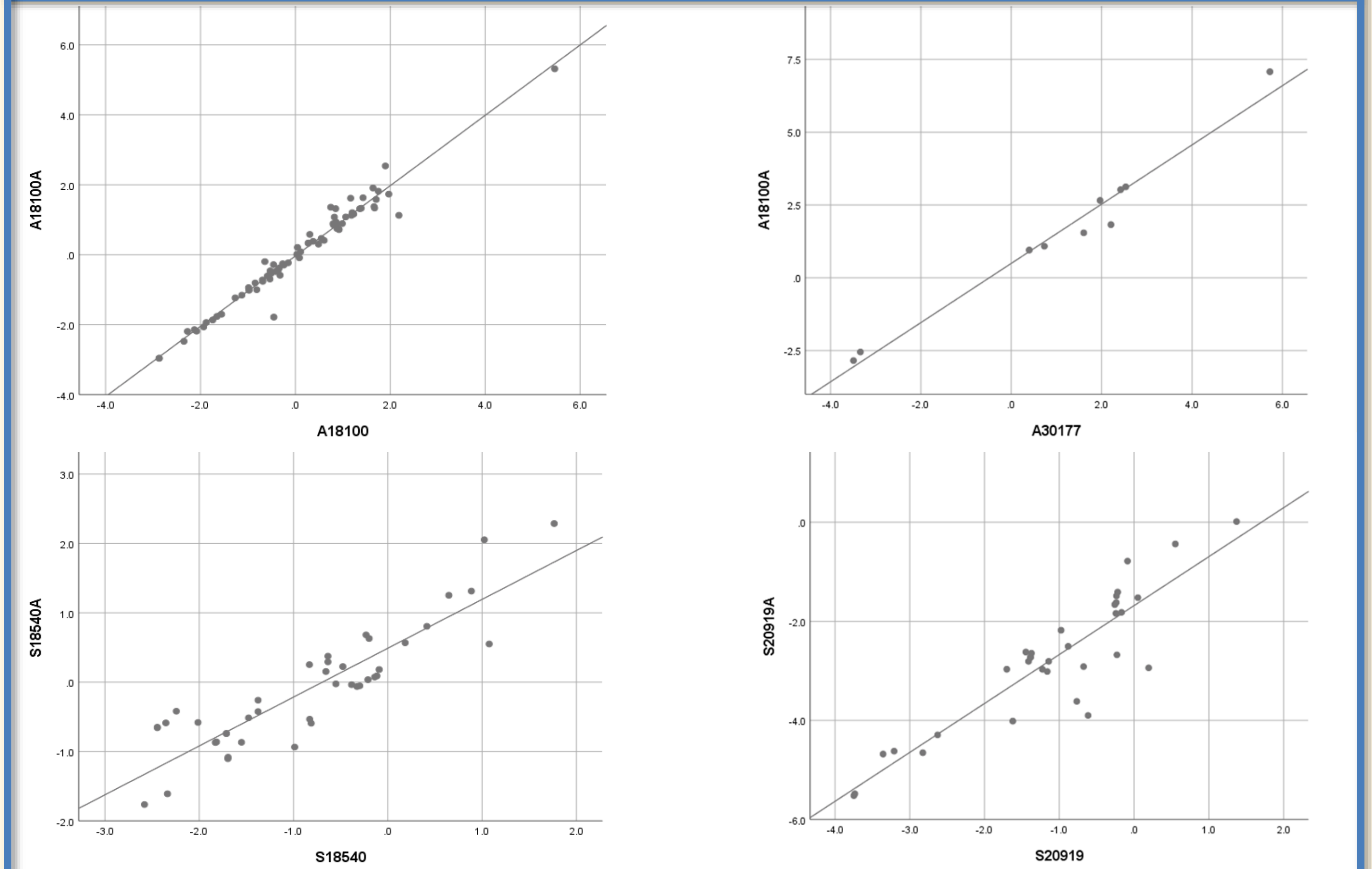
- Collaboration with Dr. Pierre Massion, a MTA is in place and is being extended to provide frozen lung tumor tissues and matched noncancerous lung tissues of stage 1 NSCLC patients, along with matched sputum samples, for testing.

- Benchmarks for success
  - Previous miRNA study (miR-21, miR-31 and miR-210) demonstrated 82% sensitivity and greater than 87% specificity for malignant pulmonary nodules in sputum.
  - The goal of the new panel of biomarkers is to have a 90% sensitivity and 95% specificity for malignant pulmonary nodules vs. 27 – 41% sensitivity for cytology.

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

TABLE 1

Examples of Reproducibility



miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

Sample Reproducibility

Sample Name	Including all targets with > 0 reads	Including targets with greater than or equal to 10 reads	# of targets with > 0 reads	# of targets with greater than or equal to 10 reads
A15081	0.929	0.986	93	72
A18100	0.950	0.993	88	68
A20336	0.932	0.962	92	74
A20413	0.944	0.986	82	61
A30177	0.788	0.986	47	10
A30213	0.815	0.888	78	50
S17578	0.959	0.980	84	64
S18339	0.893	0.945	74	35
S18540	0.624	0.891	73	42
S18663	0.795	0.898	77	47
S18736	0.841	0.784	76	47
S20496	0.876	0.990	89	64
S20919	0.885	0.901	81	33
S30216	0.716	0.778	91	77
S30220	0.945	0.987	87	65
Average	**0.859**	**0.930**	**80.8**	**53.9**

15 tumor/normal pairs run in duplicate to determine inter-run and intra-run reproducibility as quantified by the R^2 from the log-fold change. Variation in quality of tissue sample and RNA quality and quantity may account for variation in the reproducibility of the assay across patients.

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

- Variation in quality of tissue sample may account for variation in the reproducibility of assay across patients.

- In addition, Potential variation in reproducibility may be partially explained by: RNA quantity and/or quality from the extraction. We are establishing the lowest amount of RNA that can be used and additional replicate samples are currently being run.

- Analysis of NSCLC versus normal lung tissue pairs on IoT/PGM platform suggests miRNAs of potential interest with similar dysregulation as seen previously using whole-genome NGS data.

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

- We have demonstrated that our optimized test process can accurately identify elevated and reduced levels of targeted miRNAs.
- The number of total reads obtained per specimen ranged from 1,703,691 to 3,785,283 (average = 2,532,933).
- Our preliminary validation data suggests that the miRNA signature is reproducibly detected by our targeted miRNA-enriched NGS test.
- Data is promising for the identification of a targeted panel of miRNAs for the early detection of NSCLC.
- We have run over 80 tumor/normal pairs and have confirmed the reproducibility.

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

- Statistical analysis of miRNA expression was performed in an independent series of 29 NSCLC versus normal lung tissue pairs from the UMB-BRL and from Vanderbilt, to include sub-variants of the miRNAs. Testing showed reproducibility. UMB-BRL originally tested and validated 102 miRNAs in non-small cell lung cancer (NSCLC) and the matched non-cancerous tissue by using whole-genome next generation sequencing.

- The original set of 102 miRNAs were compared using NGS in an expanded cohort of 57 matched NSCLC and noncancerous tissue samples, identified miRNAs showing a statistically significant fold change (FC) in tumor relative to normal tissue

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

- Of the original set of 102 miRNAs that were originally identified using NGS in matched NSCLC and noncancerous tissues, 35 continue to be associated with a fold change (FC) in tumor compared to the paired normal sample that was statistically significant at the 5% level.

- These miRNAs now represent the final candidate biomarker panel for detection of NSCLC in sputum.

- These miRNAs that have been studied experimentally have both biologic and therapeutic implications.

- We will now initiated testing of sputum samples with matched tumor/normal samples from the same individual received from Vanderbilt and NSCLC sputum samples from the UMB-BRL

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

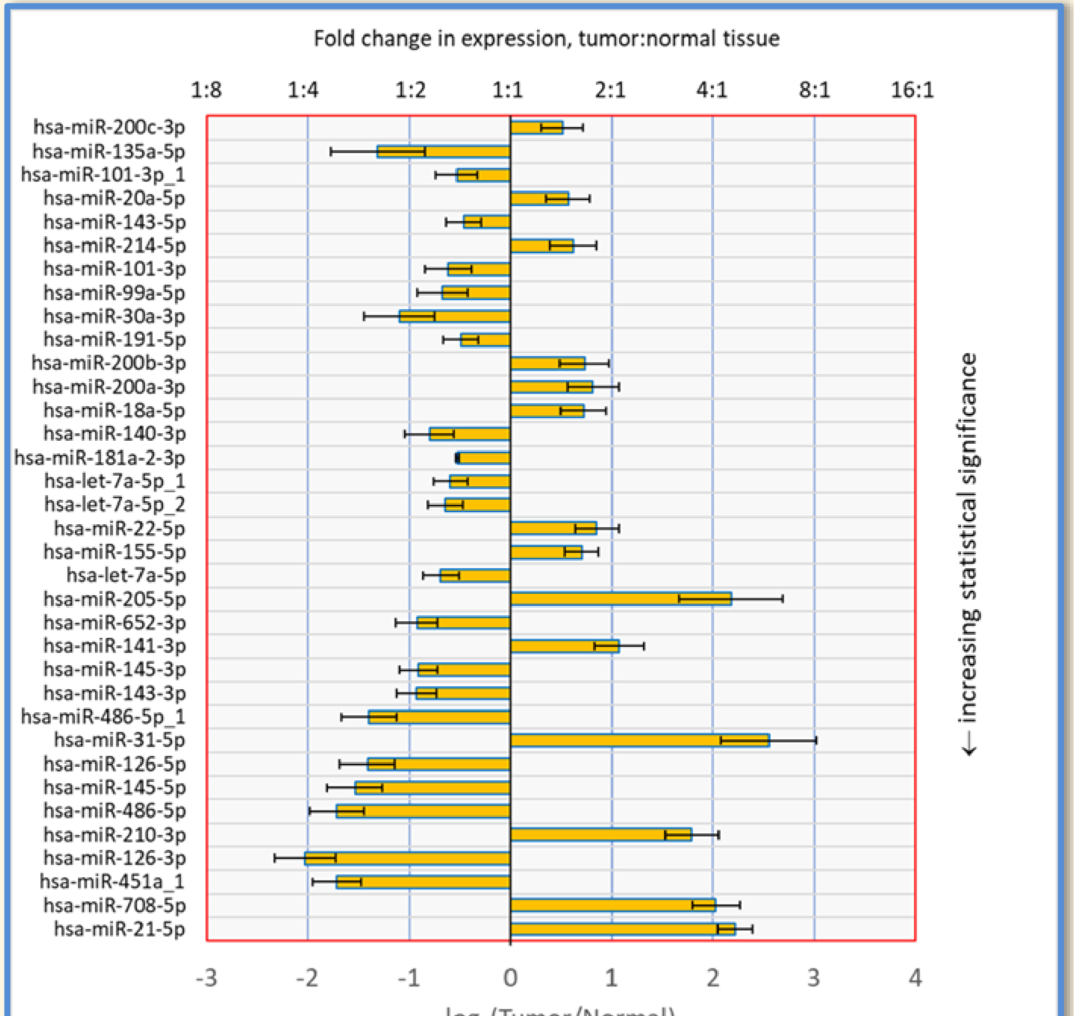


Table 1: Fold change in expression of the top 35 miRNA showing differential expression in tumor relative to normal-tissue samples paired within cases.

miRNAs as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

Future Aims

- **Aim 2:** Use the presumptive panel of the 35 miRNAs biomarkers to identify malignant pulmonary nodules, using a testing set of sputum samples to determine the diagnostic value of the biomarker panel.

miRNA as Biomarkers for Early Detection of Non-Small Cell Lung Cancer in Sputum

- **Aim 3:** Validate analytical and clinical performance of the biomarker panel in prospectively collected sputum samples with and without the Lung Flute from individuals with CT-detected pulmonary nodules .

Next steps

- The derived sputum miRNA signature will be tested in tumor/normal tissue/sputum triplets to compare with the performance of our tissue miRNA signature.

- Over- and under-expression of these candidate miRNAs relative to the population reference values will be used as a signature of NSCLC.

- We are prospectively validating the biomarker panel with study of sputum samples from individuals with lesions detected by low dose CT. We will conduct a nested case-control study to establish the sensitivity and specificity of the sputum miRNA assay using clinical imaging and pathologic diagnosis.

Long non-coding RNAs (lncRNAs) as Biomarkers for Early Detection of Non-Small cell Lung Cancer in Sputum

Feng Jiang, MD, PhD

Long non-coding RNAs (lncRNAs) as Biomarkers for Early Detection of Non-Small cell Lung Cancer in Sputum

- lncRNAs play vital roles in various biological processes, and their dysregulations contribute to the development and progression of lung tumorigenesis. We have identified 26 lung cancer–associated lncRNAs by using the same strategy with whole-genome NGS as the one we used for miRNA analysis.

Expression levels of lncRNAs in lung tumor vs. noncancerous lung tissues			
lncRNAs			
SNHG1	Up	RMRP	Up
SNHG9	Up	SOX2-OT	Up
SNHG11	Up	HNF1A-AS1	Up
LINC00461	Up	CCAT2	Up
MEG8	Up	LUADT1	Up
CAR10	Up	ZXF1	Up
MALAT1	Up	ANRIL	Up
HOTAIR	Up	SCAL1	Up
H19	Up	GAS5	Down
RGMBAS1	Up	PANDAR	Down
PVT1	Up	MEG3	Down
BCYRN1	Up	TUG1	Down
DLX6-AS1	Up	SPRY4-IT1	Down

Long non-coding RNAs (lncRNAs) as Biomarkers for Early Detection of Non-Small cell Lung Cancer in Sputum

To determine whether aberrant lncRNAs could be used as potential sputum biomarkers for lung cancer, we measured expressions of six lung cancer-associated lncRNAs in sputum of a cohort of 67 lung cancer patients and 65 cancer-free smokers with benign diseases by using ddPCR.

Long non-coding RNAs (lncRNAs) as Biomarkers for Early Detection of Non-Small cell Lung Cancer in Sputum

Expression levels of four sputum lncRNAs in lung cancer patients and controls			
lncRNAs	Mean of level in NSCLC patients	Mean of level in controls	AUC (95% CI)
SNHG1	6.65 (1.08)	5.29 (1.13)	0.80 (0.73 to 0.87)
H19	3.39 (2.63)	2.25 (0.65)	0.77 (0.68 to 0.83)
HOTAIR	1.08 (0.56)	0.29 (0.09)	0.73 (0.66 to 0.83)
MEG3	1.53 (0.92)	3.02 (1.56)	0.72 (0.63 to 0.78)
Abbreviations: AUC, the area under receiver operating characteristic curve; CI, confidence interval.			

- As shown in the table, four (SNHG1, H19, and 190 HOTAIR, and MEG3) of the six lncRNAs displayed a significantly different level in sputum of the lung cancer patients vs. cancer-free individuals, producing 0.72–0.80 AUCs (all P < 0.01).

Long non-coding RNAs (lncRNAs) as Biomarkers for Early Detection of Non-Small cell Lung Cancer in Sputum

Using logistic regression models with constrained parameters, from the four lncRNAs, we further selected three genes (SNHG1, H19, and HOTAIR) as the best biomarker panel, which produced 82.03% sensitivity and 89.23% specificity for diagnosis of lung cancer.

Summary

- **What is the intended use?**

  Screening smokers for lung cancer early detection and clinical decision of biopsy.

- **What is the type marker or biomarker panel?**

  miRNA and lncRNA.

  Since dysregulations of miRNAs and lncRNAs contribute to lung tumorigenesis via diverse pathways, combined analysis of the two different biomarkers might have a higher diagnostic performance than each type of biomarkers used alone.

  In the future we will integrate the miRNA ad lncRNA biomarkers to develop a panel of biomarkers for NSCLC early detection.

- **What is the types(s) of specimens (tissue, blood and/or urine)?**

  Sputum

- **How does the new biomarker or the new panel of biomarkers add to the current knowledge baseline?**

  Improve sensitivity of screening when combined with LDCT.

- **What are the parameters of sensitivity, specificity, PPV, NPV, AUC and how does this compare to the current standard of care?**

  Given that only three lncRNA biomarkers have 82.03% sensitivity and 89.23% specificity and integrating two different biomarkers panels may have synergistic effect. Our goal is that combined analysis of miRNA and lncRNA biomarkers might have 90% sensitivity and 95% specificity compared to LDCT 64% specificity and cytology sensitivity of 27 - 41% for a single sputum sample. Cytology is greater than 99% specificity.

UMB-BRL

Vimentin (VIM) Gene Methylation Next Generation Sequencing (NGS)

UMB-BRL
VIM Gene Methylation NGS Team

UMB-BRL

- Debra Kukuruga, MT(ASCP), PhD, D(ABHI)
  Significant Contributor

Institute for Genome Sciences

- Luke Tallon, PhD, Scientific Director
- Lisa Sadzewicz, PhD, Administrative Director
  Genomics Resource Center (GRC)
- Michelle Giglio, PhD, Associate Director
  Analysis Informatics Resource Center

Laboratory Operations

- Jeff Gray MT(ASCP), Lead Research Specialist
- Juliana Woodcock, Laboratory Manager
- Sarah Flagg MT(ASCP), CHS(ABHI), Quality Assurance Specialist
- Sandra Ott, Program Manager, GRC
- Madalyn Woodruff MT(ASCP), Technologist

NGS Bioinformatics

- Sandra Ott, Program Manager, GRC
- Ivette Santana-Cruz, Senior Bioinformatics Software Engineer
- Amol Shetty, PhD, Lead Bioinformatics Software Engineer

UMB-BRL
VIM Gene Methylation
Next Generation Sequencing (NGS)

UMB BRL: Sanford Stass, MD, Debra KuKuruga, PhD, Case Western Reserve University (CWRU): Sandy Markowitz, MD, PhD, University of Michigan: Dean Brenner, MD, PI Great Lakes New England 010 Stool Study (GLNE-010)

Background

- Aberrant methylation of exon-1 sequences within the non-transcribed vimentin gene is a novel molecular biomarker of adenomas and colon cancer and can be successfully detected in fecal DNA.

- UMB-BRL serves as a CLIA/CAP accredited reference laboratory for vimentin assay(s) developed by CWRU.

- Clinical study samples provided by the Great Lakes New England (CVC), Dean Brenner, MD, University of Michigan and SPORE specimens (for pilot assay) provided by Sandy Markowitz, MD, PhD, CWRU.

UMB-BRL
VIM Gene Methylation
Next Generation Sequencing (NGS)

- Validated a methylation specific quantitative real-time PCR assay

- NGS identifies more advanced adenomas than qPCR
  - Appears qPCR does not correlate well with NGS
  - Vimentin NGS + fecal immunochemical test together more sensitive than Cologuard for advanced adenoma and more specific than Cologard for non-adenoma controls
- UMB-BRL purchased a Illumina MiSeq to implement an assay for VIM Methylation detection using NGS.

- Expert team of medical technologists

- Bioinformaticists perform analysis

UMB-BRL
VIM Gene Methylation NGS Bioinformatics

NGS Genomics Resource Center (GRC) Pipeline Development

- Pipeline development in collaboration with Dr. Helen R. Moinova, Dr. Thomas LaFromboise, Dr. Sanford Markowitz Case Western Reserve University (CWRU).
- Pipeline tested using 2 test samples from CWRU
  - Results were compared between analysis done at IGS GRC and CWRU
- 3 months for development and testing of the pipeline
  - Oct 2018 to Jan 2019
- Further validation was done using positive and negative control samples

UMB-BRL
VIM Gene Methylation
NGS Bioinformatics

Analytical Pipeline

Demultiplexing & Trimming

- Demultiplexing of samples based on sequencing barcodes
- Trimming of reads for primer, adapter and barcode sequences

Alignment

- Alignment of reads to Human (GRCh38 build) reference genome using Bismark
- Determine the methylation status for 21 CpGs across each individual read

Methylation Profiling

- Generate methylation profiles and classify reads into “methylated” or “unmethylated” reads
- Classify samples into “positive” or “negative” based on percentage of methylated reads

UMB-BRL
VIM Gene Methylation
NGS Bioinformatics

Bioinformatic Analysis

- VIM Methylation Profiling
  - Each read strand selected for the analysis must include all 21 CpG sites between the primers within the 199 bp VIM gene amplicon (otherwise read is excluded)
  - Sample should have ≥ 10,000 reads

- 629 samples analyzed from July 2019 to September 2019
  - 22 batches of 28 samples in triplicate and 6 controls in duplicate
  - 1 batch of 13 samples in triplicate and 6 controls in duplicate
  - Control samples were correctly identified to be negative or positive for VIM methylation with a 98% concordance between duplicates
  - 90% concordance between triplicates

UMB-BRL
VIM Gene Methylation
NGS Next Steps

Analyzed NGS sequencing data submitted to DMCC for their evaluation and correlation with advanced adenomas or cancer

Additional steps of study pending this evaluation

Lung tumor-associated Microbiome as a Biomarker of Lung Cancer

Natalia von Muhlinen, PhD
Visiting Assistant Professor, UMSOM

Objectives

- Investigate the contribution of the **tumor-associated microbiome** to lung cancer.
- Investigate the use of tumor-associated **microbes** enriched in lung cancer as novel **biomarkers** of lung cancer to improve diagnosis and treatment.

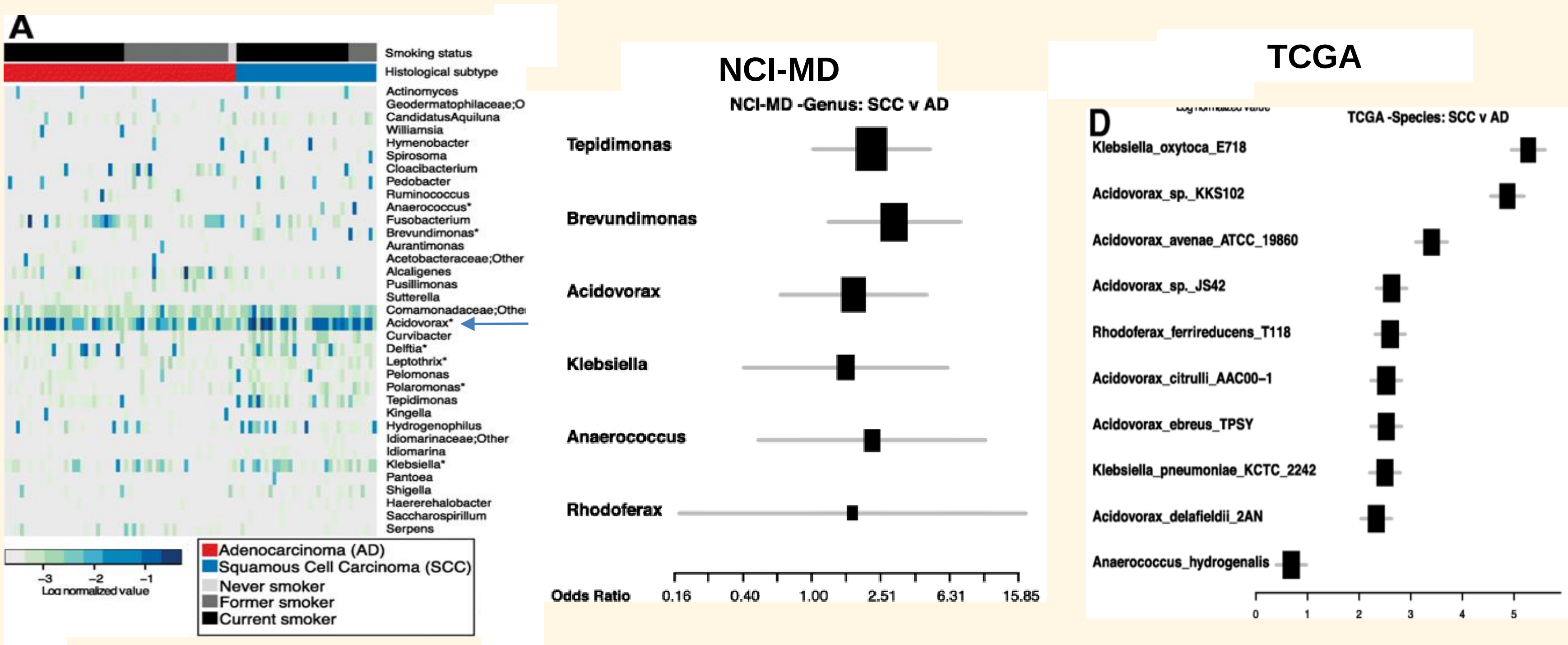
Study Design

- **Patients cohort:** frozen lung tissue from our NCI-MD Case Control Study (n = 143 tumor and n = 144 non-tumor adjacent tissues).
  - **Method of microbiome analysis:** 16S rRNA sequencing the V3–V5 16S ribosomal RNA (rRNA) bacterial gene using the Illumina MiSeq platform.
- **Validation cohort:** The Cancer Genome Atlas (TCGA) ; n = 1112 tumor and non-tumor adjacent RNA-sequencing [RNA-seq] data from tissues).
  - **Method of microbiome analysis:** After removing human reads, we used MetaPhlAn, Kraken, and PathoScope, to assign reads to taxonomy

Results

- Alfa diversity and beta diversity (changes between microbial communities) were significantly different between tumor and non-tumors, and between adenocarcinoma (AD) and squamous cell carcinoma (SCC) in the NCI-MD study and TCGA study.

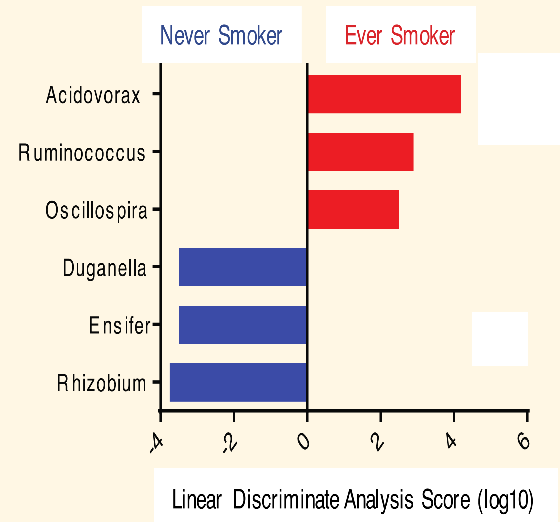
- In the NCI-MD cohort, we identified 32 **genera that were differentially abundant in SCC** (n = 47) **versus AD** (n =67) tumors (Student's t-test; MW P < 0.05), nine of which were significant after multiple testing correction (FDR) (*Acidovorax, Brevundimonas, Comamonas, Tepidimonas, Rhodoferax, Klebsiella, Leptothrix, Polaromonas, Anaerococcus*).



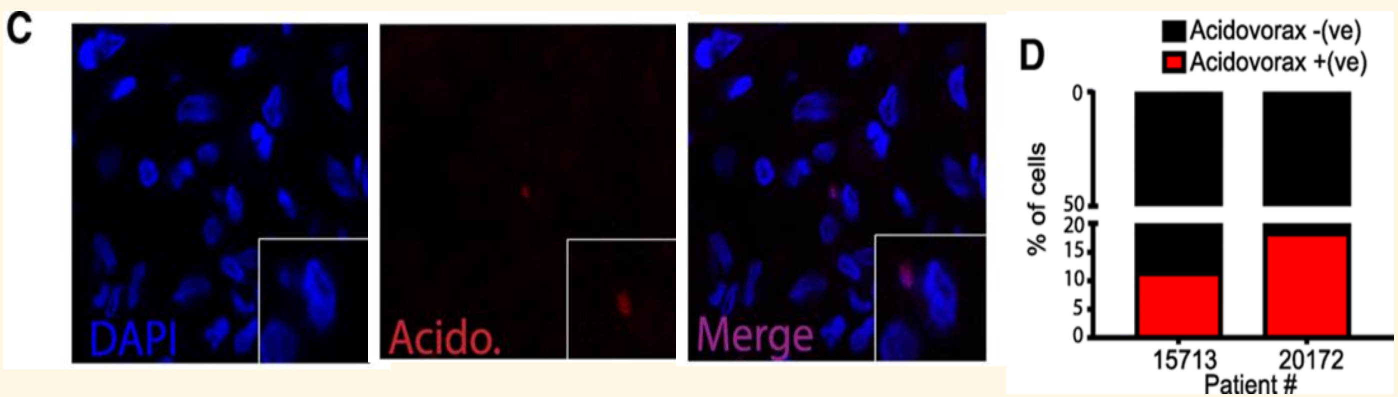
Greathouse K, …, **von Muhlinen N**,…, Harris CC, *Gen Biol* 2018

Results

- We analyzed whether the SCC-enriched taxa were associated with smoking.
- We stratified the tumor samples into never smokers (n = 7) or ever-smokers (current [n = 70] and former smokers [n =40]) using linear discriminant analysis (LEfSe) to identify smoking-associated microbial biomarkers in SCC tumors.
- We found that **Acidovorax** spp. was **more abundant** in **smokers and former smokers** versus never smokers.
- When stratified by TP53 mutations, we found tha Acidovorax spp. is enriched in tumors with mutations in TP53.



- We also **identified Acidovorax spp. in lung tumor** tissue using Acidovorax-specific probes by RNA fluorescent in-situ hybridization (**FISH**).



Greathouse K, …, **von Muhlinen N**,…, Harris CC, *Gen Biol* 2018

Conclusions

- Our results demonstrated **changes in the tumor-associated microbiome** in lung cancer.
- We showed that a microbial consortia is **associated with histological subtypes** of lung cancer and **TP53 mutations**.
- We confirmed the presence of a genus of bacteria, **Acidovorax spp.**, associated with **smoking** and **TP53 mutations**.
- These data indicate that changes in the tumor microbiome, together with tobacco smoke and mutations in TP53, **may drive lung cancer** and thus **may serve as biomarkers** of lung cancer. Further studies are necessary to address these questions.

Ongoing Studies

Functional Role of Acidovorax spp. in lung cancer

- Using conditional Kras/TP53 mutant lung cancer mouse model for **animal studies** to investigate whether Acidovorax spp. has **functional roles** in lung cancer progression.

Acidovorax and other tumor-associated microbes as biomarkers of lung cancer

- Developing **RNAScope** techniques to detect microbes in frozen tissue, paraffin-embedded tissues (FFPE), fixed frozen tissue, or tissue microarrays.
- Analyzing abundance of Acidovorax spp. and other tumor-associated microbes by RNAScope in **bronchial preneoplastic lesions and/or normal bronchial tissue**, available as FFPE sections.
- Planning to develop methods to detect these microbes in **nasal airway and sputum** for potential use as **early detection biomarkers** of lung cancer.

Collaborations:

Curtis Harris, MD – National Cancer Institute
Dean Mann, MD – University of Maryland School of Medicine, UMSOM
Kristen Leigh Greathouse, PhD – Baylor University