Identifying Clinically-Relevant Population-Specific IsomiRs in African Americans and European Americans with Lung Cancer

Thesis highlights

- Lung cancer is the leading cause of cancer-related deaths in the United States (U.S.) with over 139,000 deaths per year. African Americans (AAs) have a higher mortality and lower 5-year survival rate than European Americans (EAs).
- AAs smoke less than EAs. When they do smoke, AAs choose menthol cigarettes more than EAs (88.5% vs. 25.7%).
- Smoking-associated biomarkers have been explored at the genomic, metabolomic, and transcriptomic level, providing evidence for miRNA variants (isomiRs) as potential regulators of menthol metabolizing enzymes (MMEs).
- The miRNA and candidate isomiR, miR-374b-5p|3'a-1, had significantly higher expression in AAs that correlated with lower *CYP1B1* and *UGT2B4* expression.
- The miRNA and candidate isomiR, miR-374b-5p|3'a-1, had significantly higher expression in AAs upon menthol exposure. This did not correlate with lower *CYP1B1* and *UGT2B4* expression, suggesting other miRNAs, isomiRs, or menthol metabolizing enzymes may be involved.
- Adopting a precision medicine approach to upregulate miR-374b isomiRs may preferentially benefit AAs, and help to reduce mortality and survival disparities.

Honors Thesis Proposal by Savanna Touré ('21)
For the Academic Year (2020-2021)
Advisor: Dr. Khadijah A. Mitchell
Committee: Drs. Elaine Reynolds and Daniel Griffith

I. TABLE OF CONTENTS

II. Biographical Sketch……5
III. Abstract……....7
IV. Introduction…….……8
- A. Racial disparities in lung cancer mortality and survival
- B. Known behavioral determinants of racial disparities in lung cancer: smoking
- C. Known biological determinants of racial disparities in lung cancer: genomic and metabolomic level
- D. Novel biological determinants of racial disparities in lung cancer: transcriptomic level
- E. References

V. Chapter 1: Identifying population-specific isomiRs targeting differentially expressed *UGTs* and *CYPs* in tumors from African Americans and European Americans with lung cancer
- A. Introduction……...……..…….…13
  1. MicroRNAs in cancer and cancer disparities: diagnosis, prognosis, and treatment
  2. IsomiRs in cancer and cancer disparities: diagnosis, prognosis, and treatment
  3. Hypothesis
- B. Methods……...……...……...15
  1. Clinico-demographic data extraction for the TCGA patient cohort
  2. Differentially expressed miRNAs by race and their MME-mRNA targets
  3. Differentially expressed isomiRs by race and their MME-mRNA targets
  4. Differentially expressed candidate isomiR analysis by potential confounding variables
  5. Differentially expressed MMEs targeted by candidate isomiR
  6. Integration analysis of isomiR-, miR-, and MME-expression by race
- C. Results……...……...…….18
  1. There were significant differences in age at diagnosis, smoking status, and vital status in AA and EA NSCLC patients
  2. There are 13 upregulated miRs targeting *UGTs* and *CYPs* in AA lung cancer patients
  3. There are two differentially expressed isomiRs by race from the highly abundant miR-374b
  4. Population-specific expression of candidate isomiR is independent of age and smoking status in LUAD patients
  5. isomiR374b, miR-374b-5p|3'a-1, targets three population-specific *UGTs* and *CYPs*
  6. isomiR and mRNA expression statuses are correlated in a subset of AA patients
- D. Discussion……...……...……..……20
- E. Tables and Figures……..……...……...……..…….22
  1. Table 1: Clinico-demographic characteristics of NSCLC patients from the TCGA cohort

2. Table 2: Differential expression of miRNAs by race targeting *UGTs* and *CYPs* in NSCLC patients
3. Table 3: Differential expression of isomiRs by race targeting *UGTs* and *CYPs* in NSCLC patients
4. Figure 1: Candidate isomiR expression, chromosomal location, and sequence
5. Figure 2: miR-374b and miR-374b-5p|3'a-1 expression by race, age, and smoking status in LUAD patients
6. Figure 3. MME expression by race, age, and smoking status in LUAD patients
7. Figure 4: Correlating matched isomiR-,miR-, and mRNA- expression levels in AA and EA patients with LUAD

F. References……...……...…..…....32

VI. Chapter 2: Determining miR-374b-5p|3'a-1 and targeted *UGT2B4* response upon exposure to menthol in cell lines from lung cancer patients

A. Introduction……...……...…….……35
1. miR-374b-5p expression and implications for its isomiRs in cancer
2. *CYP1B1* expression in cancer and unaffected smokers
3. *UGT2B4* expression in cancer
4. Hypothesis

B. Methods……...……...……...37
1. Cell line selection and tissue culture
2. qRT-PCR primer design
3. Determination of the appropriate CSC, L-menthol, and mCSC dose for LUAD cell lines
4. Menthol cigarette smoke condensate treatment
5. RNA isolation, DNase I treatment, and cDNA synthesis
6. *qPCR and Delta Delta Ct analysis of CYP1B1, UGT2B4, TRPM8, miR374b, and miR374b-5p|3'a-1*

C. Results……...……...……40
1. LUAD cell lines are matched by patient age and sex
2. qRT-PCR primers specifically amplify *CYP1B1*, *UGT2B4*, *TRPM8*, miR374b, and miR374b-5p|3'a-1
3. 10 µg/mL CSC and 40 µg/mL L-menthol is the most effective dose to elicit *CYP1B1* and *TRPM8* gene expression response
4. *CYP1B1* is induced in the presence of cigarette smoke condensate and menthol
5. *TRPM8* is induced in the presence of cigarette smoke condensate and menthol
6. *UGT2B4* is induced in the presence of menthol
7. miR-374b is induced in the presence of cigarette smoke condensate and menthol
8. miR-374b-5p|3'a-1 is induced in the presence of cigarette smoke condensate and menthol

D. Discussion……...……...……..……41

E. Tables and Figures……...……...……..…43
1. Table 1: LUAD cell line characteristics
2. Table 2. qRT-PCR primers
3. Figure 1: *CYP1B1* and *TRPM8* dose response after 6 hours of cigarette smoke condensate and L-menthol treatment
4. qRT-PCR primers
5. Figure 2: *CYP1B1* gene expression in LUAD cell lines after cigarette smoke condesate and menthol treatments
6. Figure 3: *TRPM8* gene expression in LUAD cell lines after cigarette smoke condesate and menthol treatments
7. Figure 4: *UGT2B4* gene expression in LUAD cell lines after cigarette smoke condesate and menthol treatments
8. Figure 5: miR-374b expression in LUAD cell lines after cigarette smoke condesate and menthol treatments
9. Figure 6: isomiR-374b expression in LUAD cell lines after cigarette smoke condesate and menthol treatments

F. References…….…….....50
VII. Conclusions……52
VIII. Acknowledgments……..53

II. Biographical Sketch

I am Savanna Touré, a senior at Lafayette College. I will obtain my Bachelors of Science degree in Neuroscience in May 2021. This is my third semester doing lung cancer disparities research with Dr. Khadijah A. Mitchell. I met her in the spring of my sophomore year (2018-2019) when I was actively searching for a Black mentor in the Biological Sciences. I was committed to working with her because of her diligence, as well as her mentoring and outreaching to Black students and community-centered work. This was the first time that I had met someone who looks like me in STEM and focused on my research interest, racial disparities. I started my research with her in the summer of 2019 where I focused on if racial differences exist among transcriptomic factors in lung cancer patients.

My health disparities research originally developed from my research in environmental determinants of reproductive health, cultivated when working with Dr. Melissa Perry and Ms. Laura Neumann. I then began working in Dr. Mitchell's Integrative and Translational Laboratory for Applied Biology (IT LAB) during my third year of college and decided to pursue a thesis for my upcoming senior year. The following summer of 2020, I performed research under Dr. Adam Engler at the University of San Diego School of Medicine. I continued to enhance my computational skills using applications, such as R Studio, to conduct *in silico* analysis on racial differences in the expression of an oncogene in triple-negative breast cancers from African Americans and European Americans. It was here that I began to refine the statistical skills necessary to carry out my honors thesis work.

After graduation, I will be completing a two-year post baccalaureate research fellowship under Dr. Derek C. Radisky at the Mayo Clinic in Jacksonville, Fl. I am working with breast cancer biological factors that exacerbate the racial disparities between African American women

compared to European Americans. I also plan to shadow and assist clinicians in the area, which will prepare me with the tools for applying, succeeding, and graduating from an NIH-funded M.D./P.hD. program focused on population health sciences.

III. Abstract

Background: Lung cancer is the leading cause of cancer-related deaths in the US, with the most prevalent subtypes being lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC). African Americans (AAs) have worse mortality and survival compared to European-Americans (EAs). Smoking negatively impacts treatment outcomes of lung cancer patients. However, AAs smoke less. When they do smoke, AAs choose menthol cigarettes more than EAs (88.5% vs. 25.7%). In previous studies, menthol-associated genomic (*MRGPRX4*) and metabolomic (menthol-glucuronide) biomarkers have been linked to AA smokers. To my knowledge, no lung cancer study has explored population-specific transcriptomic biomarkers, specifically, microRNA (miR) variants (isomiRs) that regulate *UGTs* and *CYPs* (menthol metabolizing enzymes, MMEs) before and after menthol exposure.
Hypothesis 1: Population-specific isomiRs drive racial differences in menthol-induced *UGT* and *CYP* expression.
Hypothesis 2: isomiR-374b-5p|3'a-1 expression is associated with race-specific expression changes of *CYP1B1* and *UGT2B4* upon menthol exposure in cell lines from AA and EA LUAD patients.
Research Methods: Clinico-, miRNA-, isomiR-, and mRNA-sequencing data for 1,046 miRNAs were obtained from the Broad GDAC Firehose for 394 LUAD patients (*n*=50 AAs, 342 EAs) and 294 LUSC patients (*n*=24 AAs, 270 EAs) in the TCGA cohort. Differential expression of miRNAs, isomiRs, *UGTs*, and *CYPs* by race were determined (one way ANOVA and two-tailed t-test with Welch's corrections). A correlation matrix integrated the various transcriptomic expression. Two cell lines, H1373 (AA) and A549 (EA), were exposed to cigarette smoke condensate (CSC) and menthol. After 6-hour exposure, qRT-PCR was performed. The Delta Delta CT method was used to determine population-specific expression levels for miRNA, isomiR, and target MMEs.
Results: A total of 13 miRNAs targeting 22 *UGT* and 47 *CYP* genes were differentially expressed (DE) in LUAD (*n*=10/1,046), LUSC (*n*=4/1,046), and LUAD and LUSC (*n*=1/1,046) patients. Only 3/13 DE miRNAs, in LUAD patients, had abundant population-specific isomiRs targeting MMEs (miR374b, miR96, and miR503). The most abundant isomiRs were among the canonical miR374b. In a subset of AA patients, high expression of miR-374b-5p|3'a-1 correlated with low expression of menthol-metabolizing gene regulation. Upon menthol exposure, lung cancer cells from AAs had significantly higher expression of miR-374b-5p|3'a-1.
Conclusion: This data suggests that AA-specific miR-374b-5p|3'a-1 is induced upon menthol exposure in AAs and may provide a functional link to MME gene regulation in this population. isomiR expression was not associated with *CYP1B1* and *UGT2B4* levels, but maybe associated with *CYP2U1*. If high isomiR-374b abundance is associated with low *CYP2U1* expression, this population-specific transcriptomic change has the possibility to be a novel therapeutic option for lung cancer patients. Adopting a precision medicine approach and developing an FDA-approved targeted therapy to upregulate isomiR-374b expression can help reduce the observed disparities.

IV. Introduction

A. Racial disparities in lung cancer mortality and survival

Lung cancer is the leading cause of cancer-related mortality in the U.S., with over 131,000 deaths per year, more than breast, prostate, colorectal, and brain cancers combined[1]. Five-year lung cancer survival is 21% which is one of the worst among all cancers, including pancreas (10%), liver (20%), and esophagus (20%)[1] carcinomas. Racial disparities exist in lung cancer mortality and survival. African Americans (AAs) have a higher mortality rate (48.6 vs. 46 per 100,000) and lower 5-year survival (17.1% vs. 19.7%) compared to European Americans (EAs)[2]. Higher mortality and lower survival rates in AAs are associated with later diagnoses at stages III and IV compared to EAs, where standard curative care is limited[3]. The racial differences in mortality and survival rates have also been attributed to population-specific smoking behaviors.

B. Known behavioral determinants of racial disparities in lung cancer: smoking

In addition to causing lung cancer, smoking after diagnosis can also impact a patient's therapeutic response and influence mortality and survival rates. Among lung cancer patients who are current smokers at the time of diagnosis, over 80% continue to smoke after diagnosis [6]. AAs smoke less and are more likely to be intermittent or light smokers compared to EAs. However, they still have worse mortality and survival outcomes than their EA counterparts[3].

Notably, smoking brand preferences differ among racial groups. Roughly 88% of AAs choose menthol cigarette brands, a regular tobacco with the flavoring of menthol, compared to 26% of EAs[4]. In the 1970s, tobacco companies heavily targeted AAs to purchase and consume menthol cigarettes. Menthol cigarettes are popular for AA smokers due to their cooling effect that masks the bitter taste of other toxicants in tobacco compared to traditional cigarettes[5]. Others

have suggested this masking effect leads to deeper inhalation amongst menthol cigarette smokers, which may contribute to a more aggressive lung tumor biology and the racial disparities in mortality and survival observed[6]. Population-specific smoking behaviors, like menthol cigarette preference, have also been implicated in poor outcomes after radiation and chemotherapy treatment[6]. Smoking cessation has been linked to decreased risk of cancer recurrence, better survival, and increased therapeutic response[7]. Tailored menthol smoking cessation programs would greatly benefit this population.

Cigarette smoking can impact metabolism through cytochrome P450 (*CYP*) enzymes, a key family of enzymes responsible for breaking down toxins and drugs. Previous studies have reported l-menthol has an inhibitory effect on *CYPs*, and fails to activate the enzymes[8]. Though researchers have explored menthol cigarette smoking as a behavioral determinant, biological changes may also be involved.

C. Known biological determinants of racial disparities in lung cancer: genomic and metabolomic level

The biological role of menthol cigarette smoking is widely debated. Interestingly, African-ancestry specific menthol biomarkers have been identified at the genomic (*MRGPRX4 DNA*) and metabolomic (menthol-glucuronide) levels [9,10,11]. An African ancestry-specific G variant for the *MRGPRX4* gene has been associated with menthol smoking and its cooling effect[10,11], which may contribute to longer consumption and deeper inhalation by AA smokers compared to EA smokers. Distinct metabolomic profiles in AAs may also be involved. Menthol-glucuronide, a menthol metabolism biomarker, is higher among AA smokers compared to EA smokers[11]. A major menthol detoxifying agent is *UGT2B7*. Along with *CYPs*, UDP-glucuronosyltransferases (*UGTs*) are another family of enzymes known to break down toxins and

drugs[11]. High levels of menthol-glucuronide in AAs suggests this process may be slower to clear tobacco carcinogens in AAs.

CYPs and *UGTs* are both induced by cigarette smoking, and break down potential tobacco carcinogens, like menthol[12]. In this thesis, I will refer to both enzyme families as menthol-metabolizing enzymes (MMEs). Previous research has noted that AAs with a population-specific *CYP2D6* mutation have a slower metabolic rate than EAs [13]. Clinical practice in this population should consider *CYP2D6* mutation status for drug therapy recommendations.

In general, few studies have explored menthol-associated transcriptomic (RNA) differences by race. Specifically, racial differences in MME expression (*UGTs* and *CYPs*) are rarely explored. Recently, our lab has shown population-specific expression of MMEs among lung cancer patients, with decreased expression of both *CYP* and *UGT* mRNAs in AAs[14]. Furthermore, the regulation of MME mRNAs through microRNAs and their variants may also be implicated in the racial differences at the transcriptomic level.

D. Novel biological determinants of racial disparities in lung cancer: transcriptomic level

Transcriptomic level changes are molecular indicators of genomic differences between AA and EA populations. Our lab has found differences of the coding transcriptome in lung tumors and normal tissues by race, identifying 40 genes that were population-specific between AAs compared with EAs[15]. Only recently has the non-coding transcriptome been examined as potential drivers of transcriptomic differences seen between AAs and EAs with lung cancer. For example, we have shown that high expression of miR-191 in lung tumors from AAs is associated with targeting and downregulation of (DE) the tumor suppressor *DUSP7* gene compared to EAs[15]. MicroRNAs drive the expression patterns of DE genes, thus, providing more support into

the study of population-specific expression of MMEs among lung cancer patients and implications for precision medicine interventions.

E. References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021 Jan;71(1):7-33. doi: 10.3322/caac.21654. Epub 2021 Jan 12. PMID: 33433946.
2. American Cancer Society (2019) Cancer Facts & Figures for African Americans 2019–2021. American Cancer Society, Atlanta.
3. Ryan BM. Lung cancer health disparities. Carcinogenesis. 2018 May 28;39(6):741-751. doi: 10.1093/carcin/bgy047. PMID: 29547922; PMCID: PMC5972630.
4. El-Toukhy S, Sabado M, Choi K. Trends in Susceptibility to Smoking by Race and Ethnicity. Pediatrics. 2016 Nov;138(5):e20161254. doi: 10.1542/peds.2016-1254. Epub 2016 Oct 17. Erratum in: Pediatrics. 2019 Sep;144(3): PMID: 27940778; PMCID: PMC5079079.
5. Alexander LA, Trinidad DR, Sakuma KL, Pokhrel P, Herzog TA, Clanton MS, Moolchan ET, Fagan P. Why We Must Continue to Investigate Menthol's Role in the African American Smoking Paradox. Nicotine Tob Res. 2016 Apr;18 Suppl 1(Suppl 1):S91-101. doi: 10.1093/ntr/ntv209. PMID: 26980870; PMCID: PMC6367903.
6. Hsu PC, Lan RS, Brasky TM, Marian C, Cheema AK, Ressom HW, Loffredo CA, Pickworth WB, Shields PG. Menthol Smokers: Metabolomic Profiling and Smoking Behavior. Cancer Epidemiol Biomarkers Prev. 2017 Jan;26(1):51-60. doi: 10.1158/1055-9965.EPI-16-0124. Epub 2016 Sep 14. PMID: 27628308; PMCID: PMC5386404.
7. Cataldo JK, Dubey S, Prochaska JJ. Smoking cessation: an integral part of lung cancer treatment. Oncology. 2010;78(5-6):289-301. doi: 10.1159/000319937. Epub 2010 Aug 11. PMID: 20699622; PMCID: PMC2945268.
8. Kramlinger VM, von Weymarn LB, Murphy SE. Inhibition and inactivation of cytochrome P450 2A6 and cytochrome P450 2A13 by menthofuran, β-nicotyrine and menthol. Chem Biol Interact. 2012 May 30;197(2-3):87-92. doi: 10.1016/j.cbi.2012.03.009. Epub 2012 Apr 1. PMID: 22486895; PMCID: PMC3362486.
9. Keeler C, Max W, Yerger V, Yao T, Ong MK, Sung HY. The Association of Menthol Cigarette Use With Quit Attempts, Successful Cessation, and Intention to Quit Across Racial/Ethnic Groups in the United States. Nicotine Tob Res. 2017 Nov 7;19(12):1450-1464. doi: 10.1093/ntr/ntw215. PMID: 27613927; PMCID: PMC6251684.
10. Kozlitina J, Risso D, Lansu K, Olsen RHJ, Sainz E, et al. (2019) An African-specific haplotype in MRGPRX4 is associated with menthol cigarette smoking. PLOS Genetics 15(2): e1007916. https://doi.org/10.1371/journal.pgen.1007916.
11. Hsu PC, Lan RS, Brasky TM, Marian C, Cheema AK, Ressom HW, Loffredo CA, Pickworth WB, Shields PG. Metabolomic profiles of current cigarette smokers. Mol Carcinog. 2017 Feb;56(2):594-606. doi: 10.1002/mc.22519. Epub 2016 Aug 22. PMID: 27341184; PMCID: PMC5646689.
12. Pathania S, Bhatia R, Baldi A, Singh R, Rawal RK. Drug metabolizing enzymes and their inhibitors' role in cancer resistance. Biomed Pharmacother. 2018 Sep;105:53-65. doi: 10.1016/j.biopha.2018.05.117. Epub 2018 May 26. PMID: 29843045.

13. Bradford LD, Kirlin WG. Polymorphism of CYP2D6 in Black populations: implications for psychopharmacology. Int J Neuropsychopharmacol. 1998 Dec;1(2):173-185. doi: 10.1017/S1461145798001187. PMID: 11281961.
14. Shrestha A. Exploring the Impact of Menthol Exposure on *UGT* and *CYP* mRNA Isoform Expression in Lung Cancers from African Americans and European Americans. Unpublished. 2019.
15. Mitchell KA, Zingone A, Toulabi L, Boeckelman J, Ryan BM. Comparative Transcriptome Profiling Reveals Coding and Noncoding RNA Differences in NSCLC from African Americans and European Americans. Clin Cancer Res. 2017 Dec 1;23(23):7412-7425. doi: 10.1158/1078-0432.CCR-17-0527. PMID: 29196495.

V. Chapter 1: Identifying population-specific isomiRs targeting differentially expressed UGTs and CYPs in tumors from African Americans and European Americans with lung cancer

A. Introduction

1. MicroRNAs in cancer and cancer disparities: diagnosis, prognosis, and treatment

MicroRNAs (miRNAs) are 22 base pair sequences that target mRNAs, and subsequently lead to the silencing of expression and inhibition of translation[12]. By partially binding to the 3'untranslated region (UTRs) of target mRNA, miRNAs can deregulate or cause the degradation of the mRNA[12]. MicroRNAs are key regulators of cell proliferation, apoptosis, and tumorigenesis, with evidence to suggest their importance in cancer diagnosis, prognosis, and treatment. MicroRNAs, such as miR-181a and a miR221/miR222 cluster, have been used to diagnose and determine prognosis of patients with cancer[12]. Treatment outcomes for patients have also been impacted by miRNAs through miRNA sponges, which have the ability to continuously inhibit miRNA function. MicroRNA sponges have been used to inhibit miR-10b in tumor cells which led to a >90% reduction in lung metastases[13]. Small regulatory molecules, such as miRNAs, continue to shed light on precision medicine approaches by contributing to diagnostic, prognostic, and treatment outcomes of patients. MicroRNAs have a canonical form, also known as mature miRNA.

2. IsomiRs in cancer and cancer disparities: diagnosis, prognosis, and treatment

MicroRNAs have variants, also known as isomiRs, that can differ in length, sequence, and biology compared to their mature miRNAs[14]. There are four types of isomiR sequences: (1) 3' isomiR, (2) 5' isomiR, (3) polymorphic isomiRs (changes within the internal mature miRNA sequence), and (4) mixed type isomiRs (two or more termini and polymorphic isomiRs)[14]. The 3' and 5' isomiRs can have alterations, such as addition and deletion, at 3'or 5' end with respect to

the canonical form. It has been suggested that isomiRs play a vital role in cancer diagnosis, prognosis, and treatment. Previous studies have noted isomiRs with cancer-specific expression or sensitivity that would allow it to serve as a better potential diagnostic biomarker compared to mRNA and mature miRNA[15,16]. Over 32 TCGA cancer types were specified by the presence or absence of 7,466 isomiRs[16]. This has implicated isomiRs as potential diagnostic and prognostic classifiers[16]. As novel diagnostic and prognostic tools, isomiRs have also been evaluated in potential precision medicine interventions. In previous studies, isomiR-specific sponges were used to inhibit the regulatory effect of isomiRs on genes associated with various diseases[17]. The potential to target oncogene regulation through canonical miRs and isomiRs remains promising and may address cancer racial disparities at the biological level.

IsomiRs may play a key role in cancer disparities. Population-specific isomiR expression has been suggested to interfere with normal breast molecular physiology, potentially associating racial differences seen with triple-negative breast cancer (TNBC)[15]. Differentially expressed (DE) isomiRs regulate differentially wired (DW) mRNAs in normal and tumor tissues, further impacting metastatic-associated and key cancer pathways[15]. This suggests that DE isomiRs by race may distinguish cancer types with greater specificity. Few studies have examined the role of DE isomiRs in lung cancer disparities[16,18,19]. Profiling population-specific isomiRs targeting DE MMEs by race may provide further insights into using isomiR-specific sponges for patients with lung cancer.

3. Hypothesis

I hypothesize that population-specific isomiRs drive racial differences in menthol-induced *UGT* and *CYP* expression.

B. Methods

1. Clinico-demographic data extraction for the TCGA patient cohort

Non-small cell lung cancer (NSCLC) is the most common lung cancer type (85% cases), with two predominant subtypes: lung adenocarcinoma (LUAD, 40%) and lung squamous cell carcinoma (LUSC, 25%)[20] . The clinical information for 74 African American (AA) and 612 European American (EA) NSCLC patients in The Cancer Genome Atlas (TCGA) study (50 AA and 342 EA LUAD patients; 24 AA and 270 EA LUSC patients) were downloaded from cBioportal for Cancer Genomics (https://www.cbioportal.org). Chi-squared and Kruskal Wallis statistical tests were performed to determine any confounding variables.

2. Differentially expressed miRNAs by race and their MME-mRNA targets

Annotated clinico-demographic information was subsequently merged with normalized miRNA- and mRNA-sequencing data downloaded from the Broad GDAC Firehose (http://gdac.broadinstitute.org/). All normalized data were imported into Partek Genomic Suite 7.0. The Partek Genomics Suite 7.0 miRNA Expression Workflow was used to carry out a differential expression by race with a one-way analysis of variance for all miRNAs (n = 1,046), P-value and False Discovery Rate <0.05. TargetScanHuman 7.2 was used to find target *UGT* and *CYP* mRNAs for the DE miRNAs. The DE miRNAs that did not regulate both MMEs were removed from further analysis.

3. Differentially expressed isomiRs by race and their MME-mRNA targets

Annotated clinico-demographic information was subsequently merged with normalized isomiR-sequencing data downloaded from the Broad GDAC Firehose (http://gdac.broadinstitute.org/). The isomiR-sequencing data was filtered to only include mature variants of DE canonical miRNAs. A nomenclature system from Telonis and colleagues[15] was

used to distinguish isomiRs. First, the canonical miRNA name was found in the miRBase database version 22.1 (http://www.mirbase.org/) (*e.g.* miR-374b). Next, the miRBase accession number was used to identify which arm of the premiRNA gives rise to the isomiR (*e.g.* accession number: MIMAT0004955, miR-374b-5p). The isomiR was determined relative to the canonical miRNA sequence using the human reference genome (*e.g.* miR-374b-5p|+2|-1).

I devised a nomenclature system to more clearly specify 3' and 5' isomiR positions, and any associated base pair additions and deletions. The isomiR coordinates were aligned with the GRCh37/hg19 human genome assembly (*e.g.* chrX:73,438,382-73,438,453) (https://genome.ucsc.edu/cgi-bin/hgGateway). The 3' and 5' end of the isomiR was listed. Additions were denoted by "a" and deletions by "t" for termination/deletion, followed by a numeral that reflected the number of base pair changes (*e.g.* miR-374b-5p|3'a-1|). Polymorphic isomiRs were filtered out.

A criterion for low and high isomiR expression, sensitivity, and specificity was adapted from Magee and colleagues[21]. Sensitivity is the ability to detect bonafide base changes in isomiRs as opposed to background sequencing errors. Specificity is the ability to differentiate between isomiRs from the same canonical miRNA. Lowly expressed miRs were characterized by $\leq$1.0 Read Per Million (RPM), high sensitivity, and low specificity. Highly expressed miRs had $\geq$ 5.0 RPM, low sensitivity, and high specificity. To balance between sensitivity and specificity of the isomiR, mean expression $\geq$3 RPM was used, which was moderately expressed, and assumed to have high sensitivity and high specificity. GraphPad Prism 9.0 was used to identify DE isomiRs (mean expression $\geq$3 RPM) by performing unpaired t-tests with Welch's corrections.

4. Differentially expressed candidate isomiR analysis by potential confounding variables

A candidate DE isomiR was selected based on meeting the following summary criteria: DE canonical miRNA by race (P+FDR <0.05) that targeted both *UGTs* and *CYPs* and DE isomiR by race (P <0.05) with ≥3 RPM and high sensitivity and specificity. The isomiR with the greatest RPM abundance was selected based on the assumption it had the strongest association with cancer. An unpaired t-test with Welch's correction in GraphPad Prism 9.0 was used to evaluate candidate DE isomiR expression in AAs and EAs with NSCLC stratified by relevant clinico-demographic variables.

5. Differentially expressed MMEs targeted by candidate isomiR

Annotated clinico-demographic information was subsequently merged with normalized mRNA-sequencing data downloaded from the Broad GDAC Firehose (http://gdac.broadinstitute.org/). The mRNA-sequencing data was filtered to only include the DE candidate isomiR target *UGTs* (n = 2) and *CYPs* (n = 16). An unpaired t-test with Welch's correction in GraphPad Prism 9.0 was used to evaluate DE *UGT* and *CYP* expression in AAs and EAs with NSCLC stratified by relevant clinico-demographic variables.

6. Integration analysis of miRNA-, isomiR-, and MME-expression by race

A total of 28/50 AA and 124/342 EA LUAD patients had all three data types (miRNA-, isomiR-, and MME mRNA-sequencing). Graph Pad Prism 9.0 was used to perform a correlation matrix analysis based on DE miRNA, isomiR, and MME mRNA expression. The p-values, correlation statuses, and R squared values were obtained from the data. Heat maps were generated for all patients, and then evaluated further by race.

C. Results

1. There were significant differences in age at diagnosis, smoking status, and vital status in AA and EA NSCLC patients

AAs with LUAD had significantly lower age of diagnosis compared to EAs (~60 vs. ~66) (P = <0.0001, Table 1, left). There was a significant racial difference in smoking status among LUAD patients (P = 0.0389, Table 1, left). There was a significant difference between vital status and race of LUSC patients (P = 0.0248, Table 1, right).

2. There are 13 upregulated miRs targeting UGTs and CYPs in AA lung cancer patients

There are 17/1,046 DE miRNAs by race. All DE miRNAs by race were upregulated in AAs compared to EAs (Table 2, third column). Of the 17 DE miRNAs, only 13 miRNAs targeted both *UGTs* and *CYPs*. Four miRNAs (miR-548j, miR-422a, miR-212, and miR-1912) were removed from further analyses due to not meeting the criteria of targeting both *UGT* and *CYP* enzymes. There were 10/13 LUAD-specific miRNAs, 2/13 were LUSC-specific miRNAs, and one was a pan-cancer miR present in both histologies (miR-1304) (Table 2).

3. There are two differentially expressed isomiRs by race from the highly abundant miR-374b

Three DE miRNAs by race (miR96, miR374b, miR503) had seven abundant candidate isomiRs (mean expression ≥3 RPM) (miR96: 2, miR374b: 4, miR503: 1). After performing differential expression by race for candidate isomiRs, miR96 was excluded for lack of population-specific isomiR expression. The remaining three candidate isomiRs are seen in the third column of Table 3. It is known that isomiRs can target different mRNAs than their canonical miRNA. The three candidate isomiRs have the same predicted MME targets as their canonical miRNAs (Table 3, right). The most highly abundant of the three candidate isomiRs

was miR-374b-5p|3'a-1, which is a LUAD specific isomiR (Figure 1A). The miR-374b-5p and its isomiRs are located on the q arm of the X chromosome (Figure 1B). The Human Genome Build GRCh37/hg19 showed an additional base on the 3'end of 5p (miR-374b-5p|3'a-1) (Figure 1C).

4. Population-specific expression of candidate isomiR is independent of age and smoking status in LUAD patients

Interactions with race were evaluated by stratifying patients based on potential confounding variables (age and smoking status, Table 1). Overall, the miR-374b and miR-374b-5p|3'a-1 isomiR showed significantly higher expression in AA patients (Figure 2, top). Age and smoking status did not change the direction of the race-specific expression pattern of miR-374b-5p|3'a-1. The racial difference is more pronounced in people below 70 years of age (Figure 2, bottom left). The racial difference is significant in current smokers compared to other groups (Figure 2, bottom right). Because there were only two AA never smokers, meaningful comparisons could not be made with the EA never smokers.

5. isomiR374b (miR-374b-5p|3'a-1) targets three population-specific UGTs and CYPs

Three MMEs (*UGT2B4, CYP1B1*, and *CYP2U1*) showed significantly decreased expression in AAs. Age and smoking status did not change the direction of the race-specific expression patterns. AAs were consistently downregulated when compared to EAs, independent of age and smoking status (Figure 3A-C, left). *UGT2B4* and *CYP2U1* had pronounced racial differences in patients below 70 years of age (Figure 3A and C, top right). *UGT2B4* had a more pronounced decrease in current smokers, while *CYP2U1* had differences in former smokers (Figure 3A and C, bottom right). *CYP1B1* had the same general trend of lower expression in AAs, however, this was not statistically significant.

6. isomiR and mRNA expression statuses are correlated in a subset of AA patients

The correlation matrix of transcriptomic expression showed no significant interaction between the miRNA, isomiR, *UGT2B4*, *CYP1B1*, and *CYP2U1* expression in AAs and EAs (Figure 4). In EAs, there were low miRNA and isomiR expressions that seemed to be associated with moderately high *CYP2U1* and high *CYP1B1* expression (Figure 4, top). AA patients had higher miR-374b expression that was also associated with lower *UGT2B4* and *CYP1B1* expression (Figure 4, bottom left). A subset of AA patients with the highest expression of miR-374b and its isomiR correlated with lower expression of *UGT2B4*, *CYP2U1*, and *CYP1B1*. This pattern is to be expected due to previous findings.

D. **Discussion**

Differential expression of isomiRs from canonical miRNAs showed racial differences between AAs and EAs. These patterns did not differ when accounting for age and smoking status, which supports my hypothesis regarding a race-specific influence. Among current and former smokers, there were significant racial differences in miR-374b-5p|3'a-1 expression. This could possibly be due to the active regulation of metabolizing enzymes (MMEs) that work to break down cigarette toxins, which shows a more pronounced racial difference.

In a subset of five AA patients with the highest expression of miR-374b-5p|3'a1, isomiR patterns were correlated with the downregulated expression of MMEs, *UGT2B4* and *CYP2U1*. This subset group could have biological factors that contribute to the expression patterns seen: specific genetic ancestral composition, smoking statuses, and treatment outcomes. The high expression of isomiRs have been seen to exert their binding affinity to the target mRNAs more effectively, thus either deregulating or degrading it[13]. The lack of this expression pattern within a sample set of the 23 remaining patients limited the scope of the analyses. Nonetheless, these

results may produce clinical implications behind the association of isomiRs with the expression of menthol metabolizers. The chapter 2 research focuses on racial differences in menthol-induced miR-374b-5p|3'a-1 and MME expression in cell lines from AA and EA LUAD patients. This will provide more insight into the lung cancer race disparities discussed in this chapter.

E. Tables and Figures

Table 1: Clinico-demographic characteristics of NSCLC patients from the TCGA cohort						
	AA LUAD (*n*=50)	**EA LUAD (*n*=342)**	***P***	**AA LUSC (*n*=24)**	**EA LUSC (*n*=270)**	***P***
Age (years)#			*0.0001182*			*0.861*
Mean (SD)	59.86 (10.18)	66.01 (9.91)		67.5 (7.33)	69 (8.82)	
Range	39-80	40-88		56-83	41.5-85	
Sex (%)^			*0.951*			*0.155*
Female	27 (54.0)	187 (54.7)		10 (41.67)	70 (25.93)	
Male	23 (46.0)	156 (45.3)		14 (58.33)	200 (74.07)	
Smoking Status (%)^			*0.0389*			*0.375*
Never Smoker	3 (6.0)	52 (15.2)		1 (4.35)	5 (1.92)	
Current Smoker	16 (32.0)	76 (22.2)		9 (39.13)	88 (33.72)	
Former Smoker	29 (58.0)	202 (59.1)		13	168 (64.37)	
Unknown	2 (4.0)	12 (3.5)		1 (4.35)	9 (3.33)	
Stage (%)^			*0.932*			*0.245*
I*	25 (50.0)	188 (55.0)		13 (54.17)	126 (47.19)	
II**	14 (28.0)	82 (24.0)		5 (20.83)	103 (38.58)	
III***	7 (14.0)	55 (16.1)		6 (25)	36 (13.48)	
IV	2 (4.0)	14 (4.1)			2 (0.75)	
NA	2 (4.0)	3 (0.88)				
Vital Status (%)^			*0.668*			*0.0248*
Dead	16 (32.0)	124 (36.3)		16 (66.67)	110 (40.74)	
Alive	34 (68.0)	218 (63.7)		8 (33.33)	160 (59.26)	
* Stage IA and IB, **Stage IIA and IIB, ***Stage IIIA, ^ Chi-squared test, #Kruskal Wallis Test						

Highlighted cells show significant differences ($P<0.05$).

Table 2. Differential expression of miRNAs by race targeting UGTs and CYPs in NSCLC patients

P value +FDR <0.05				Menthol-Metabolizing Enzymes (MMEs)	
HGNC Gene Symbol	**P-value (AA vs. EA)**	**AA vs. EA**	**Sub-Type**	***UGTs***	***CYPs***
miR-1304*	7.49E-17	AA up vs EA	LUAD	*UGT2B10, UGT3A2*	*CYP2U1,CYP1B1,CYP4F12,CYP3A7, CYP3A,CYP2S1,CYP19A1,CYP3A4, CYP46A1,CYP20A1,CYP27B1,CYP11, CYP27C1,CYP51A1,CYP2B6,CYP4F3, CYP7A1,CYP26B1,CYP2C19*
miR-3667	3.80E-06	AA up vs EA	LUAD	*UGT2B15,UGT2B10, UGT3A1,UGT2B4*	*CYP2U1,CYP2B6,CYP2C8,CYP2A6, CYP2A7,CYP2A13,CYP4V2,CYP3A5, CYP11A1,CYP27B1,CYP26B1,CYP8B1, CYP27C1,CYP20A1*
miR-491	2.75E-05	AA up vs EA	LUAD	*UGT1A10,UGT1A3, UGT2A3,UGT1A9, UGT1A4,UGT1A7, UGT1A8,UGT1A5, UGT1A6,UGT1A1, UGT8*	*CYP1B1,CYP2C19,CYP26B1,CYP20A1, CYP2C8,CYP17A1,CYP3A7,CYP3A4, CYP7B1,CYP39A1,CYP4A11,CYP4A22, CYP1A1,CYP51A1*
miR-96	9.33E-05	AA up vs EA	LUAD	*UGT2B7, UGT2B28, UGT8,UGT2B4*	*CYP3A5,CYP2E1,CYP7A1,CYP4Z1, CYP7B1,CYP4F3,CYP4F2,CYP19A1*
miR-374b	0.00016249	AA up vs EA	LUAD	*UGT2B4,UGT2B11*	*CYP2U1,CYP1B1,CYP26A1,CYP7B1, CYP26B1,CYP51A1,CYP2C8,CYP2J2, CYP39A1,CYP4F11,CYP2S1,CYP4A22, CYP4F2,CYP4F3,CYP4A11, CYP1A2*
miR-1252	0.000351896	AA up vs EA	LUAD	*UGT2A3,UGT3A2, UGT3A1,UGT2B4*	*CYP2U1,CYP2C19,CYP27B1,CYP4F22, CYP4F12,CYP20A1,CYP3A5,CYP4F3, CYP2J2,CYP7B1,CYP2A13,CYP2W1, CYP3A5,CYP4B1,CYP4V2,CYP11B1, CYP26B1*
miR-1537	0.000379101	AA up vs EA	LUAD	*UGT3A2*	*CYP27B1*
miR-3685	0.000395766	AA up vs EA	LUAD	*UGT2B11,UG2B4*	*CYP2C18,CYP8B1,CYP39A1,CYP4F2, CYP24A1,CYP51A1,CYP20A1,CYP4F3, CYP27C1,CYP4V2,CYP27B1*
miR-3611	0.000406077	AA up vs EA	LUAD	*UGT2A3, UGT2B28, UGT2B10,UGT2B15, UGT2B7,UGT2B17, UGT2A2,UGT8, UGT3A1,UGT2A1,*	*CYP24A1,CYP4F3,CYP7B1,CYP2S1, CYP46A1,CYP2B6,CYP4A22,CYP4A11, CYP20A1,CYP4F12,CYP2C8,CYP2A7, CYP3A5,CYP27B1,CYP8B1*

				UGT2B11	
miR-1243	0.000520395	AA up vs EA	LUAD	*UGT2B7,UGT2B28, UGT3A1*	*CYP3A5,CYP7B1,CYP20A1,CYP27B1, CYP3A5*
miR-503	0.000646378	AA up vs EA	LUAD	*UGT3A1,UGT2B4*	*CYP1B1, CYP2C19,CYP26B1,CYP7B1, CYP11B2,CYP2A7,CYP20A1*
miR-1304*	3.77E-11	AA up vs EA	LUSC	*UGT2B10, UGT3A2*	*CYP2U1, CYP1B1, CYP4F12,CYP3A7, CYP3A5,CYP2S1,CYP19A1,CYP3A4, CYP46A1,CYP20A1,CYP27B1,CYP11B1 CYP27C1,CYP51A1,CYP2B6,CYP4F3, CYP7A1,CYP26B1,CYP2C19*
miR-3929	7.26E-06	AA up vs EA	LUSC	*UGT2B10,UGT3A2*	*CYP1B1,CYP20A1,CYP1A2,CYP2C19, CYP1A1,CYP8B1,CYP4F3,CYP4F31P, CYP11A1,CYP4V2,CYP2B6,CYP11B1, CYP11B2,CYP21A2,CYP27C1,CYP51A1 ,CYP2U1,CYP2C8,CYP4F11*
miR-1298	1.55E-05	AA up vs EA	LUSC	*UGT2B28,UGT3A1, UGT2B11*	*CYP2A7,CYP51A1,CYP4B1,CYP27B1, CYP2U1,CYP26B1,CYP2A7*
miR-1912	0.000102689	AA up vs EA	LUSC		*CYP20A1,CYP3A5,CYP27C1,CYP2S1, CYP4B1,CYP19A1,CYP3A4,CYP3A5, CYP26B1,CYP27B1*

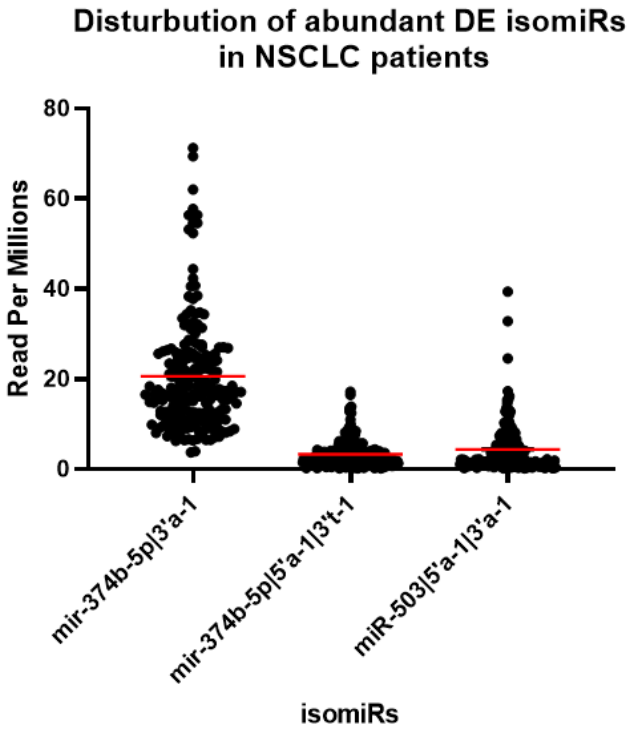
Highlighted cells show higher mean expression in AA patients ($P < 0.05$).

Table 3. Differential expression of isomiRs by race targeting *UGTs* and *CYPs* in NSCLC patients

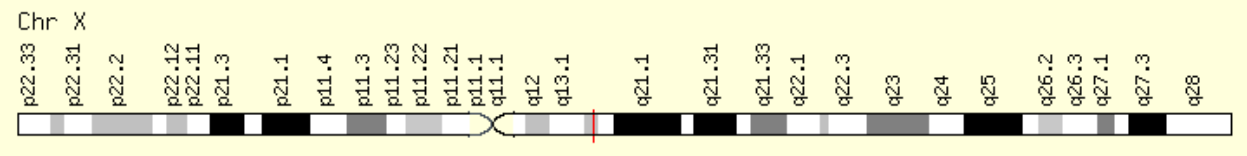
P value +FDR <0.05				Menthol-Metabolizing Enzymes (MMEs)	
HGNC Gene Symbol	**Abundant isomiR ID**	**DE isomiR ID**	**P value**	***UGTs***	***CYPs***
miR-96	miR-96-5p|3'a-1 miR-96-5p|5't-1|3'a-1			*UGT2B7* *UGT2B28* *UGT8,UGT2B4*	*CYP3A5,CYP2E1,* *CYP7A1,CYP4Z1,* *CYP7B1,CYP4F3,* *CYP4F2,CYP19A1*
miR-374b	miR-374b-5p|3'a-1 miR-374b-5p|5'a-1|3'a-1 miR-374b-5p|5't-1|3'a-1 miR-374b-5p|5't-2|3'a-1	miR-374b-5p|3'a-1	0.0163	*UGT2B4* *UGT2B11*	*CYP2U1,CYP1B1,* *CYP26A1,CYP7B1,* *CYP26B1,CYP51A1,* *CYP2C8,CYP2J2,* *CYP39A1,CYP4F11,* *CYP2S1,CYP4A22,* *CYP4F2,CYP4F3,* *CYP4A11,CYP7B1,* *CYP1B1,CYP1A2*
		miR-374b-5p|5'a-1|3'a-1	0.007	*UGT2B4* *UGT2B11*	*CYP2U1,CYP1B1,* *CYP26A1,CYP7B1,* *CYP26B1,CYP51A1,* *CYP2C8,CYP2J2,* *CYP39A1,CYP4F11,* *CYP2S1,CYP4A22,* *CYP4F2,CYP4F3,* *CYP4A11,CYP1A2*
miR-503	miR-503-5p|3'a-1	miR-503-5p|3'a-1	0.0472	*UGT3A1* *UGT2B4*	*CYP1B1,CYP2C19,* *CYP26B1,CYP7B1,* *CYP11B2,CYP2A7,* *CYP20A1*

Highlighted cells show significant differences among abundant DE isomiRs (P<0.05).

A)



B)



C)

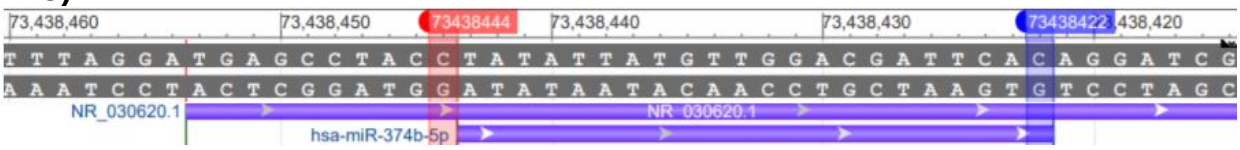


Figure 1. Candidate isomiR expression, chromosomal location, and sequence. (A) Scatter plot of abundant DE isomiR expression in NSCLC patients. The red line indicates the mean. (B) Chromosome map of DE miR-374b location on the human X chromosome shown with a red line. (C) UCSC Genome Browser Track highlighting the sequence and direction of transcription shown in red (start) and dark blue (end) for miR-374b-5p, the position of the DE candidate isomiR.

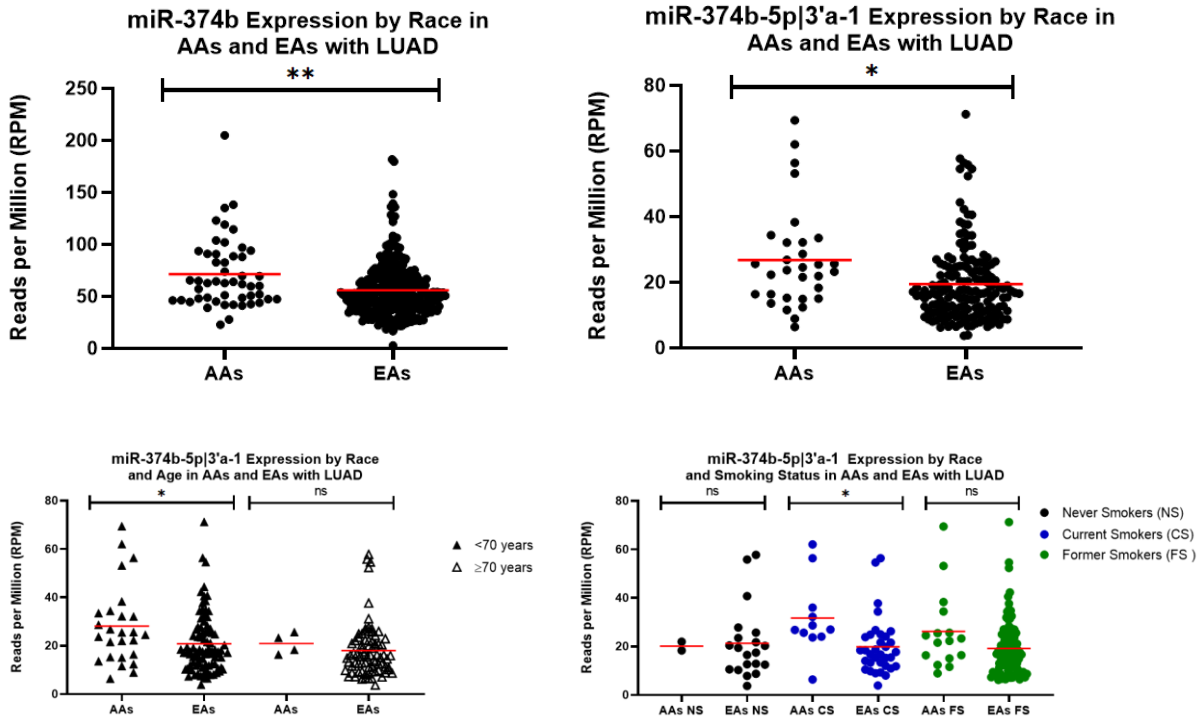


Figure 2. miR-374b and miR-374b-5p|3'a-1 expression by race, age, and smoking status in LUAD patients. Scatter plot showing a two-way interaction of patient race and age (<70 years and ≥70 years, bottom left) and patient race and smoking status (never smokers, current smokers, and former smokers, bottom right).
t-test with a Welch's correction, ns > 0.05, * < 0.05, ** < 0.01

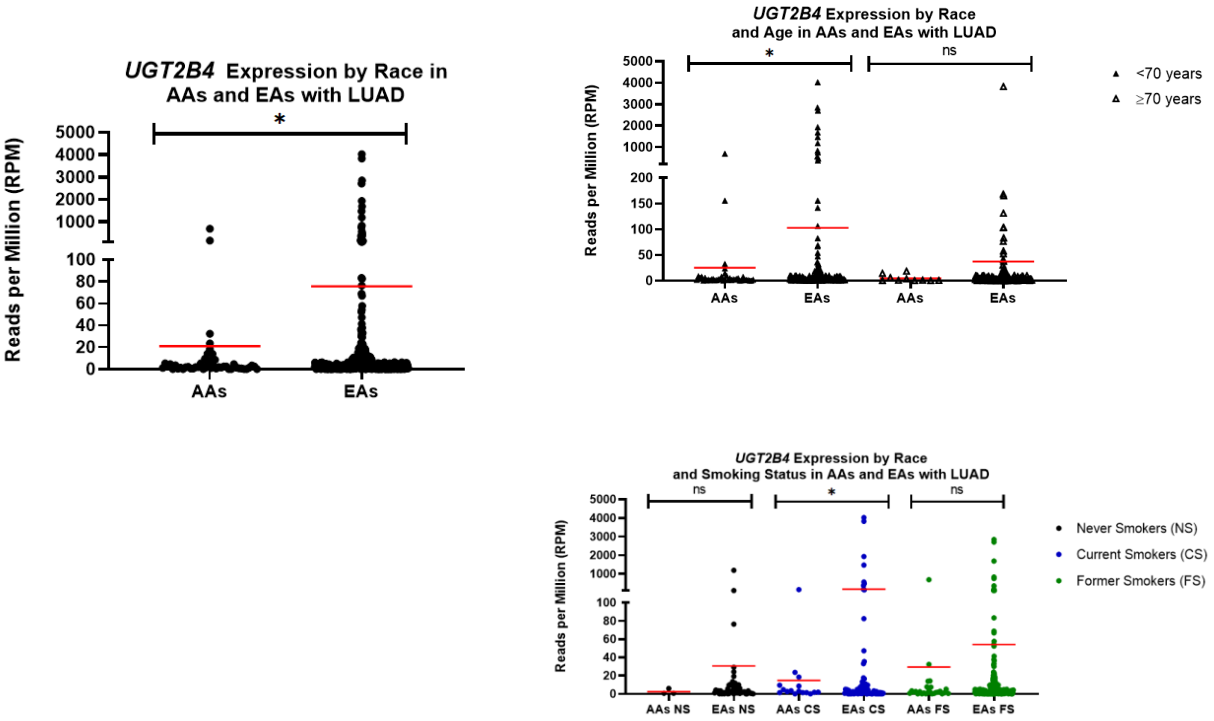


Figure 3. MME expression by race, age, and smoking status in LUAD patients. Scatter plot showing a two-way interaction of patient race and age (<70 years and ≥70 years, top right) and patient race and smoking status (never smokers, current smokers, and former smokers, bottom right) for (A) *UGT2B4*, (B) *CYP1B1*, and (C) *CYP2U1*.
t-test with a Welch's correction, ns > 0.05, * < 0.05, ** < 0.01

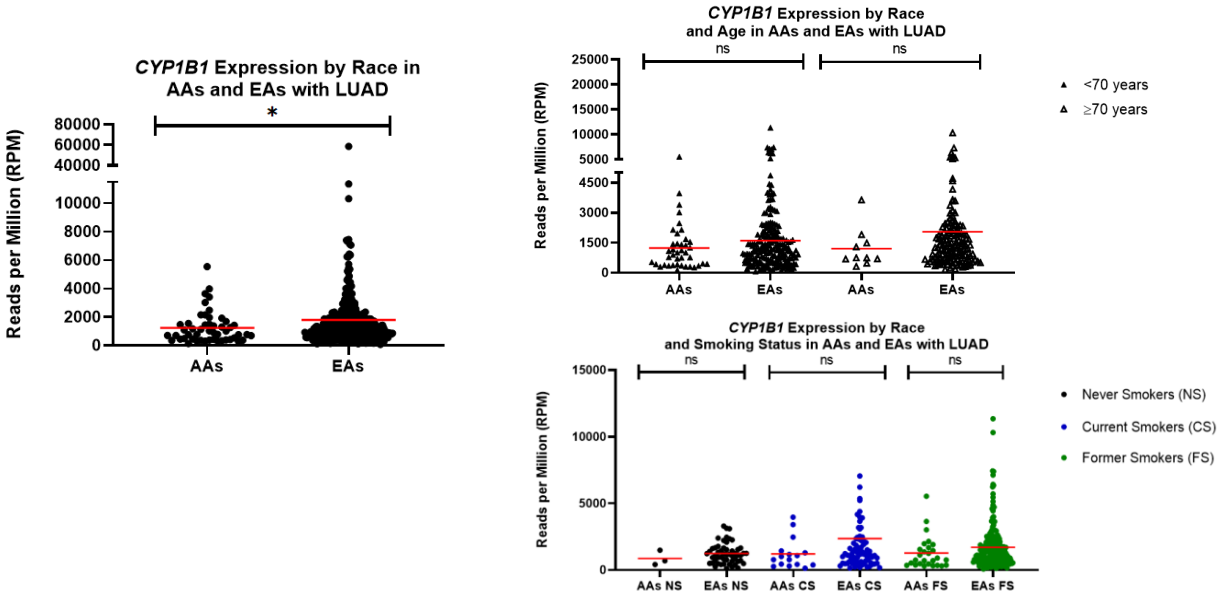


Figure 3. MME expression by race, age, and smoking status in LUAD patients. Scatter plot showing a two-way interaction of patient race and age (<70 years and ≥70 years, top right) and patient race and smoking status (never smokers, current smokers, and former smokers, bottom right) for (A) *UGT2B4*, (B) *CYP1B1*, and (C) *CYP2U1*.
t-test with a Welch's correction, ns > 0.05, * < 0.05, ** < 0.01

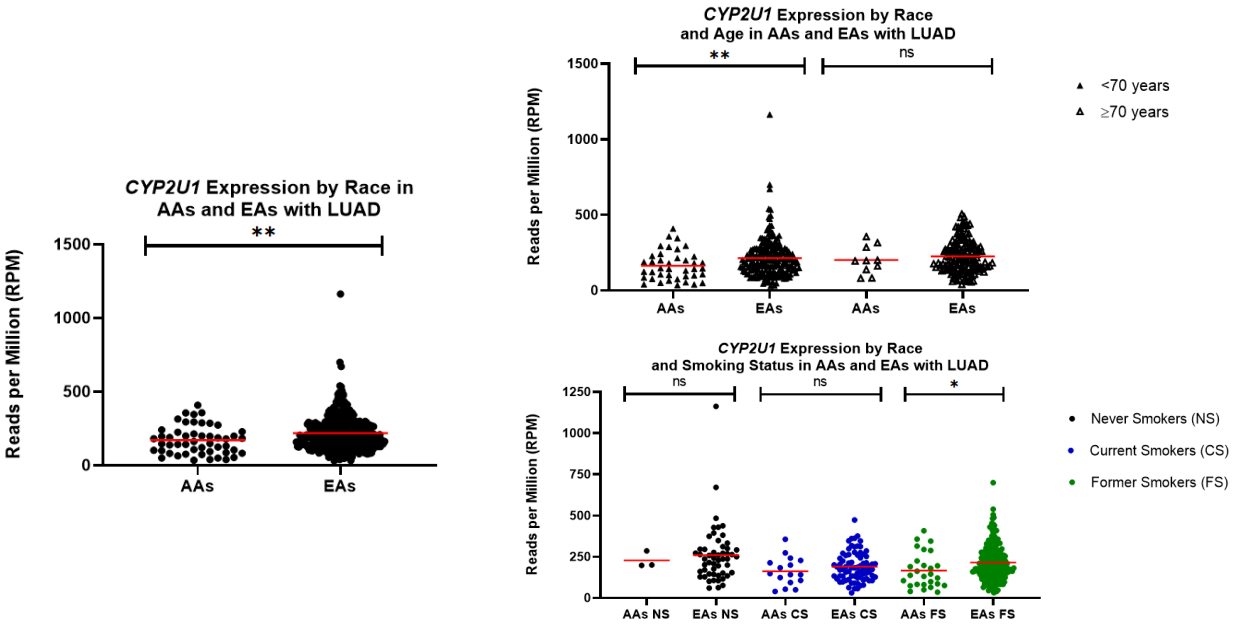


Figure 3. MME expression by race, age, and smoking status in LUAD patients. Scatter plot showing a two-way interaction of patient race and age (<70 years and ≥70 years, top right) and patient race and smoking status (never smokers, current smokers, and former smokers, bottom right) for (A) *UGT2B4*, (B) *CYP1B1*, and (C) *CYP2U1*.
t-test with a Welch's correction, ns > 0.05, * < 0.05, ** < 0.01

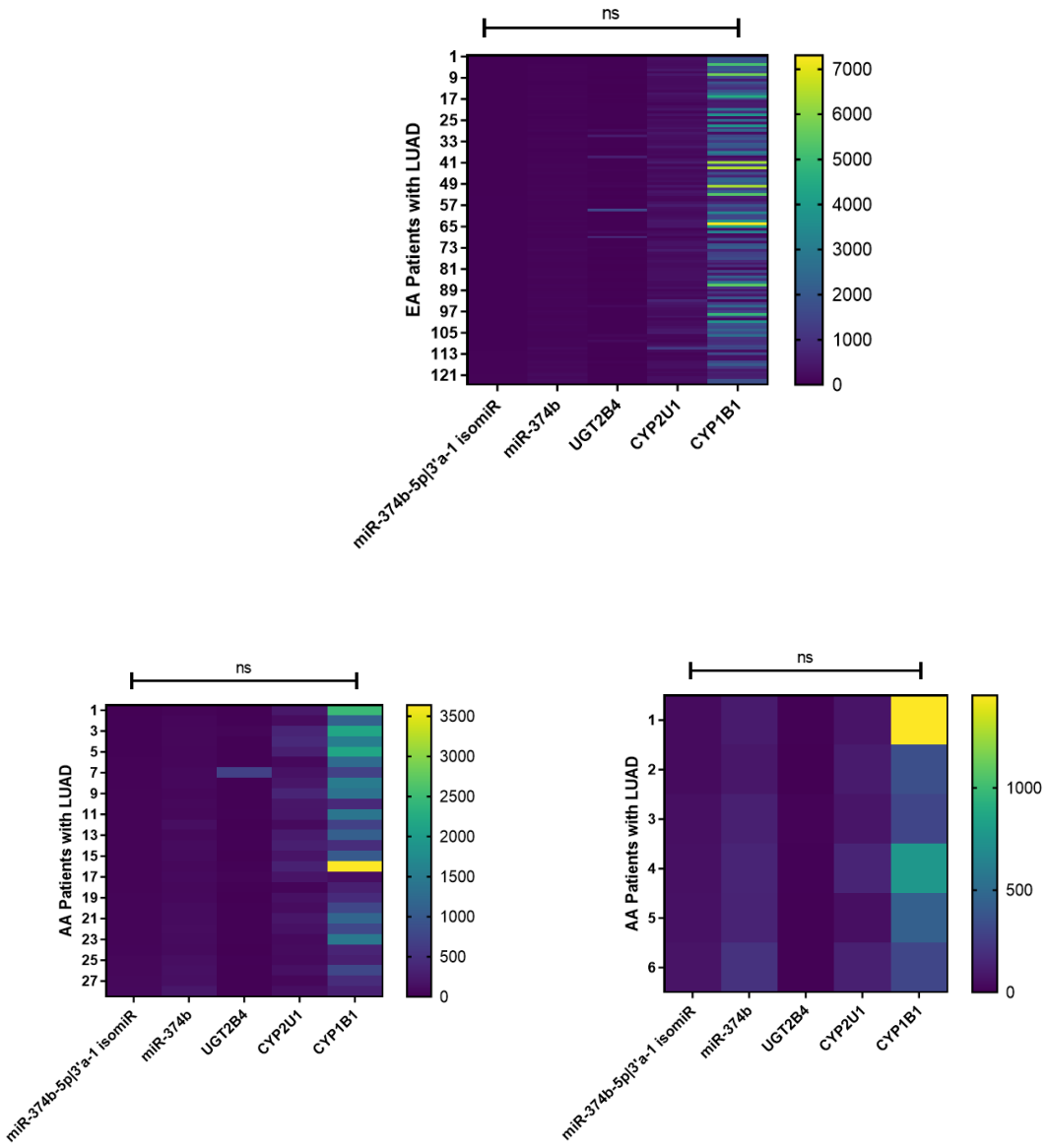


Figure 4. Correlating matched isomiR-,miR-, and mRNA- expression levels in AA and EA patients with LUAD. Heat map showing the cluster of expression levels by isomiR statuses for each patient in Reads Per Million (RPM). The top figure is EA expression cluster (n=124), bottom left is AA expression cluster (n=28), and bottom right is subset of AA expression cluster with the highest levels of isomiR 374b-5p|3'a-1 (n=5). (column effect ns > 0.05, * < 0.05).

F. References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021 Jan;71(1):7-33. doi: 10.3322/caac.21654. Epub 2021 Jan 12. PMID: 33433946.
2. American Cancer Society (2019) Cancer Facts & Figures for African Americans 2019–2021. American Cancer Society, Atlanta.
3. Ryan BM. Lung cancer health disparities. Carcinogenesis. 2018 May 28;39(6):741-751. doi: 10.1093/carcin/bgy047. PMID: 29547922; PMCID: PMC5972630.
4. El-Toukhy S, Sabado M, Choi K. Trends in Susceptibility to Smoking by Race and Ethnicity. Pediatrics. 2016 Nov;138(5):e20161254. doi: 10.1542/peds.2016-1254. Epub 2016 Oct 17. Erratum in: Pediatrics. 2019 Sep;144(3): PMID: 27940778; PMCID: PMC5079079.
5. Alexander LA, Trinidad DR, Sakuma KL, Pokhrel P, Herzog TA, Clanton MS, Moolchan ET, Fagan P. Why We Must Continue to Investigate Menthol's Role in the African American Smoking Paradox. Nicotine Tob Res. 2016 Apr;18 Suppl 1(Suppl 1):S91-101. doi: 10.1093/ntr/ntv209. PMID: 26980870; PMCID: PMC6367903.
6. Keeler C, Max W, Yerger V, Yao T, Ong MK, Sung HY. The Association of Menthol Cigarette Use With Quit Attempts, Successful Cessation, and Intention to Quit Across Racial/Ethnic Groups in the United States. Nicotine Tob Res. 2017 Nov 7;19(12):1450-1464. doi: 10.1093/ntr/ntw215. PMID: 27613927; PMCID: PMC6251684.
7. Kozlitina J, Risso D, Lansu K, Olsen RHJ, Sainz E, et al. (2019) An African-specific haplotype in MRGPRX4 is associated with menthol cigarette smoking. PLOS Genetics 15(2): e1007916. https://doi.org/10.1371/journal.pgen.1007916.
8. Hsu PC, Lan RS, Brasky TM, Marian C, Cheema AK, Ressom HW, Loffredo CA, Pickworth WB, Shields PG. Metabolomic profiles of current cigarette smokers. Mol Carcinog. 2017 Feb;56(2):594-606. doi: 10.1002/mc.22519. Epub 2016 Aug 22. PMID: 27341184; PMCID: PMC5646689.
9. Pathania S, Bhatia R, Baldi A, Singh R, Rawal RK. Drug metabolizing enzymes and their inhibitors' role in cancer resistance. Biomed Pharmacother. 2018 Sep;105:53-65. doi: 10.1016/j.biopha.2018.05.117. Epub 2018 May 26. PMID: 29843045.
10. Shrestha A. Exploring the Impact of Menthol Exposure on *UGT* and *CYP* mRNA Isoform Expression in Lung Cancers from African Americans and European Americans. Unpublished. 2019.
11. Mitchell KA, Zingone A, Toulabi L, Boeckelman J, Ryan BM. Comparative Transcriptome Profiling Reveals Coding and Noncoding RNA Differences in NSCLC from African Americans and European Americans. Clin Cancer Res. 2017 Dec 1;23(23):7412-7425. doi: 10.1158/1078-0432.CCR-17-0527. PMID: 29196495.
12. Wang, Wei, and Yun-ping Luo. “MicroRNAs in breast cancer: oncogene and tumor suppressors with clinical potential.” Journal of Zhejiang University. Science. B vol. 16,1 (2015): 18-31. doi:10.1631/jzus.B1400184
13. Ma L, Reinhardt F, Pan E, Soutschek J, Bhat B, Marcusson EG, Teruya-Feldstein J, Bell GW, Weinberg RA. Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model. Nat Biotechnol. 2010 Apr;28(4):341-7. doi: 10.1038/nbt.1618. Epub 2010 Mar 28. PMID: 20351690; PMCID: PMC2852471.
14. Lan C, Peng H, McGowan EM, Hutvagner G, Li J. An isomiR expression panel based novel breast cancer classification approach using improved mutual information. BMC Med

Genomics. 2018 Dec 31;11(Suppl 6):118. doi: 10.1186/s12920-018-0434-y. PMID: 30598116; PMCID: PMC6311920.
15. Telonis AG, Rigoutsos I. Race Disparities in the Contribution of miRNA Isoforms and tRNA-Derived Fragments to Triple-Negative Breast Cancer. Cancer Res. 2018 Mar 1;78(5):1140-1154. doi: 10.1158/0008-5472.CAN-17-1947. Epub 2017 Dec 11. PMID: 29229607; PMCID: PMC5935570.
16. Telonis AG, Magee R, Loher P, Chervoneva I, Londin E, Rigoutsos I. Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. Nucleic Acids Res. 2017 Apr 7;45(6):2973-2985. doi: 10.1093/nar/gkx082. PMID: 28206648; PMCID: PMC5389567.
17. Tan GC, Chan E, Molnar A, Sarkar R, Alexieva D, Isa IM, Robinson S, Zhang S, Ellis P, Langford CF, Guillot PV, Chandrashekran A, Fisk NM, Castellano L, Meister G, Winston RM, Cui W, Baulcombe D, Dibb NJ. 5' isomiR variation is of functional and evolutionary importance. Nucleic Acids Res. 2014 Aug;42(14):9424-35. doi: 10.1093/nar/gku656. Epub 2014 Jul 23. PMID: 25056318; PMCID: PMC4132760.
18. Guo L, Liang T, Yu J, Zou Q. A Comprehensive Analysis of miRNA/isomiR Expression with Gender Difference. PLoS One. 2016 May 11;11(5):e0154955. doi: 10.1371/journal.pone.0154955. PMID: 27167065; PMCID: PMC4864079.
19. Salem O, Erdem N, Jung J, Münstermann E, Wörner A, Wilhelm H, Wiemann S, Körner C. The highly expressed 5'isomiR of hsa-miR-140-3p contributes to the tumor-suppressive effects of miR-140 by reducing breast cancer proliferation and migration. BMC Genomics. 2016 Aug 8;17:566. doi: 10.1186/s12864-016-2869-x. PMID: 27502506; PMCID: PMC4977694.
20. Denisenko TV, Budkevich IN, Zhivotovsky B. Cell death-based treatment of lung adenocarcinoma. Cell Death Dis. 2018 Jan 25;9(2):117. doi: 10.1038/s41419-017-0063-y. PMID: 29371589; PMCID: PMC5833343.
21. Magee R, Loher P, Londin E, Rigoutsos I. Threshold-seq: a tool for determining the threshold in short RNA-seq datasets. Bioinformatics. 2017 Jul 1;33(13):2034-2036. doi: 10.1093/bioinformatics/btx073. PMID: 28203700; PMCID: PMC5870860.
22. Sun D, Wang X, Sui G, Chen S, Yu M, Zhang P. Downregulation of miR-374b-5p promotes chemotherapeutic resistance in pancreatic cancer by upregulating multiple anti-apoptotic proteins. Int J Oncol. 2018 May;52(5):1491-1503. doi: 10.3892/ijo.2018.4315. Epub 2018 Mar 14. PMID: 29568910; PMCID: PMC5873836.
23. Chang JT, Wang F, Chapin W, Huang RS. Identification of MicroRNAs as Breast Cancer Prognosis Markers through the Cancer Genome Atlas. PLoS One. 2016 Dec 13;11(12):e0168284. doi: 10.1371/journal.pone.0168284. PMID: 27959953; PMCID: PMC5154569.
24. Hanniford D, Zhong J, Koetz L, Gaziel-Sovran A, Lackaye DJ, Shang S, Pavlick A, Shapiro R, Berman R, Darvishian F, Shao Y, Osman I, Hernando E. A miRNA-Based Signature Detected in Primary Melanoma Tissue Predicts Development of Brain Metastasis. Clin Cancer Res. 2015 Nov 1;21(21):4903-12. doi: 10.1158/1078-0432.CCR-14-2566. Epub 2015 Jun 18. PMID: 26089374; PMCID: PMC4631639.
25. Li J, Zhang X, Tang J, Gong C. MicroRNA-374b-5p Functions as a Tumor Suppressor in Non-Small Cell Lung Cancer by Targeting FOXP1 and Predicts Prognosis of Cancer Patients. Onco Targets Ther. 2020 May 15;13:4229-4237. doi: 10.2147/OTT.S243221. PMID: 32523358; PMCID: PMC7237128.

26. Wang Y, Yu L, Wang T. MicroRNA-374b inhibits the tumor growth and promotes apoptosis in non-small cell lung cancer tissue through the p38/ERK signaling pathway by targeting JAM-2. J Thorac Dis. 2018 Sep;10(9):5489-5498. doi: 10.21037/jtd.2018.09.93. PMID: 30416798; PMCID: PMC6196188.
27. Xia J, Zhang W. A meta-analysis revealed insights into the sources, conservation and impact of microRNA 5'-isoforms in four model species. Nucleic Acids Res. 2014 Feb;42(3):1427-41. doi: 10.1093/nar/gkt967. Epub 2013 Oct 30. PMID: 24178030; PMCID: PMC3919606.
28. Jones NR, Lazarus P. UGT2B gene expression analysis in multiple tobacco carcinogen-targeted tissues. Drug Metab Dispos. 2014 Apr;42(4):529-36. doi: 10.1124/dmd.113.054718. Epub 2014 Jan 23. PMID: 24459179; PMCID: PMC3965906.
29. Dates CR, Fahmi T, Pyrek SJ, Yao-Borengasser A, Borowa-Mazgaj B, Bratton SM, Kadlubar SA, Mackenzie PI, Haun RS, Radominska-Pandya A. Human UDP-Glucuronosyltransferases: Effects of altered expression in breast and pancreatic cancer cell lines. Cancer Biol Ther. 2015;16(5):714-23. doi: 10.1080/15384047.2015.1026480. PMID: 25996841; PMCID: PMC4622877.
30. Wijayakumara DD, Mackenzie PI, McKinnon RA, Hu DG, Meech R. Regulation of UDP-Glucuronosyltransferases UGT2B4 and UGT2B7 by MicroRNAs in Liver Cancer Cells. J Pharmacol Exp Ther. 2017 Jun;361(3):386-397. doi: 10.1124/jpet.116.239707. Epub 2017 Apr 7. PMID: 28389526.

VI. Chapter 2: Determining miR-374b-5p|3'a-1 and targeted *UGT2B4* response upon exposure to menthol in cell lines from lung cancer patients

A. Introduction

1. miR-374b-5p expression and implications for its isomiRs in cancer

Overexpression of miR-374b-5p has been implicated in several cancers, including in breast, head and neck, gastric, prostate, and melanoma[1]. Furthermore, the down regulation of miR-374b-5p has also been associated with cancer[1]. The role of miR-374b-5p expression is different in various types of tumors and needs to be further investigated. The expression of miR-374b-5p has shown promise as a diagnostic and prognostic marker [2,3]. Few studies have investigated miR-374b-5p and its effect on non-small cell lung cancer (NSCLC) cells, and none have examined racial differences. Previous research showed the downregulation of miR-374b-5p in tumors compared to normal tissues, providing more insight into its tumor suppressor ability[4,5].

Though research has not been conducted on miR-374b-5p isomiRs, they may also be important for gene expression and miRNA-mRNA targeting. The miRNA seed region is important in regulating the target genes of the canonical miRNA. Both the canonical miRNA and its isomiRs contain the same miRNA seed region. A study has shown isomiRs can interrupt the canonical miRNA function, which in turn impacts their ability to regulate the target genes[6]. It is possible that miR-374b-5p may have isomiRs that regulates its target genes, like *CYP1B1* and *UGT2B4*.

2. CYP1B1 expression in cancer and unaffected smokers

CYPs are responsible for breaking down toxins including cigarette smoke and menthol. *CYP1B1* has decreased expression in African Americans (AAs) compared to European Americans (EAs) with NSCLC (Unpublished, Shrestha). Cigarette smoking impacts gene

expression profiles[7,8]. Leeuween and colleagues examined gene expression in blood cells from smokers and nonsmokers, where they found *CYP1B1* was differentially expressed (DE) and upregulated in smokers compared to nonsmokers[7].

3. UGT2B4 expression in cancer

The *UGT2B* sub-family, along with *CYPs*, are responsible for menthol glucuronidation and control the speed of metabolism for compounds such as bile, steroids, chemotherapeutic agents, and drugs[9]. Lung, aerodigestive, and pancreatic cancers are associated with smoking and lower *UGT2B* expression[7]. *UGT2B4* has been implicated in risk and development of breast, pancreatic, and liver tumorigenesis[10,11]. To my knowledge, no mechanistic studies in lung cancer cell lines have examined *CYP1B1* and *UGT2B4* expression upon menthol cigarette smoke exposure.

4. Hypothesis

I hypothesize isomiR-374b-5p|3'a-1 expression is associated with race-specific expression changes of *CYP1B1* and *UGT2B4* upon menthol exposure in cell lines from AA and EA LUAD patients.

B. Methods

1. Cell line selection and tissue culture

An AA (NCI-H1373) and EA (A549) cell line pair was purchased from National Cancer Institute (https://dtp.cancer.gov/repositories/) and American Type Culture Collection (https://www.atcc.org/). The cell line pair was matched based on age, sex, histology, and culture properties. The selected cell lines were were cultured with complete medium, containing RPMI-1640 medium, 10% fetal bovine serum, 1% glutamine, and 1% penicillin-streptomycin. The incubation conditions were 37.0°C and 5.0% CO_2 with subculturing every 2 to 3 days.

2. qRT-PCR primer design

Quantitative reverse transcription PCR (qRT-PCR) primers for *CYP1B1*, *UGT2B4*, *TRPM8* (biological control for menthol exposure), *GAPDH* (housekeeping control for mRNA expression), candidate miR-374b-5p, and isomiR-374b-5p|3'a-1 were designed using NCBI Gene and Primer-BLAST. Each target mRNA was searched in NCBI Gene with a *Homo sapiens* filter. The RefSeq transcript accession numbers were obtained; under the featured section, Primer-BLAST was selected. The following parameters were chosen: PCR product size: Min-Max 70-200; Min-Max 50-65; Exon junction span: Primer must span an exon-exon junction. For miRNA and isomiRs, the sequences were obtained from NCBI Gene transcript and expanded downstream for the 3' addition nucleotide. *U6* (housekeeping control for miRNA expression) primers were provided by the Mir-X miRNA First Strand Synthesis Kit and TB Green (Cat# 638316).

Primers were resuspended in DNAse and RNAse free water, based on the amount of nmol provided, to make 100 µM stock solutions. The 10 µM of forward and reverse primers were combined in a PCR tube with an additional 180 µL of RNase free water.

3. Determination of the appropriate CSC, L-menthol, and mCSC dose for LUAD cell lines

Since no lung cancer study has used menthol cigarette smoke condensate (mCSC), a trial was conducted with varying levels of cigarette smoke condensate (CSC) and L-menthol to determine ideal treatment concentrations. A dose-response experiment of CSC and L-menthol was performed on the A549 cell line to determine appropriate concentrations to elicit the expression of *CYP1B1* and *TRPM8*. In the clinically relevant range of 0-100 µg/mL CSC, doses in the assay included 0, 5, and 10 µg/mL[14,15]. These doses have maintained 90% or greater

viability of A549 cells and were used to avoid high apoptosis. L-menthol doses in the assay contained 0, 10, 20, and 40 µg/mL[16]. Cells were grown to 10^6 per well in a 6-well plate and treated for 6 hours. Cells underwent total RNA isolation, cDNA synthesis, and quantitative PCR (qRT-qPCR) as described in Methods section five. The Delta Delta Cycle threshold (ΔΔ Ct) method was used to analyze the gene expression status for *CYP1B1*, *TRPM8*, and *GAPDH*. Graphpad Prism 9.0 was used to create bar graphs of the dose responses.

4. Menthol cigarette smoke condensate treatment

The CSC solution had a stock concentration of 40 mg/mL and was purchased from Murty Pharmaceuticals. L-menthol was obtained from Sigma-Aldrich (Cat# PHR1116). NCI-H1373 and A549 cell lines were seeded in a 6-well plate at a concentration of 10^6 cells per well and were treated for 6 hours with 10 µg/mL CSC, 40 µg/mL L-menthol, and 50 µg/mL mCSC[17,14], similar to the average commercial menthol cigarette concentration (2.9 to 19.6 mg/cigarette)[18]. Cells were harvested in TRI reagent and RNA was isolated. Two biological replicates were completed.

5. RNA isolation, DNase I treatment, and cDNA synthesis

Zymo Direct-zol RNA MiniPrep kit (Cat# 11-330T) was used to isolate small and large (up to 10 µg) RNAs resuspended in TRI reagent, this was in accordance with the manufacturer's protocol. DNAse I treatment (400 µl RNA Wash Buffer, 5 µl DNase I (6 U/µl), and 75 µl DNA Digestion Buffer) was applied to remove remaining genomic DNA contents from the samples. Quantification of RNA occurred with the Eppendorf BioSpectrometer. All isolated RNA were stored at -20ºC.

The Applied Biosystems High-Capacity RNA-to-cDNA Kit (Cat# 4387406) was used for mRNA to cDNA synthesis, following the manufacturer's protocol. A total concentration of 1,000

ng of RNA was used and resuspended in RNAse-free water to a final volume of 9 µL. In a PCR tube, 10 µL 2X RT Buffer Mix and 1 µL 20X RT enzyme Mix were added to the 9 µL solution, this yielded a working volume of 20 µL per treatment. Then, the PCR tubes were used in the Applied Biosystems MiniAmp Thermal Cycler with the following conditions: 37 ºC for 60 minutes, 95 ºC for 5 minutes and a 4 ºC infinite hold. The Mir-X miRNA First Strand Synthesis Kit and TB Green (Cat# 638316) was used for miR and isomiR-cDNA synthesis, following the provided protocol.

6. qPCR and Delta Delta Ct analysis of CYP1B1, UGT2B4, TRPM8, miR374b, and miR374b-5p|3'a-1

A total of 50 ng (5 ng/uL) of cDNA template was used for quantitative PCR. The total reaction volume for each well was 28 µL (2 µL of forward and reverse primers, 6 µL of RNase-free water, and 10 µL Applied Biosystems SYBR Green PowerUp Master Mix (Cat# 4387406)). The mRNA qPCR was performed using the Applied Biosystems QuantStudio 3 machine and underwent the following conditions: 50 ºC for 2 minutes, 95 ºC for 2 minutes and 40 cycles of 95 ºC for 15 seconds, 58 ºC for 30 seconds and 72 ºC for 1 minute, followed by standard melt curve conditions. This process was repeated for qPCR of the miR and isomiR using the Mir-X miRNA qRT-PCR TB Green Kit (Cat#638316) with the following conditions: 95 ºC for 10 seconds, 40 cycles of 95 ºC for 5 seconds and 58 ºC for 20 seconds, followed by standard melt curve conditions. Gene and miR expression amplification plots using the Delta Delta Ct method were obtained and used to analyze expression statuses for treated and untreated cell lines.

C. Results

1. LUAD cell lines are matched by patient age and sex

Two lung cancer cell lines from EA and AA LUAD patients were selected. The pair was matched by age (58 vs 56 years) and sex (both males) (Table 1).

2. qRT-PCR primers specifically amplify *CYP1B1, UGT2B4, TRPM8,* miR374b, and miR374b-5p|3'a-1

Eight forward and reverse primers were designed for four mRNAs. Only forward primers were designed for the miRNA and isomiR according to the manufacturer's protocol. The reverse primers and *U6* primers were proprietary and provided by the company (Table 2).

3. 10 µg/mL CSC and 40 µg/mL L-menthol is the most effective combined dose to elicit CYP1B1 and TRPM8 gene expression response

After completing a dose-response experiment, 10 µg/mL CSC and 40 µg/mL induced *CYP1B1*. While 40 µg/mL L-menthol induced *TRPM8* expression (Figure 1). The expression changes for treated groups were more than 4 fold higher than untreated.

4. CYP1B1 is induced in the presence of cigarette smoke condensate and menthol

CYP1B1 showed greater fold changes under all three treatment conditions compared to untreated groups (Figure 2). AA CSC treated cells had a lower expression of *CYP1B1* compared to EA CSC treated cells. AA L-menthol treated cells had a higher expression of *CYP1B1*. AA mCSC treated cells had greater *CYP1B1* expression compared to EA mCSC treated cells. As expected from the dose response experiment, the CSC treatment resulted in noticeable induction of *CYP1B1*, which is known to have increased expression in cigarette smokers within an hour upon exposure.

5. TRPM8 is induced in the presence of cigarette smoke condensate and menthol

TRPM8 showed a greater fold change under all three treatment conditions compared to untreated (Figure 3). However, this trend was more pronounced in the AA cell line. Interestingly, the highest expression was in L-menthol treated cells.

6. UGT2B4 is induced in the presence of menthol

UGT2B4 showed an increased fold change in only the menthol treated groups (Figure 4). This finding was most noticeable in the AA cells. It is possible that *UGT2B4* induction may be due to miR374b and its isomiR.

7. miR-374b is induced in the presence of cigarette smoke condensate and menthol

miR-374b showed a greater fold change under all three treatment conditions compared to untreated (Figure 5). However, this result was more prominent in the AA cell line.

8. miR-374b-5p|3'a-1 is induced in the presence of cigarette smoke condensate and menthol

miR-374b-5p|3'a-1 showed a greater fold change under all three treatment conditions compared to untreated (Figure 6). However, this result was more prominent in the AA cell line.

D. Discussion

In this chapter, I hypothesized isomiR-374b-5p|3'a-1 expression was associated with race-specific *CYP1B1* and *UGT2B4* expression changes after being exposed to menthol. Overall, I observed the upregulation of canonical miRNA and isomiR-374b expression in AA cells when exposed to menthol (L-menthol alone or mCSC). By contrast, EA cells had no expression of miR-374b or its isomiR, even in the presence of any menthol exposure. The racial difference in isomiR expression was significant, further supporting my hypothesis in chapter 1.

In L-menthol and mCSC treated AA lung cancer cells, increased isomiR expression corresponded with high *CYP1B1* and *UGT2B4* expression. These expression patterns were contrary to what I expected. Chapter 1 demonstrated a lower expression of *UGT2B4* in lung tumors from AAs compared with EAs. The chapter 2 finding may be explained by the other DE by race and abundant miR-374b isomiR, miR-374b-5p|5't-2|3'a-1. This isomiR also contains a miR-374b seed region that could target *CYP1B1* and *UGT2B4*. The concentrations I used may not be adequate to induce this other isomiR. Additionally, there are other miRNAs that target *CYP1B1* and *UGT2B4*. Upon cigarette smoke exposure, some miRNAs may have decreased expression that would correspond with higher menthol metabolizing enzymes (MMEs). This assumption is supported by a recent cancer study that showed miR-217 was downregulated in the presence of cigarette smoke, which correlated with an increased expression of *KLK7*[19].

The *TRPM8* receptor was upregulated in lung cancer cells upon menthol exposure, supporting previous research[12,20]. AA cells had higher *TRPM8* expression compared to EA cells. The AA cell line was from a patient with a history of smoking. AAs are more likely to be menthol smokers, including the AA patient in which the NCI-H1373 cell line was derived. These AA cells may be more sensitive to L-menthol and mCSC due to previous menthol exposure. The EA cell line had no record of smoking history. Surprisingly, the CSC also upregulated *TRPM8*. This may be due to the fact that non-mentholated cigarettes contain <2.9 mg of menthol per cigarette[15]. This is compared to the average commercial menthol cigarette, which is between 2.9-19.6 mg per cigarette[15].

The limitations to my study are the use of only one AA and EA cell line pair and performing analysis of two biological replicates. In the future, I would expand to include more cell line pairs and biological triplicates. A future direction would include exploring the role of

miR-374b-5p|3'a-1 and *CYP2U1* in lung cancer. *CYP2U1* is DE by race and targeted by miR-374b and its isomiRs. In a recent study, *CYP2U1* overexpression was associated with poor prognosis and lower diseases-specific survival in triple negative breast cancer, a very aggressive subtype with a known health disparity in mortality between AA and EA women[20,21].

E. Tables and Figures

Table 1. LUAD cell line characteristics			
Cell Line Name	**Race**	**Age**	**Sex**
A549	EA	58	male
NCI-H1373	AA	56	male

Table 2. qRT-PCR primers		
ID	**Forward Sequence**	**Reverse Sequence**
CYP1B1	GCCACTATCACTGACATCTTCGG	CACGACCTGATCCAATTCTGCC
UGT2B4	CTTTAGGACTCAATACTCGGCTG	CTCATAGATGCCATTGGCTCCAC
TRPM8	AATTGCGATGCTGAGGGCTA	CACGAGCAGCAAATGTGTGT
GAPDH	CTCCTGTTCGACAGTCAGCC	ACCAAATCCGTTGACTCCGAC
miR-374b	ATATAATACAACCTGCTAAGTG	Provided by Mir-X miRNA qRT-PCR TB Green Kit (TakaRaBio/638316)
miR-374b-5p |3'a-1	ATATAATACAACCTGCTAAGTGT	Provided by Mir-X miRNA qRT-PCR TB Green Kit (TakaRaBio/638316)
U6	Provided by Mir-X miRNA qRT-PCR TB Green Kit (TakaRaBio/638316)	Provided by Mir-X miRNA qRT-PCR TB Green Kit (TakaRaBio/638316)

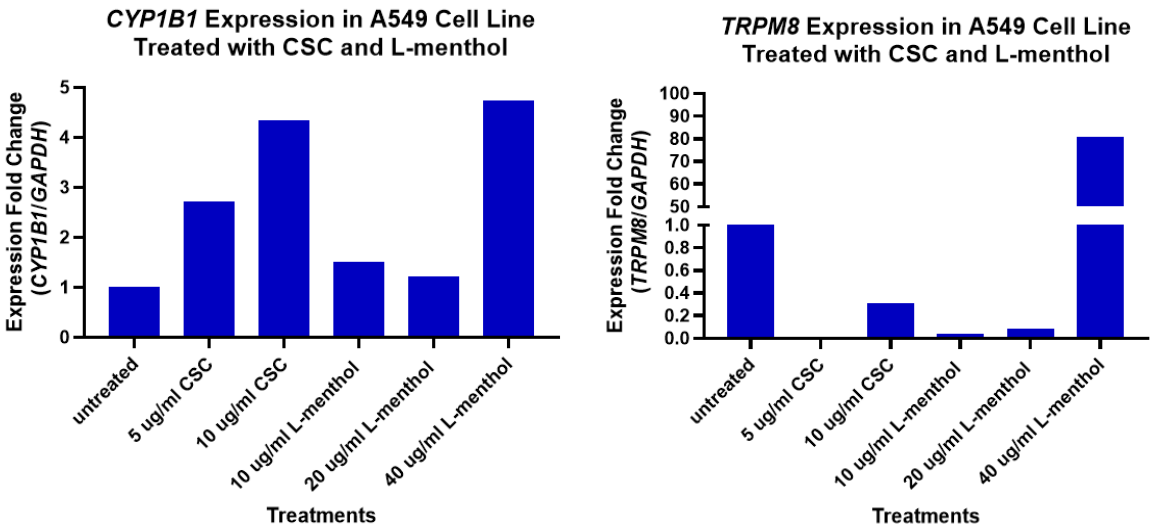


Figure 1. *CYP1B1* and *TRPM8* dose response after 6 hours of cigarette smoke condensate and L-menthol treatment. Quantitative reverse transcription PCR results showed expression fold change for five different treatment groups.

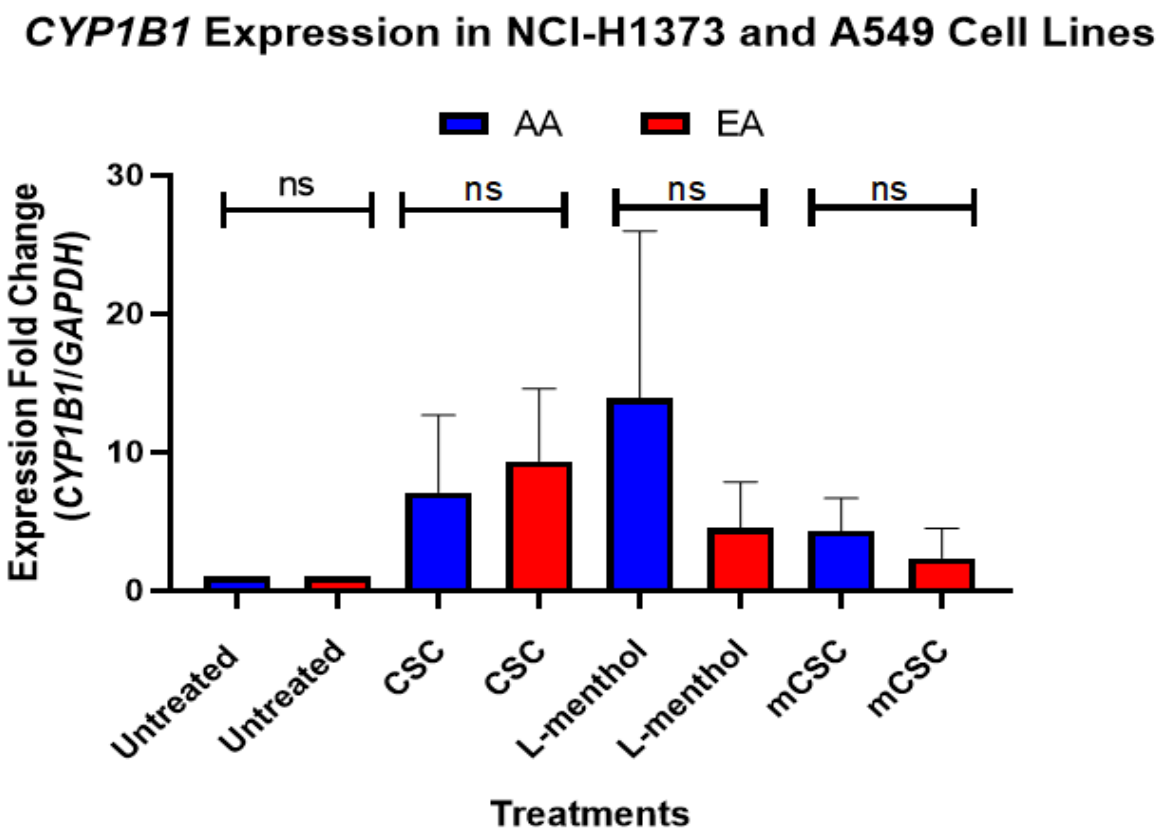


Figure 2. *CYP1B1* gene expression in LUAD cell lines after cigarette smoke condesate and menthol treatments. Quantitative reverse transcription PCR analysis showing relative *CYP1B1* gene expression levels after 6 hours of exposure to cigarette smoke condensate, L-menthol, and menthol cigarette smoke condesate. The error bars represent standard error from two biological replicates and three technical triplicates (ns = $P > 0.05$).

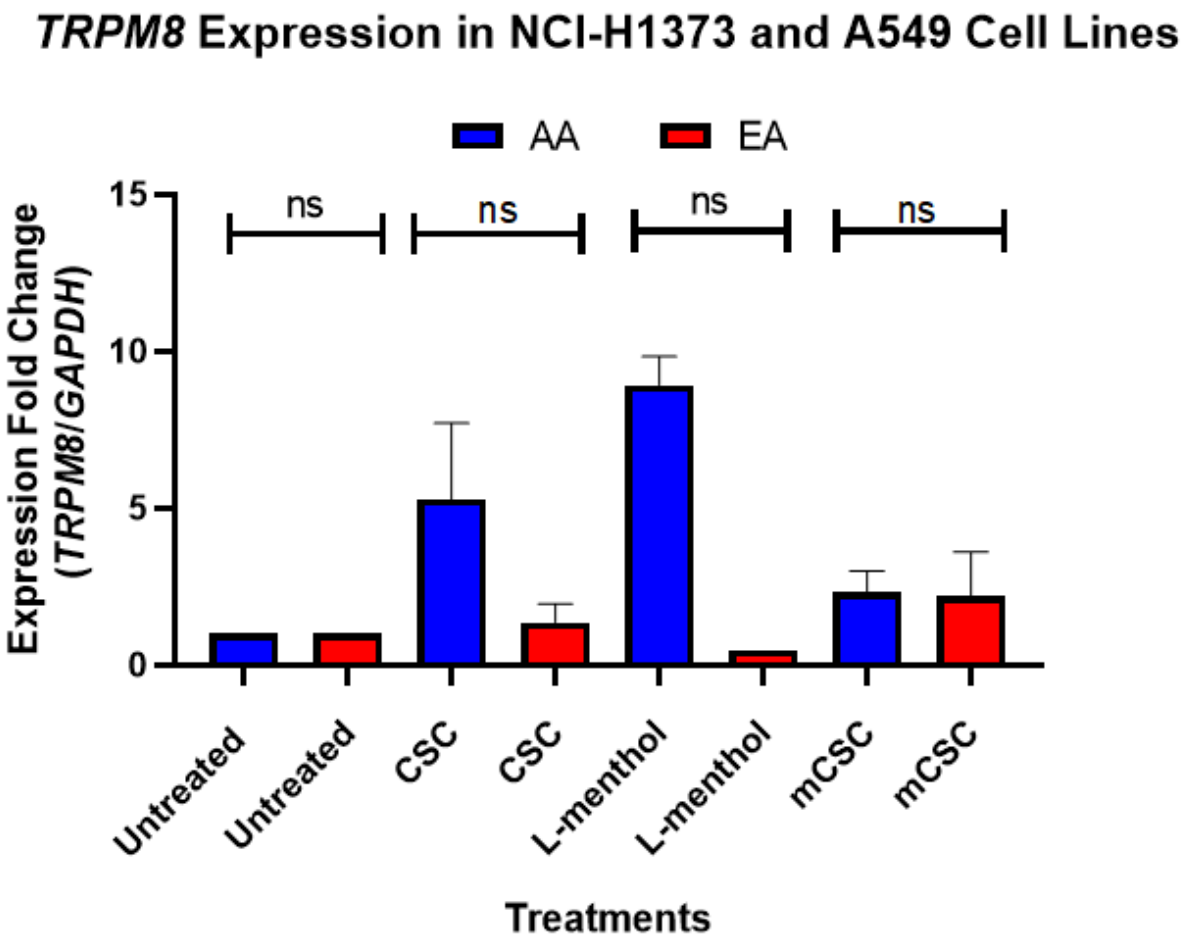


Figure 3. *TRPM8* gene expression in LUAD cell lines after cigarette smoke condesate and menthol treatments. Quantitative reverse transcription PCR analysis showing relative *TRPM8* gene expression levels after 6 hours of exposure to cigarette smoke condensate, L-menthol, and menthol cigarette smoke condesate. The error bars represent standard error from two biological replicates and three technical triplicates (ns = $P > 0.05$).

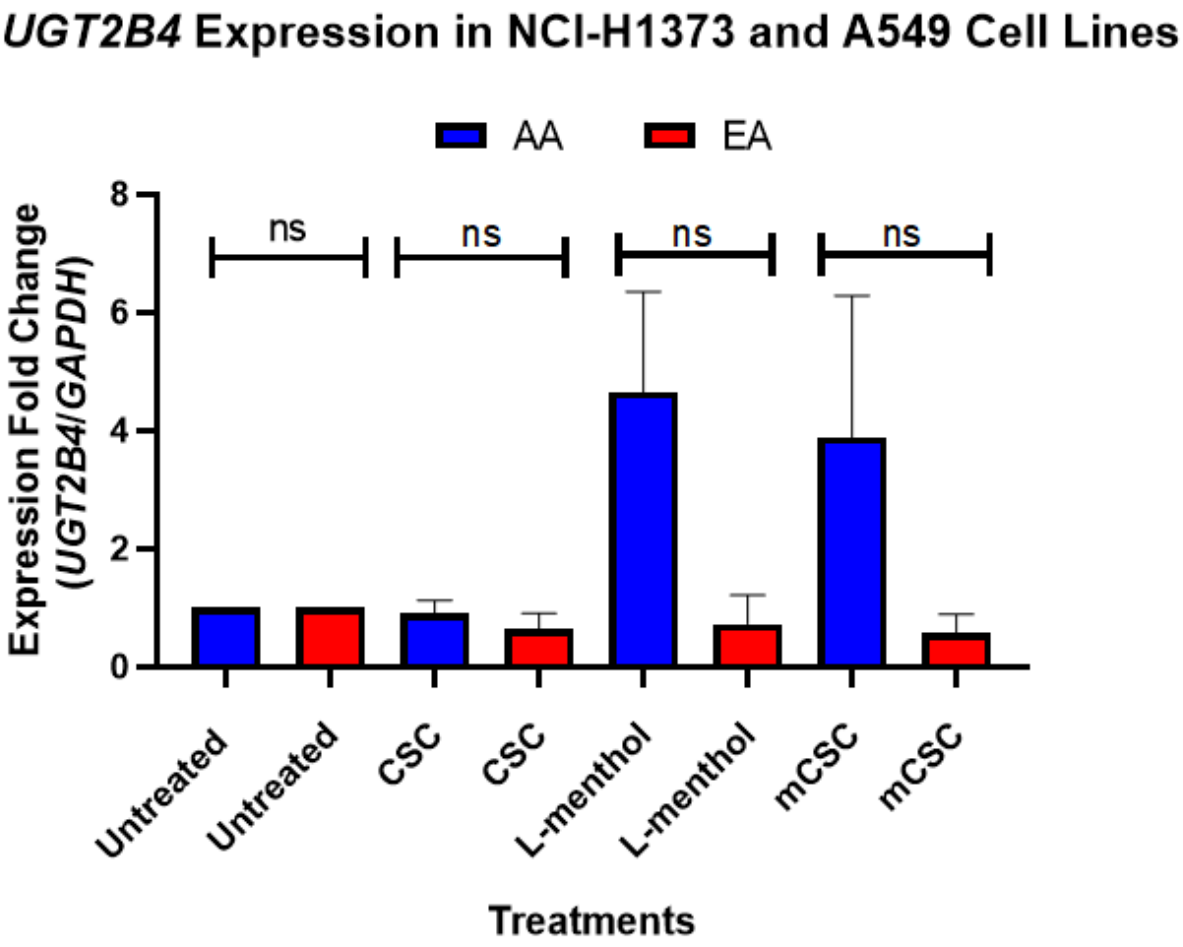


Figure 4. *UGT2B4* gene expression in LUAD cell lines after cigarette smoke condesate and menthol treatments. Quantitative reverse transcription PCR analysis showing relative *UGT2B4* gene expression levels after 6 hours of exposure to cigarette smoke condensate, L-menthol, and menthol cigarette smoke condesate. The error bars represent standard error from two biological replicates and three technical triplicates (ns = $P > 0.05$).

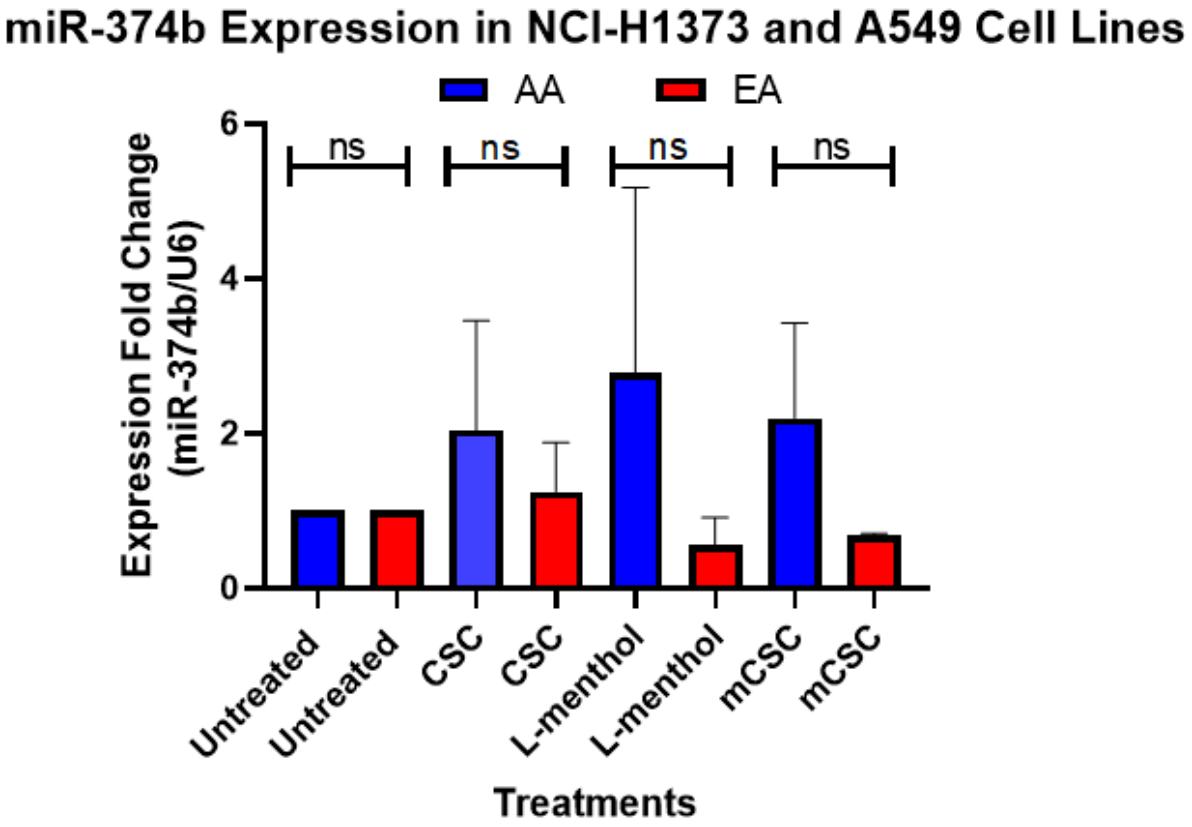


Figure 5. miR-374b expression in LUAD cell lines after cigarette smoke condesate and menthol treatments. Quantitative reverse transcription PCR analysis showing relative miR-374b expression levels after 6 hours of exposure to cigarette smoke condensate, L-menthol, and menthol cigarette smoke condesate. The error bars represent standard error from two biological replicates and three technical triplicates (ns = $P > 0.05$).

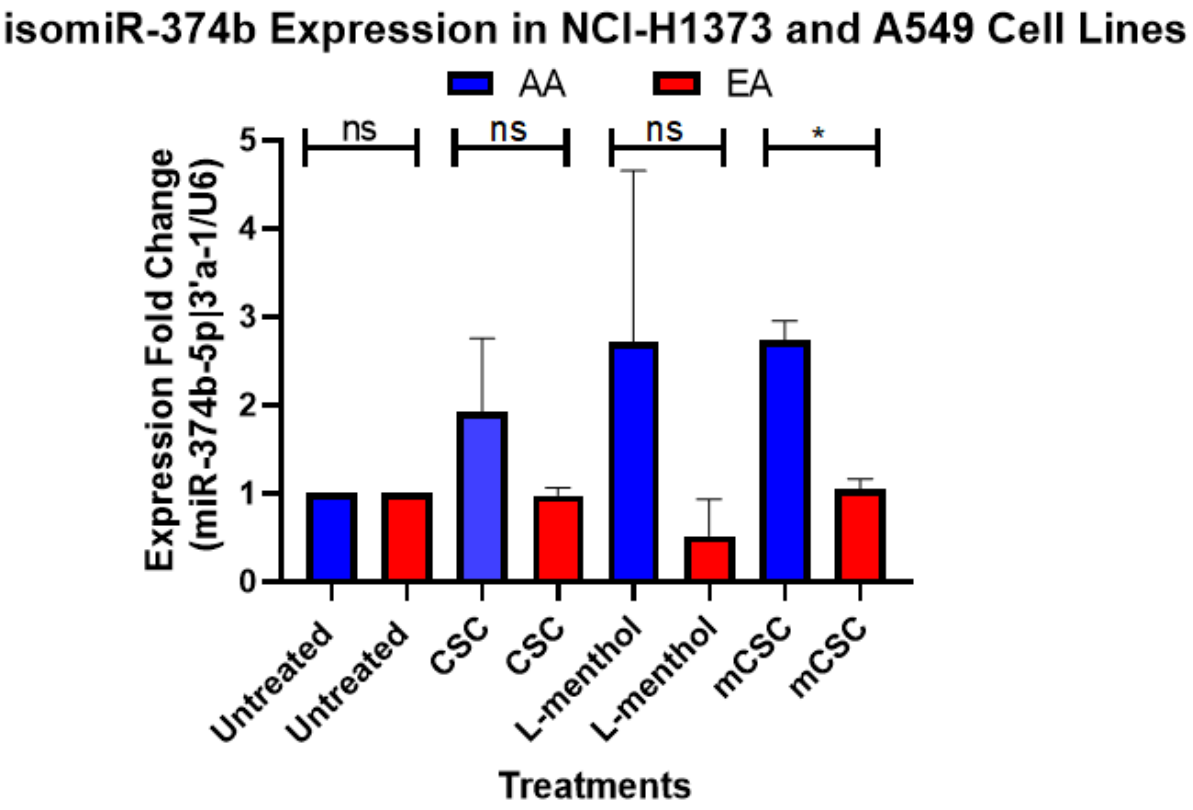


Figure 6. isomiR-374b expression in LUAD cell lines after cigarette smoke condesate and menthol treatments. Quantitative reverse transcription PCR analysis showing relative isomiR-374b expression levels after 6 hours of exposure to cigarette smoke condensate, L-menthol, and menthol cigarette smoke condesate. The error bars represent standard error from two biological replicates and three technical triplicates (ns = P > 0.05, s=<0.05).

References

1. Sun D, Wang X, Sui G, Chen S, Yu M, Zhang P. Downregulation of miR-374b-5p promotes chemotherapeutic resistance in pancreatic cancer by upregulating multiple anti-apoptotic proteins. Int J Oncol. 2018 May;52(5):1491-1503. doi: 10.3892/ijo.2018.4315. Epub 2018 Mar 14. PMID: 29568910; PMCID: PMC5873836.
2. Chang JT, Wang F, Chapin W, Huang RS. Identification of MicroRNAs as Breast Cancer Prognosis Markers through the Cancer Genome Atlas. PLoS One. 2016 Dec 13;11(12):e0168284. doi: 10.1371/journal.pone.0168284. PMID: 27959953; PMCID: PMC5154569.
3. Hanniford D, Zhong J, Koetz L, Gaziel-Sovran A, Lackaye DJ, Shang S, Pavlick A, Shapiro R, Berman R, Darvishian F, Shao Y, Osman I, Hernando E. A miRNA-Based Signature Detected in Primary Melanoma Tissue Predicts Development of Brain Metastasis. Clin Cancer Res. 2015 Nov 1;21(21):4903-12. doi: 10.1158/1078-0432.CCR-14-2566. Epub 2015 Jun 18. PMID: 26089374; PMCID: PMC4631639.
4. Li J, Zhang X, Tang J, Gong C. MicroRNA-374b-5p Functions as a Tumor Suppressor in Non-Small Cell Lung Cancer by Targeting FOXP1 and Predicts Prognosis of Cancer Patients. Onco Targets Ther. 2020 May 15;13:4229-4237. doi: 10.2147/OTT.S243221. PMID: 32523358; PMCID: PMC7237128.
5. Wang Y, Yu L, Wang T. MicroRNA-374b inhibits the tumor growth and promotes apoptosis in non-small cell lung cancer tissue through the p38/ERK signaling pathway by targeting JAM-2. J Thorac Dis. 2018 Sep;10(9):5489-5498. doi: 10.21037/jtd.2018.09.93. PMID: 30416798; PMCID: PMC6196188.
6. Xia J, Zhang W. A meta-analysis revealed insights into the sources, conservation and impact of microRNA 5'-isoforms in four model species. Nucleic Acids Res. 2014 Feb;42(3):1427-41. doi: 10.1093/nar/gkt967. Epub 2013 Oct 30. PMID: 24178030; PMCID: PMC3919606.
7. van Leeuwen DM, van Agen E, Gottschalk RW, Vlietinck R, Gielen M, van Herwijnen MH, Maas LM, Kleinjans JC, van Delft JH. Cigarette smoke-induced differential gene expression in blood cells from monozygotic twin pairs. Carcinogenesis. 2007 Mar;28(3):691-7. doi: 10.1093/carcin/bgl199. Epub 2006 Oct 19. PMID: 17056606.
8. Church, Timothy R et al. “Interaction of CYP1B1, cigarette-smoke carcinogen metabolism, and lung cancer risk.” International journal of molecular epidemiology and genetics vol. 1,4 295-309. 5 Aug. 2010
9. Jones NR, Lazarus P. UGT2B gene expression analysis in multiple tobacco carcinogen-targeted tissues. Drug Metab Dispos. 2014 Apr;42(4):529-36. doi: 10.1124/dmd.113.054718. Epub 2014 Jan 23. PMID: 24459179; PMCID: PMC3965906.
10. Dates CR, Fahmi T, Pyrek SJ, Yao-Borengasser A, Borowa-Mazgaj B, Bratton SM, Kadlubar SA, Mackenzie PI, Haun RS, Radominska-Pandya A. Human UDP-Glucuronosyltransferases: Effects of altered expression in breast and pancreatic cancer cell lines. Cancer Biol Ther. 2015;16(5):714-23. doi: 10.1080/15384047.2015.1026480. PMID: 25996841; PMCID: PMC4622877.
11. Wijayakumara DD, Mackenzie PI, McKinnon RA, Hu DG, Meech R. Regulation of UDP-Glucuronosyltransferases UGT2B4 and UGT2B7 by MicroRNAs in Liver Cancer Cells. J Pharmacol Exp Ther. 2017 Jun;361(3):386-397. doi: 10.1124/jpet.116.239707. Epub 2017 Apr 7. PMID: 28389526.

12. Nagaraj NS, Beckers S, Mensah JK, Waigel S, Vigneswaran N, Zacharias W. Cigarette smoke condensate induces cytochromes P450 and aldo-keto reductases in oral cancer cells. Toxicol Lett. 2006 Aug 20;165(2):182-94. doi: 10.1016/j.toxlet.2006.03.008. Epub 2006 Apr 18. PMID: 16713138; PMCID: PMC5774676.
13. Hellermann GR, Nagy SB, Kong X, Lockey RF, Mohapatra SS. Mechanism of cigarette smoke condensate-induced acute inflammatory response in human bronchial epithelial cells. Respir Res. 2002;3(1):22. doi: 10.1186/rr172. Epub 2002 Jul 10. PMID: 12204101; PMCID: PMC150508.
14. Fan L, Balakrishna S, Jabba SV, Bonner PE, Taylor SR, Picciotto MR, Jordt SE. Menthol decreases oral nicotine aversion in C57BL/6 mice through a TRPM8-dependent mechanism. Tob Control. 2016 Nov;25(Suppl 2):ii50-ii54. doi: 10.1136/tobaccocontrol-2016-053209. Epub 2016 Oct 3. PMID: 27698211; PMCID: PMC5496986.
15. Ai J, Taylor KM, Lisko JG, Tran H, Watson CH, Holman MR. Menthol Content in US Marketed Cigarettes. Nicotine Tob Res. 2016 Jul;18(7):1575-80. doi: 10.1093/ntr/ntv162. Epub 2015 Aug 9. PMID: 26259988; PMCID: PMC4747842.
16. Nagathihalli NS, Massion PP, Gonzalez AL, Lu P, Datta PK. Smoking induces epithelial-to-mesenchymal transition in non-small cell lung cancer through HDAC-mediated downregulation of E-cadherin. Mol Cancer Ther. 2012 Nov;11(11):2362-72. doi: 10.1158/1535-7163.MCT-12-0107. Epub 2012 Aug 29. PMID: 22933707; PMCID: PMC3969334.
17. Liu B, Fan L, Balakrishna S, Sui A, Morris JB, Jordt SE. TRPM8 is the principal mediator of menthol-induced analgesia of acute and inflammatory pain. Pain. 2013 Oct;154(10):2169-2177. doi: 10.1016/j.pain.2013.06.043. Epub 2013 Jun 29. PMID: 23820004; PMCID: PMC3778045.
18. Parashar G, Parashar NCh, Capalash N. – (-) Menthol Induces Reversal of Promoter Hypermethylation and Associated Up-Regulation of the FANCF Gene in the SiHa Cell Line. Asian Pac J Cancer Prev. 2017 May 1;18(5):1365-1370. doi: 10.22034/APJCP.2017.18.5.1365. PMID: 28612587; PMCID: PMC5555548.
19. Xi S, Inchauste S, Guo H, Shan J, Xiao Z, Xu H, Miettenen M, Zhang MR, Hong JA, Raiji MT, Altorki NK, Casson AG, Beer DG, Robles AI, Bowman ED, Harris CC, Steinberg SM, Schrump DS. Cigarette smoke mediates epigenetic repression of miR-217 during esophageal adenocarcinogenesis. Oncogene. 2015 Oct 29;34(44):5548-59. doi: 10.1038/onc.2015.10. Epub 2015 Feb 23. PMID: 25703328; PMCID: PMC6301032.
20. Luo B, Chen C, Wu X, Yan D, Chen F, Yu X, Yuan J. Cytochrome P450 2U1 Is a Novel Independent Prognostic Biomarker in Breast Cancer Patients. Front Oncol. 2020 Aug 5;10:1379. doi: 10.3389/fonc.2020.01379. PMID: 32850442; PMCID: PMC7419690.
21. Alonso-Trujillo M, Muñiz-González AB, Martínez-Guitarte JL. Endosulfan exposure alters transcription of genes involved in the detoxification and stress responses in Physella acuta. Sci Rep. 2020 May 12;10(1):7847. doi: 10.1038/s41598-020-64554-8. PMID: 32398709; PMCID: PMC7217849.

VII. Conclusions

In chapter 1, population-specific *CYP1B1* and *UGT2B4* mRNA expression were associated with isomiR-374b expression in LUAD patients. AA LUAD patients had lower MME and higher isomiR expression compared to EAs. A subset of AA patients had a positive isomiR-mRNA relationship. These expression patterns were solely due to race, unaffected by age and smoking factors. The candidate isomiR, miR-374b-5p|3'a-1, was specific to AAs.

Chapter 2 provided mechanistic effects of menthol exposure on lung cancer cells from AA and EA patients. To my knowledge, it is the first lung cancer study that has utilized menthol cigarette smoke condensate to assess racial differences in MME and isomiR expression. Similar to chapter 1, I found that miR-374b-5p|3'a-1 was specific to AAs. This expression was not associated with *CYP1B1* and *UGT2B4* levels, but maybe associated with *CYP2U1*. If high isomiR-374b abundance is associated with low *CYP2U1* expression, this population-specific transcriptomic change has the possibility to be a novel therapeutic option for lung cancer patients. A subset of LUAD patients had a high isomiR-374b expression signature, but the majority had low expression. By adopting a precision medicine approach and developing an FDA-approved targeted therapy to upregulate isomiR-374b expression.

VIII. Acknowledgments

I would like to thank the Lafayette College Department of Biology for the funding and support for my honors thesis project. I also want to thank my committee members, Dr. Elaine Reynolds in the biology department and Dr. Daniel Griffith in the chemistry department, for dedicating their time, knowledge, and unwavering encouragement. Thank you to my family, friends, and supporters who have given me unconditional wisdom and love. Lastly, I would like to thank my mentor, Dr. Khadijah A. Mitchell, for her endless support, guidance, and effort to make this project and my goals a reality.