Exploring *PDE4* Somatic Copy Number Variations, Gene Expression, and Induced Overexpression in Lung Cancer Health Disparities

Proposal highlights

- Lung cancer is the leading cause of cancer-related death in the United States, and it disproportionately affects African American (AA) males.
- The *PDE4* family as a whole is upregulated in a number of cancers. Specifically, *PDE4D* overexpression has been linked to sex and racial disparities in lung cancer.
- Population-specific differences in lung cancer biology may be the result of somatic copy number variation (SCNV) and alternative splicing events in *PDE4* genes.
- The *PDE4* family promotes EMT (epithelial–mesenchymal transition), a process associated with cancer initiation, invasion, metastasis, and resistance to therapy, in lung epithelial cells.
- PDE4 inhibitors, previously employed to treat the lung condition COPD, have seen success in reduction of lung carcinogenesis in murine models.
- Some anticancer drugs vary in efficacy and safety by sex and African ancestry, suggesting AA males may have clinically relevant differences in PDE4 expression patterns than EA males.
- AAs and other racial minorities are underrepresented in cancer drug clinical trials, underscoring the need for more research in this population.
- This honors thesis explored the relationship between *PDE4* SCNV profiles, gene and isoform expression, and experimentally induced overexpression response by sex and
- race. This study is significant for precision medicine purposes to ensure AA male lung cancer patients benefit from treatment advances to the same extent as their EA counterparts.

Joelle Rabin-Court
Committee: Drs. Robert Kurt, Eric Ho and Zoe Boekelheide
Advisor: Dr. Khadijah A. Mitchell
Department of Biology, Lafayette College

TABLE OF CONTENTS

I. BIOGRAPHICAL SKETCH

I am Joelle Rabin-Court, a senior at Lafayette College. I will obtain my Bachelor of Science degree in Biology in June 2021. This is my third year doing research with Dr. Khadijah A. Mitchell. I first learned about her research after her TedxLafayette talk in the Spring of 2018. I was eager to learn more about precision medicine and contribute to cancer health disparities research. I started my research with her the following fall, focusing on microRNA determinants of lung cancer sex disparities.

I was first introduced to health disparities research as a volunteer in the Maternal Fetal Medicine wing of Mount Sinai West. I spent the summer pulling charts and inputting data for projects investigating age-related response to Pitocin in labor induction and racial disparities in the treatment of severe hypertension with labetalol. When I returned to campus, I began working in Dr. Mitchell's Integrative and Translational Laboratory for Applied Biology--first as an Independent Study student, then as an EXCEL Scholar and now as an Honors Thesis student. This work has cultivated a deep interest in public health within me. Last year, I ran a campus-wide anti-vaping public health campaign at the peak of e-cigarette and vaping associated lung injury (EVALI) reports. This year, amidst the pandemic, I pursued a contact tracing certification and joined the NARAP's COVID-19 Potential Tracer Registry. I also participated in a virtual externship with a public health nurse.

After graduation, I will work as either a research support assistant or a clinical care coordinator, prior to applying to MD-MPH programs.

II. ABSTRACT

Background: Lung cancer is the second most common cancer and the leading cause of cancer-related death in the US, and disproportionately affects African American (AA) males. Smoking is the strongest risk factor for lung cancer development and negative treatment outcomes. Overall, AAs smoke less than European Americans (EAs), and even in never smokers, suggesting biological determinants. Somatic Copy Number Variations (SCNVs), the most common structural variation in the human genome, have been linked to a number of cancers. Recent studies show potential *PDE4D* involvement in lung cancer health disparities involving males with high West African ancestry. Additionally, alternative splicing contributes to disease progression in AA males with prostate cancer. Similar isoform and cancer progression relationships may also exist in AA males with lung cancer. No study to date has integrated *PDE4* SCNVs, gene and isoform expression profiles, and patient survival by race and sex.

Hypothesis: *PDE4* SCNVs, gene and isoform profiles vary significantly between AA and EA males with lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), with survival implications and an association with epithelial-mesenchymal transition (EMT).

Methods: Downloaded *PDE4* SCNV (Broad GDAC Firehose) and mRNA and isoform sequencing data (TSVdb) for the TCGA cohort. Merged (LUAD n = 23 AA and 170 EA males; LUSC n = 16 AA and 254 EA males) and performed differential analyses using two-tailed t-tests with Welch corrections (GraphPad Prism 8). Compared disease-free survival (cBioPortal) based on SCNV, *PDE4* gene expression and isoform expression status. Performed plasmid overexpression in AA (NCI-H1373) and EA (A549) male LUAD cell lines and measured expression of EMT phenotypic markers, *VIM* and *ECAD* via qRT-PCR.

Results: *PDE4C* demonstrated significant differential SCNV (P<0.05) and isoform expression (P≤0.01) by sex and race across both histologies. Fewer amplifications and greater deletions were observed in AA males, relative to EA males, corresponded with lower *PDE4C* gene and isoform (uc002nii, uc002nil, uc010xqc, uc010ebk, and uc010ebm) expression. In LUAD patients, high *PDE4C* expression correlated with greater median survival, particularly amongst AA males (~62 months vs ~30 months for AA high/low expressors, compared with ~38 months and ~28 months for EA high/low expressors, respectively). In LUAD cells, *PDE4C* overexpression was not associated with the EMT phenotype, regardless of race.

Conclusions: AA and EA *PDE4C* SCNV, gene and isoform signatures suggest racial differences in lung tumor biology. However, *PDE4C* overexpression's effect on the EMT phenotype does not drive trends in patient survival.

III. INTRODUCTION

A. Lung cancer incidence and mortality statistics by sex, race and histology in the US

Lung cancer is the second most commonly diagnosed cancer, with an estimated 235,760 (12.4%) new cases in 2021.[1] It is also leading cause of cancer-related death for both males and females in the United States (US), accounting for an estimated 131,880 (21.7%) deaths in 2021.[1] According to the American Cancer Society, non-small cell lung cancer (NSCLC) is the most common type of lung cancer, making up approximately 85% of cases.[2] Within NSCLC histologies, lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) are the two most prevalent subtypes.[2] LUAD is the most common subtype among non-smokers with characteristic mutation profiles, whereas LUSC is tightly linked to smoking and secondhand smoke exposure.[2]

Across all populations, African American (AA) males have the highest rates of lung cancer incidence and mortality when compared to any other sex and race group.[3] From 2013-2017, 77.6 AA males per 100,000 were diagnosed with lung cancer; of which, approximately 31.0 were diagnosed with LUAD and 20.0 with LUSC.[3,4] Comparatively, 67.4 European American (EA) males per 100,000 were diagnosed with lung cancer, 23.2 with LUAD and 14.0 with LUSC.[3,4] Over the same time frame, 58.8 AA males per 100,000 died from lung cancer, compared to 49.4 EA males respectively.[3]

B. Behavioral and biological determinants of non-small cell lung cancer (NSCLC)

Many different factors, both behavioral and biological, contribute to lung cancer development. While exposure to radon gas, secondhand smoke, asbestos and air pollution, as well as family history, and inflammatory-based respiratory diseases like chronic obstructive pulmonary disease (COPD) all affect lung cancer risk, tobacco-smoking is by far the strongest

risk factor for lung cancer development.[5] Furthermore, smoking after cancer diagnosis is associated with more negative outcomes: increased treatment toxicity, greater risk of treatment failure, higher incidence of second tumors and shorter survival.[6] However, observed racial disparities cannot be explained by cigarette smoking alone. AAs smoke significantly less than EAs, not just in cigarettes per day, but also per lifetime.[7] Even AA never smokers have higher rates of lung cancer when compared with EAs.[7] This suggests potential biological determinants, such as genetics. One gene is *CHRNA5*.[8] *CHRNA5*, which encodes the α5 subunit of neuronal nicotinic acetylcholine receptors, was found to be upregulated 30-fold in LUAD compared to normal lung tissue.[8] In total, 18 lung cancer susceptibility loci have been identified, 10 of which recently discovered by McKay and colleagues.[9] Their study's sample, while very large, only included individuals of European descent, further underscoring the importance of this investigation.[9]

C. PDE4 somatic copy number variations (SCNVs) as emerging biological determinants of NSCLC health disparities

Another potential biological determinant could be somatic copy number variation (SCNV). SCNVs are the most common structural variation in the human genome and range at least 1 kb in length.[10] Most significantly, SCNVs have been linked to a number of cancers.[10] SCNV amplifications are vital in activating oncogenes and SCNV deletions play a role in inactivating tumor suppressor genes.[10]

Cyclic nucleotide phosphodiesterases (PDE) are metallohydrolyases that bind and degrade the intracellular secondary messengers cAMP and cGMP.[11] The PDE4 subfamily, in particular, has been implicated in a number of inflammatory diseases and malignancies, including COPD, which doubles the likelihood that an individual will develop lung cancer.[11–13] In

fact, the condition are more commonly seen together than apart.[13] A recent study has shown *PDE4D* genetic changes and gene products may be involved in lung cancer health disparities, particularly in AA males.[14] Therefore, SCNVs of *PDE4* genes may be driving gene expression differences.

D. Sex and racial differences in NSCLC treatment

1. Sex and racial differences in anticancer drug response

NSCLC patients can be treated with either surgery, radiation therapy, chemotherapy (drugs that prevent any active cell division), targeted therapy (drugs that target specific molecular changes in tumors), immunotherapy (drugs that enhance the immune system's ability to fight cancer), or a combination of the above.[15] Prior studies with anticancer drugs (i.e. bevacizumab, cisplatin, paclitaxel, doxorubicin) have found differential drug efficacy and safety by sex and race.[16,17] Doxorubicin, in particular, has demonstrated increased cardiotoxicity in African Americans[17] and females,[18] as well as increased clearance in males.[18]

2. African American underrepresentation in cancer clinical trials

Many cancer clinical trials lack adequate minority representation, which decreases the applicability of their results while increasing the likelihood that population-specific differences in drug efficacy will go undiscovered. Regardless of cancer type, AAs are less likely to be enrolled in clinical trials compared to EAs.[19] In lung cancer clinical trials, specifically, AAs constitute 5.4% of trial enrollees (compared to 10.6% of 2013 cancer prevalence), whereas EAs constitute 79.8% (compared to 82.7% of 2013 cancer prevalence).[19] Due to AA underrepresentation in lung cancer clinical trials, it is unclear if a differential population response exists for anticancer drugs until long after they are on the market. More work is needed in this area to close the knowledge gap.

3. Potential difference in PDE4 inhibitor response by sex and race

Recently, PDE4 inhibitors (PDE4i) have demonstrated great clinical promise. They are FDA-approved to treat COPD, as well as asthma and arthritis.[20] They have also been shown to attenuate lung carcinogenesis and tumor angiogenesis in murine models.[21] However, there have been no clinical studies done looking at the impact of the inhibitor in lung cancer. And, thus far, subtype selective inhibitors have only been successfully developed for PDE4B and PDE4D.[11] Further development of PDE4i for lung cancer treatment must account for potential population differences in expression by sex and race.

F. References

1. American Cancer Society. Cancer Facts & Figures 2020. *Am. Cancer Soc.* 12 (2020).
2. Herbst, R. S., Morgensztern, D. & Boshoff, C. The biology and management of non-small cell lung cancer. *Nature* **553**, 446–454 (2018).
3. US Cancer Statistics Working Group. US Cancer Statistics Data Visualizations Tool, based on 2019 submission data (!999-2017). *US Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute* https://gis.cdc.gov/cancer/uscs/dataviz.html (2020).
4. Surveillance Research Program, National Cancer Institute. SEER*Explorer: An interactive website for SEER cancer statistics. https://seer.cancer.gov/explorer/ (2020).
5. Schwartz, A. G. & Cote, M. L. Epidemiology of Lung Cancer. in *Lung Cancer and Personalized Medicine: Current Knowledge and Therapies* (eds. Ahmad, A. & Gadgeel, S.) 21–41 (Springer International Publishing, 2016). doi:10.1007/978-3-319-24223-1_2.
6. Jassem, J. Tobacco smoking after diagnosis of cancer: clinical aspects. *Transl. Lung Cancer Res.* **8**, S50–S58 (2019).
7. Ryan, B. M. Lung cancer health disparities. *Carcinogenesis* **39**, 741–751 (2018).
8. Falvella, F. S. *et al.* Multiple isoforms and differential allelic expression of CHRNA5 in lung tissue and lung adenocarcinoma. *Carcinogenesis* **34**, 1281–1285 (2013).
9. McKay, J. D. *et al.* Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat. Genet.* **49**, 1126–1132 (2017).
10. Shao, X. *et al.* Copy number variation is highly correlated with differential gene expression: a pan-cancer study. *BMC Med. Genet.* **20**, 175 (2019).
11. Hsien Lai, S., Zervoudakis, G., Chou, J., Gurney, M. E. & Quesnelle, K. M. PDE4 subtypes in cancer. *Oncogene* **39**, 3791–3802 (2020).
12. Lin, D.-C. *et al.* Genomic and functional characterizations of phosphodiesterase subtype 4D in human cancers. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6109–6114 (2013).
13. Seijo, L. M. & Zulueta, J. J. Understanding the Links Between Lung Cancer, COPD, and Emphysema: A Key to More Effective Treatment and Screening. *Oncol. Williston Park N* **31**, 93–102 (2017).

14. Mitchell, K. A. *et al.* Relationship between West African ancestry with lung cancer risk and survival in African Americans. *Cancer Causes Control CCC* **30**, 1259–1268 (2019).
15. Duma, N., Santana-Davila, R. & Molina, J. R. Non-Small Cell Lung Cancer: Epidemiology, Screening, Diagnosis, and Treatment. *Mayo Clin. Proc.* **94**, 1623–1640 (2019).
16. Huang, R. S., Kistner, E. O., Bleibel, W. K., Shukla, S. J. & Dolan, M. E. Effect of population and gender on chemotherapeutic agent–induced cytotoxicity. *Mol. Cancer Ther.* **6**, 31–36 (2007).
17. O'Donnell, P. H. & Dolan, M. E. Cancer Pharmacoethnicity: Ethnic Differences in Susceptibility to the Effects of Chemotherapy. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **15**, 4806–4814 (2009).
18. Kim, H.-I., Lim, H. & Moon, A. Sex Differences in Cancer: Epidemiology, Genetics and Therapy. *Biomol. Ther.* **26**, 335–342 (2018).
19. Duma, N. *et al.* Representation of Minorities and Women in Oncology Clinical Trials: Review of the Past 14 Years. *J. Oncol. Pract.* **14**, e1–e10 (2018).
20. Li, H., Zuo, J. & Tang, W. Phosphodiesterase-4 Inhibitors for the Treatment of Inflammatory Diseases. *Front. Pharmacol.* **9**, (2018).
21. Yeo, C. D. *et al.* Roflumilast treatment inhibits lung carcinogenesis in benzo(a)pyrene-induced murine lung cancer model. *Eur. J. Pharmacol.* **812**, 189–195 (2017).

IV. CHAPTER 1: Exploring The Relationship Between *PDE4* SCNV Profiles, Gene and Isoform Expression in Lung Cancer Sex and Racial Disparities

A. Introduction

1. PDE4 genes associated with with lung cancer

A number of studies have investigated *PDE4* genes in relation to lung cancer. The majority of which have uncovered an oncogenic role, marked by overexpression in cancer cells and tissues. Interestingly, more than one member of the gene family was found to be involved in epithelial-mesenchymal transition (EMT), a key event in cancer development that enables invasion and metastasis.[1] EMT is marked by the loss of e-cadherin (gene name, *CDH1*), a junction protein known for its role in cell-cell adhesion, and the acquisition of vimentin (gene name, *VIM*), known for its role in cell migration.

PDE4A was found to be upregulated in lung cancer cell lines.[2] One study showed that (TGF)-β1-mediated *PDE4A* upregulation induces EMT in alveolar epithelial cells.[1] Another study demonstrated that hypoxia-induced *PDE4A* upregulation results in increased tumor cell proliferation.[3] *PDE4B* was also upregulated specifically in NSCLC tissues.[4] In fact, *PDE4B* inhibition with 1,2,3-triazole derivatives had cytotoxic effects on A549 lung adenocarcinoma cells.[5] *PDE4C* was similarly found to be overexpressed in lung cancer tissues, particularly in lung adenocarcinoma samples.[4] Though the least commonly studied in the family, one study concluded that *PDE4C* might even be a novel target of a mutant p53 transcription factor.[2,6] *PDE4D*, the most widely examined of the family, was also found to be upregulated in lung cancer cells.[2] It, like *PDE4A*, was linked to the EMT pathway, tumor cell proliferation and differentiation, and thus lung cancer development and progression.[1] Correspondingly, PDE4i treatment was found to reduce vascular endothelial growth factor (VEGF) secretion, angiogenesis, cell proliferation and growth in human lung tumors.[3]

2. PDE4 SCNVs may influence gene and isoform expression

There is a strong relationship between SCNVs and differential gene expression in NSCLC.[7] As such, SCNVs may drive differences in *PDE4* gene and isoform expression, and thus disease profiles. No study to date has comprehensively evaluated the role of *PDE4* SCNVs in lung cancer associated tumor suppressors or oncogenes, though homozygous deletions of *PDE4D* have led to speculation that the gene may act as a tumor suppressor in esophageal adenocarcinoma.[8]

3. Alternative splicing events confer drug resistance and patient survival

In a number of solid cancers, alternative splicing events contribute to racial disparities in disease progression.[9] Alternative splicing events are when multiple protein isoforms are produced from a single gene. Each isoform may function differently, and an isoform switch occurs when the relative contribution of isoforms to their parent gene expression changes significantly between conditions.[9] Specifically, in prostate cancer, high expression of short isoform PIK3CD-S in AA males is linked to cancer aggressiveness and drug resistance.[10] Comparable *PDE4* isoform and cancer progression relationships may exist in a lung cancer health disparities context as well.

4. Hypothesis

To our knowledge, no study to date has characterized the *PDE4* SCNV profiles, gene and isoform expression, and survival by sex and race. I hypothesize that *PDE4* SCNV, gene and isoform profiles vary significantly between AA and EA males with LUAD and LUSC, with survival implications.

B. Methods

1. Clinical data extraction and statistical analysis for TCGA patient cohort

The Cancer Genome Atlas (TCGA) SCNV and mRNA sequencing data were downloaded from Broad GDAC Firehose (https://gdac.broadinstitute.org/) and annotated. TCGA mRNA isoform sequencing data were downloaded from the TCGA Splicing Variants Database (http://tsvdb.com/plot.html) and filtered by sex and race for 23 AA males, 207 EA males, 30 AA females and 241 EA females with LUAD, as well as 16 AA males, 254 EA males, 14 AA females and 93 EA females with LUSC. Fisher's exact t test, chi square test, and t test were performed to identify confounding clinical and demographic variables. Paired two-tailed t-tests in GraphPad Prism 9.0 were used to compare interactions between race, age and smoking status in LUAD patients, and race, sex and stage in LUSC patients.

2. SCNV and expression analysis for PDE4 genes based on sex and race, by histology in TCGA patient cohort, not stratified by amplification or deletion status.

Performed differential analyses for all *PDE4* genes according to SCNV, mRNA, and mRNA isoform expression status in the LUAD and LUSC TCGA sub cohorts. GraphPad Prism 9.0 was used to carry out two-tailed t-tests with Welch corrections.

3. SCNV and expression analysis for PDE4 genes based on sex and race, by histology in TCGA patient cohort, stratified by amplification or deletion status.

Merged all SCNV, gene and isoform data for candidate *PDE4* gene of interest. Performed differential analysis of gene and isoform expression for gene of interest, stratified by SCNV status (amplification or deletion) and histology for TCGA patient cohort in GraphPad Prism 9.0 using two-tailed t-test with Welch corrections.

4. Disease-specific survival analysis based on PDE4 gene expression.

a) By SCNV status

AA and EA patients were stratified by SCNV status into groups with amplifications and deletions. cBioPortal (https://www.cbioportal.org/) was used to obtain Kaplan-Meier five-year disease-free survival curves for AAs and EAs patients with amplification and deletion, respectively. Median months disease-free values were also collected.

b) By expression status

AA and EA patients were stratified by median candidate gene expression into either low- or high-expressors. cBioPortal was used to obtain Kaplan-Meier five-year disease-free survival curves for AA and EA low-expressing and high-expressing patients, respectively. Median months disease-free values were also collected.

C. Results

There are significant differences in the patient clinical characteristics by age and vital status in the LUAD subgroup, and by sex and vital status in the LUSC subgroup of the TCGA patient cohort.

In LUAD patients, there was significant variation with regards to age (P=0.0001) and vital status (P=0.0001) (Table 1). Two-tailed t-test patient race and age showed no significant interactions between these characteristics (P>0.05, Fig. 1). Our results are in line with previous studies have shown that AAs, on average, are diagnosed with lung cancer at a younger age.[11] In LUSC patients, there was significant variation with regards to sex (P=0.0052) and vital status (P=0.0006) (Table 1). Interaction analyses were not performed by sex in LUSC patients as this variable was controlled for throughout the study.

Of the four PDE4 genes, only PDE4C demonstrated significant differential SCNV and expression by sex and race.

Analysis of Table 2 found significant differences in *PDE4A* SCNV between AA and EA males with LUSC, but no corresponding differences in gene or isoform expression. In contrast, Table 3 found significant differences by race in *PDE4B* gene and isoform expression, but not SCNV, in male patients with LUSC. Notably, Table 4 revealed significant differences in *PDE4C* SCNV, and gene and isoform expression by race in male patients with both NSCLC histologies. In LUAD patients, there were sex-specific racial differences in *PDE4C* gene and isoform expression, specifically with regards to isoforms uc002nii and uc002nil (Table 4). In LUSC patients, there were just sex-specific racial differences in isoform expression, specifically isoforms uc010xqc, uc010ebk and uc010ebm (Table 4). For the aforementioned gene and isoforms, AAs consistently had lower mean expression than their EA counterparts (Table 4). Lastly, Table 5 found significant differences by race in *PDE4D* gene and isoform expression, but not in SCNV, in male patients with LUAD.

PDE4C amplifications may drive gene and isoform uc002nil expression in EA males with LUAD, but not LUSC.

In LUAD patients, significantly higher amplification of *PDE4C* (Fig. 2a) in EAs compared to AAs corresponded with significantly higher *PDE4C* (Fig. 2b) and isoform uc002nil (Fig. 2d) expression. There was a greater frequency of *PDE4C* deletions in AAs (87% AA vs 70% EA), and a lower frequency of *PDE4C* amplifications (13% AA vs 30% EA). However, the median *PDE4C* deletions were similar in both races (Fig. 3a). Interestingly, the median gene expression was lower in AA males (Fig. 3b), potentially driven by significantly decreased

uc002nil expression (Fig. 3d). Isoform uc002nii had extremely low expression in both populations, suggesting minimal biological relevance (Fig. 3c).

Similar to LUAD findings, there were fewer *PDE4C* amplifications in AA LUSC patients (19% AA vs 42% EA) and a greater frequency of *PDE4C* deletions (81% AA vs 58% EA). There was no significant racial difference in amplifications, however, this trend was observed in LUSC patients with *PDE4C* deletions (Fig. 4a and Fig. 5a). Significant differences in *PDE4C* deletions and gene isoform expression were observed in AAs and EAs. However, most of the patients have mRNA isoform sequencing reads of less than five, which is suspected to be biologically irrelevant (Fig. 4b-d and Fig. 5b-d).[12] Because of this low abundance, *PDE4C* expression is not a useful biomarker to stratify LUSC patients, but may have clinical utility for LUAD patients.

PDE4C amplification and high expression is associated with improved survival for EA males, and PDE4C high expression is also linked to better survival in AA males.

LUAD patients with *PDE4C* amplifications have greater survival than patients with *PDE4C* deletions (Fig. 6a-b). This is more evident in EAs, because they have a greater proportion of patients with amplifications (Fig. 6a-b). EAs with amplifications had a median survival more than two times as long as EAs with deletions (Table 6). The same relationship could not be determined based on AA SCNV status, due to the limited number of AA LUAD patients with amplifications (Table 6).

LUAD patients with high *PDE4C* expression had greater survival than patients with low *PDE4C* expression (Fig. 7), regardless of race. AAs with high *PDE4C* expression have a median survival over 2 times greater than AAs with low *PDE4C* expression (Table 7). EAs with high *PDE4C* expression have a median survival nearly 1.5 times greater than EAs with low *PDE4C*

expression (Table 7). Similarly, high expression of the biologically relevant *PDE4C* isoform (uc002nil, Fig. 2d) is associated with more than double the median survival in AA LUAD patients (Fig. 8, Table 7) and nearly two-fold greater median survival in EA LUAD patients (Fig. 8, Table 7).

D. Discussion

When focusing on amplifications, 30% of EA LUAD patients had *PDE4C* amplifications, compared to 13% of AA patients. Patients with *PDE4C* amplifications are more likely to have higher *PDE4C* mRNA and mRNA isoform expression, which is associated with longer patient survival. When focusing on deletions, 70% EA LUAD patients had *PDE4C* deletions, compared with 87% of AA patients. These deletions were associated with decreased mRNA and mRNA isoform expression, and shorter patient survival. On the other hand, LUSC patients are unlikely to have *PDE4C* SCNVs affecting *PDE4C* gene and isoform expression in a substantial way. *PDE4C* SCNVs do not appear to be involved in LUSC racial disparities.

Compared with other *PDE4* family members, *PDE4C* is a relatively overlooked precision medicine target. Both EA and AA patients with LUAD could benefit from *PDE4C* replacement therapy, or a targeted therapy that would upregulate *PDE4C*. This effect would be more pronounced in AAs, who are less likely to have the *PDE4C* amplifications and more likely to have *PDE4C* deletions, and associated expression changes, thus reducing the disparity. Interestingly, this is a different pattern from other *PDE4* family members. *PDE4A*, *B,* and *D* overexpression is associated with poor survival outcomes.[1,2,4] Further research on the mechanistic effect of *PDE4C* overexpression will provide insight into its role in lung cancer sex and race disparities, as well as future treatment options. A limitation to this study is that few AA males

with LUAD had *PDE4C* amplifications. Additionally, there was limited SCNV data available for LUAD cell lines from AAs.

E. Tables and Figures

Table 1: Clinical and Demographic Characteristics of Study Subjects with NSCLC in the TCGA Cohort.

Table 1. Clinical and Demographic Characteristics of Study Subjects in the TCGA Cohort.						
	LUAD			**LUSC**		
	AA (n = 53)	**EA (n = 448)**	***P***	**AA (n = 30)**	**EA (n = 347)**	***P***
Age (years) #			*0.0001*			*0.5630*
Mean (SD)	59.8 (±10)	66.0 (±10)		67.9 (±8)	67.0 (±9)	
Range	39-80	38-88		52-83	40-85	
Sex (%) ^			*0.7761*			*0.0052*
Female	30 (57)	241 (54)		14 (47)	93 (27)	
Male	23 (43)	207 (46)		16 (53)	254 (73)	
Stage (%) ^ *			*0.8603*			*0.0554*
Early (I-II)	41 (77)	349 (78)		22 (73)	293 (84)	
Late (III-IV)	10 (19)	94 (21)		8 (27)	50 (15)	
Unknown	2 (4)	6 (1)		0 (0)	4 (1)	
Smoking Status (%) + *			*0.0568*			*0.1612*
Never Smoker	3 (6)	68 (15)		1 (3)	10 (3)	
Current Smoker	17 (32)	98 (22)		2 (7)	56 (16)	
Reformed smoker	31 (58)	268 (60)		16 (53)	165 (48)	
Unknown	2 (4)	14 (3)		11 (37)	116 (33)	
Vital Status (%) ^			*0.0001*			*0.0006*
Dead	19 (36)	285 (64)		20 (67)	147 (42)	
Alive	34 (64)	163 (36)		10 (33)	200 (58)	
^ Fisher's exact test, + Chi square test, # t test, * unknown patients removed from significance						

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

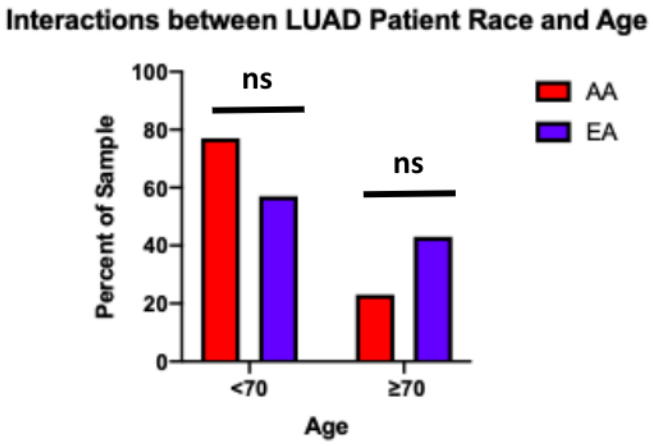

Figure 1. Analysis of potential confounding variables in the LUAD patient cohort. Bar graph showing two-tailed t-test comparing patient race and age (<70 and ≥70). (column effect ns > 0.05).

Table 2: *PDE4A* SCNV, gene and isoform expression significance by sex and race in TCGA patients with LUAD and LUSC.

	LUAD				LUSC			
Grouping	**SCNV**	**Gene**	**Isoforms**		**SCNV**	**Gene**	**Isoforms**	
AA Males vs EA Males	0.64	0.93	uc002moj	0.4*	0.03	0.57*	uc002moj	0.83
			uc002mok	0.77			uc002mok	0.96*
			uc002mol	0.53			uc002mol	0.87*
			uc002mom	0.46			uc002mom	0.85
			uc002moo	0.95*			uc002moo	0.13*
AA Females vs EA Females	0.44	0.97	uc002moj	0.69*	0.70	0.47*	uc002moj	0.62*
			uc002mok	0.35			uc002mok	0.84
			uc002mol	0.75			uc002mol	0.62*
			uc002mom	0.65			uc002mom	0.97*
			uc002moo	0.38*			uc002moo	0.80*

Highlighted cells show significant differences by sex and race, ($p<0.05$), * indicates higher mean expression in AA patients.

Table 3: *PDE4B* SCNV, gene and isoform expression significance by sex and race in TCGA patients with LUAD and LUSC.

	LUAD				LUSC			
Grouping	**SCNV**	**Gene**	**Isoforms**		**SCNV**	**Gene**	**Isoforms**	
AA Males vs EA Males	0.35	0.82*	uc001dcn	0.81*	0.51	0.01	uc001dcn	0.02
			uc009war	0.48*			uc009war	0.48*
			uc001dco	0.33*			uc001dco	0.73
			uc001dcp	0.51*			uc001dcp	0.16
			uc001dcq	0.88*			uc001dcq	0.01
			uc009was	<0.01			uc009was	0.02
AA Females vs EA Females	0.39	0.84*	uc001dcn	0.01	0.97	0.28*	uc001dcn	0.81*
			uc009war	0.53			uc009war	0.67*
			uc001dco	0.58*			uc001dco	0.39
			uc001dcp	0.05			uc001dcp	0.30*
			uc001dcq	0.89*			uc001dcq	0.37*
			uc009was	0.92*			uc009was	0.03*

Highlighted cells show significant differences by sex and race, ($p<0.05$), * indicates higher mean expression in AA patients.

Table 4: *PDE4C* SCNV, gene and isoform expression significance by sex and race in TCGA patients with LUAD and LUSC.

	LUAD				LUSC			
Grouping	**SCNV**	**Gene**	**Isoforms**		**SCNV**	**Gene**	**Isoforms**	
AA Males vs EA Males	0.03	0.05	uc0100xqc	0.53	0.01	0.27	uc010xqc	0.01
			uc002nig	0.10			uc002nig	0.47*
			uc002nih	0.84*			uc002nih	0.89*
			uc010ebk	0.51*			uc010ebk	<0.0001
			uc002nif	0.37*			uc002nif	0.11
			uc002nii	0.01			uc002nii	0.63
			uc010ebl	0.84*			uc010ebl	0.08
			uc002nik	0.32			uc002nik	0.31
			uc002nil	<0.001			uc002nil	0.54
			uc010ebm	0.67			uc010ebm	<0.001
			uc002nim	0.96			uc002nim	0.31
AA Females vs EA Females	0.18	0.03	uc0100xqc	0.02	0.80	0.51*	uc010xqc	0.01
			uc002nig	0.44			uc002nig	0.24*
			uc002nih	0.35*			uc002nih	0.66
			uc010ebk	0.66*			uc010ebk	0.35*
			uc002nif	0.38			uc002nif	0.59*
			uc002nii	0.03			uc002nii	0.44
			uc010ebl	0.91*			uc010ebl	0.37*
			uc002nik	<0.0001			uc002nik	N/A
			uc002nil	<0.01			uc002nil	0.81
			uc010ebm	0.27*			uc010ebm	0.43*
			uc002nim	0.49*			uc002nim	0.77

Highlighted cells show significant differences by sex and race ($p<0.05$), * indicates higher mean expression in AA patients.

Table 5: *PDE4D* SCNV, gene and isoform expression significance by sex and race in TCGA patients with LUAD and LUSC.

	LUAD				LUSC			
Grouping	**SCNV**	**Gene**	**Isoforms**		**SCNV**	**Gene**	**Isoforms**	
AA Males vs EA Males	0.78	0.05	uc003jrv	<0.001	0.77	0.76*	uc003jrv	0.50*
			uc003jsa	0.52			uc003jsa	0.71*
			uc003jrs	0.19			uc003jrs	0.97*
			uc003jrt	0.20			uc003jrt	0.74
			uc003jru	0.89			uc003jru	0.13
			uc003jrw	0.47			uc003jrw	0.42*
			uc003jrx	0.12			uc003jrx	<0.0001
			uc003jry	N/A			uc003jry	N/A
			uc003jrz	0.68*			uc003jrz	0.62*
			uc003jsb	0.73*			uc003jsb	0.75*
			uc010iwi	0.53			uc010iwi	0.75
			uc003jsc	0.43			uc003jsc	0.35*
			uc003jse	0.87*			uc003jse	0.66*
			uc003iwj	0.03			uc003iwj	0.73*
AA Females vs EA Females	0.29	0.51*	uc003jrv	0.79*	0.31	0.24*	uc003jrv	0.86*
			uc003jsa	<0.01			uc003jsa	0.60
			uc003jrs	0.99			uc003jrs	0.68
			uc003jrt	0.84			uc003jrt	0.88*
			uc003jru	0.59*			uc003jru	0.84*
			uc003jrw	0.49			uc003jrw	0.02
			uc003jrx	0.36*			uc003jrx	0.21
			uc003jry	N/A			uc003jry	N/A
			uc003jrz	0.34*			uc003jrz	0.06*
			uc003jsb	0.49			uc003jsb	0.01*
			uc010iwi	0.03			uc010iwi	0.09
			uc003jsc	0.46*			uc003jsc	0.73*
			uc003jse	0.13			uc003jse	0.90
			uc003iwj	0.38			uc003iwj	0.42

Highlighted cells show significant differences by sex and race ($p<0.05$), * indicates higher mean expression in AA patients

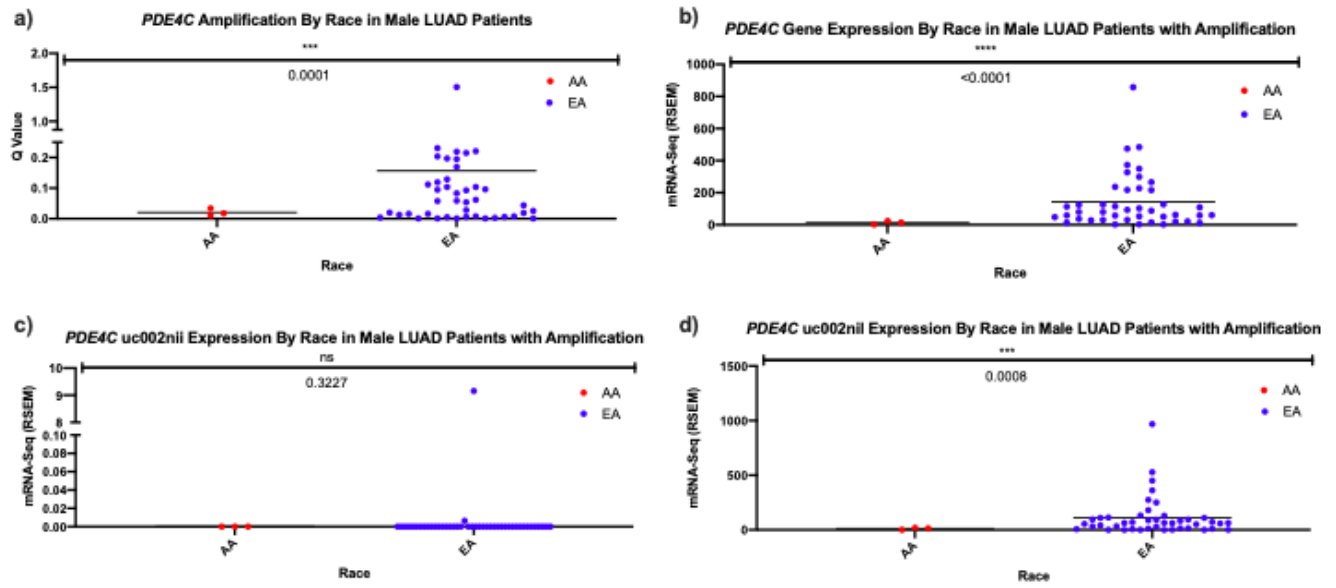

Figure 2. *PDE4C* amplification, and gene and isoform expression by amplification in TCGA patients with LUAD. Scatter dot plot showing (a) *PDE4C* amplification Q values for AA (red) and EA (blue) male LUAD patients. Plots showing mRNA-Seq RSEM values for (b) *PDE4C* gene, (c) isoform uc002nii, and (d) isoform uc002nil for AA (red) and EA (blue) male patients. (unpaired two-tailed t test with Welch correction ***P<0.001, **** P<0.0001)

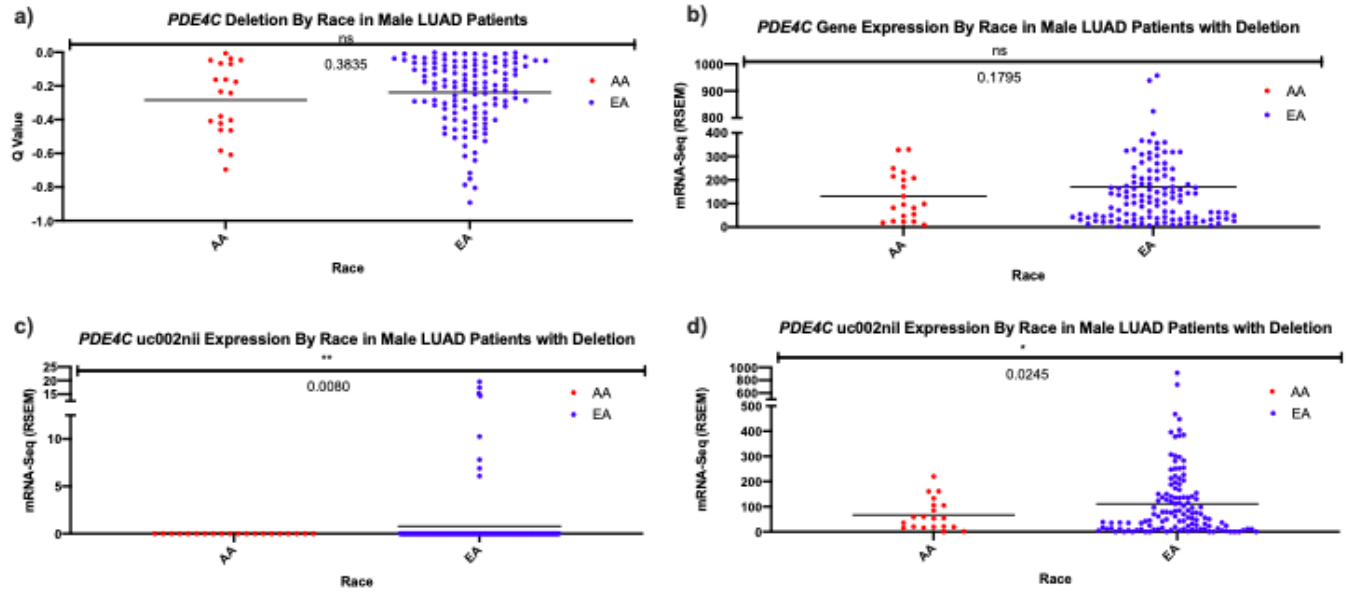

Figure 3. *PDE4C* deletion, and gene and isoform expression by deletion in TCGA patients with LUAD. Scatter dot plot showing (a) *PDE4C* deletion Q values for AA (red) and EA (blue) male LUAD patients. Plots showing mRNA-Seq RSEM values for (b) *PDE4C* gene, (c) isoform uc002nii, and (d) isoform uc002nil for AA (red) and EA (blue) male patients. (unpaired two-tailed t test with Welch correction ***P<0.001, **** P<0.0001)

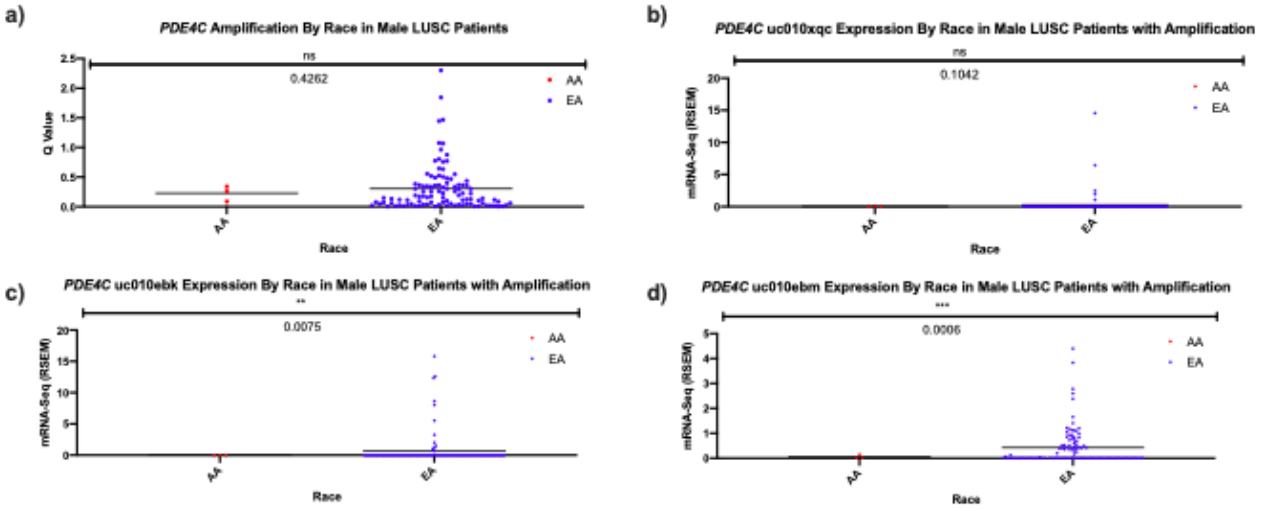

Figure 4. *PDE4C* amplification, and gene and isoform expression by amplification in TCGA patients with LUSC. Scatter dot plot showing (a) *PDE4C* amplification Q values for AA (red) and EA (blue) LUSC male patients. Plots showing mRNA-Seq RSEM values for (b) isoform uc010xqc, (c) isoform uc010ebk and (d) isoform uc010ebm for AA (red) and EA (blue) male patients. (unpaired two-tailed t test with Welch correction **<0.01, ***P<0.001)

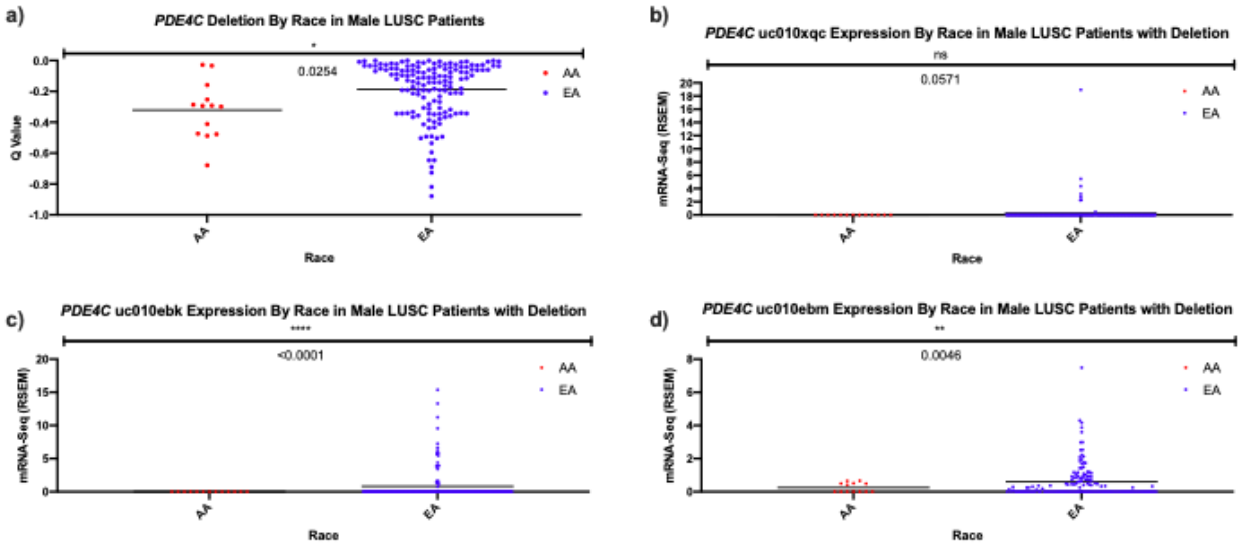

Figure 5. *PDE4C* deletion, and gene and isoform expression by deletion in TCGA patients with LUSC. Scatter dot plot showing (a) *PDE4C* deletion Q values for AA (red) and EA (blue) LUSC male patients. Plots showing mRNA-Seq RSEM values for (b) *PDE4C* gene, (b) isoform uc010ebk and (d) isoform uc010ebm for AA (red) and EA (blue) male patients with deletion. (unpaired two-tailed t test with Welch correction *<0.05, **<0.01, ****P<0.001)

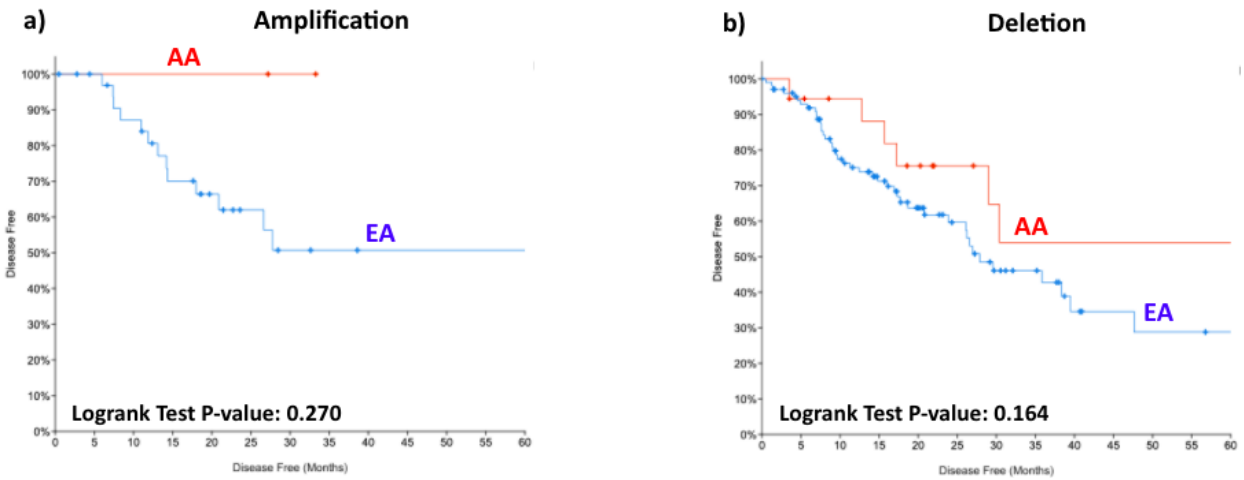

Figure 6. 5 year disease-free survival analysis based on a) amplification and b) deletion of *PDE4C* in TCGA patients with LUAD. Disease-specific survival analysis based on *PDE4C* SCNV status in AA (red) and EA (blue) patients with a) amplification and b) deletion.

Table 6: Median months disease-free based on SCNV status (amplification or deletion) of AA and EA male TCGA patients with LUAD.

	Race	Cases, Total	Cases, Relapsed/ Progressed	Median Months Disease-Free
Amplification	(A) AA	2	0	N/A
	(B) EA	35	15	67.18
Deletion	(A) AA	18	8	62.19
	(B) EA	103	44	27.89

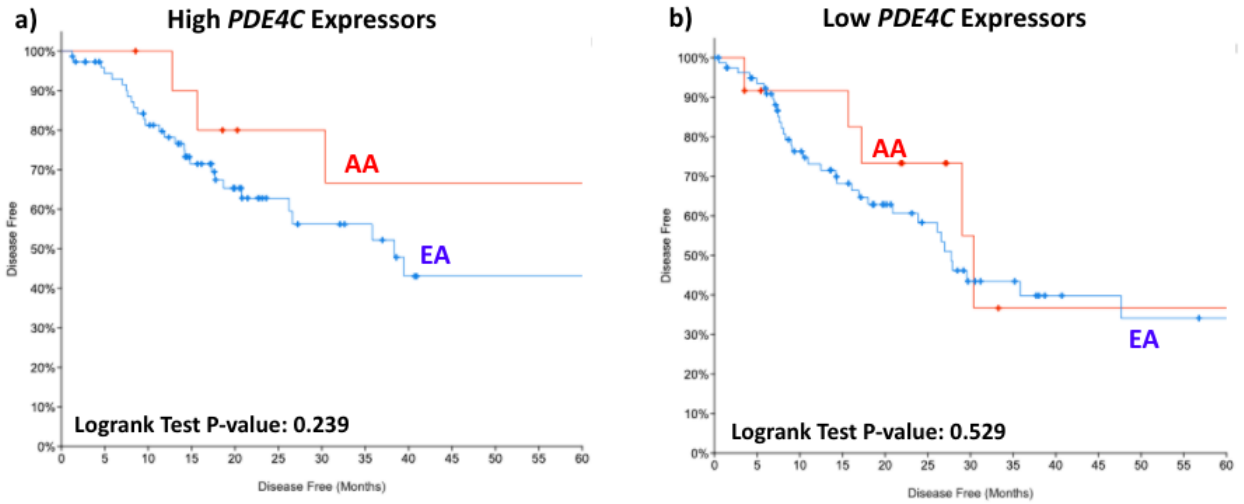

Figure 7. 5 year disease-free survival analysis based on a) high and b) low *PDE4C* gene expression in TCGA patients with LUAD. Disease-specific survival analysis based on *PDE4C* gene expression in AA (red) and EA (blue) patients in a) high- and b) low-expressors.

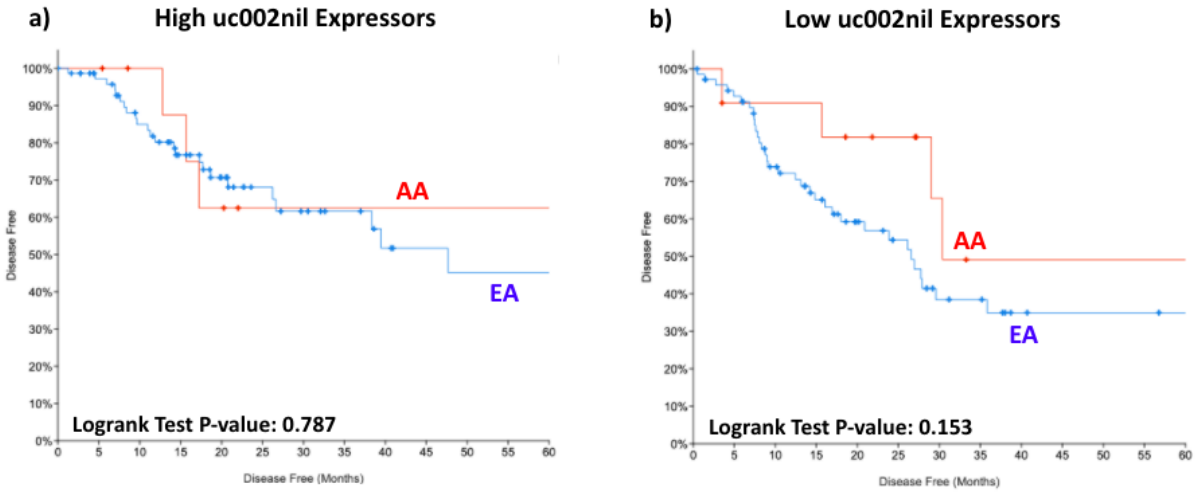

Figure 8. 5 year disease-free survival analysis based on a) high and b) low *PDE4C* uc002nil expression in TCGA patients with LUAD. Disease-specific survival analysis based on isoform uc002nil expression in AA (red) and EA (blue) patients in a) high- and b) low-expressors.

Table 7: Median months disease-free based on *PDE4C* gene and isoform uc002nil expression status of AA and EA male TCGA patients with LUAD.

		PDE4C			isoform uc002nil		
	Race	Cases, Total	Cases, Relapsed/ Progressed	Median Months Disease-Free	Cases, Total	Cases, Relapsed/ Progressed	Median Months Disease-Free
High	(A) AA	11	5	62.19	10	5	62.19
	(B) EA	74	30	38.37	75	26	47.63
Low	(A) AA	12	5	30.39	11	4	30.39
	(B) EA	81	36	27.76	71	35	26.58

F. References

1. Hsien Lai, Samuel, Guston Zervoudakis, Jesse Chou, Mark E. Gurney, and Kelly M. Quesnelle. "PDE4 Subtypes in Cancer." *Oncogene* 39, no. 19 (May 2020): 3791–3802.
2. Kolosionek, Ewa, Rajkumar Savai, Hossein Ardeschir Ghofrani, Norbert Weissmann, Andreas Guenther, Friedrich Grimminger, Werner Seeger, Gamal Andre Banat, Ralph Theo Schermuly, and Soni Savai Pullamsetti. "Expression and Activity of Phosphodiesterase Isoforms during Epithelial Mesenchymal Transition: The Role of Phosphodiesterase 4." *Molecular Biology of the Cell* 20, no. 22 (November 2009): 4751–65.
3. Pullamsetti, S. S., G. A. Banat, A. Schmall, M. Szibor, D. Pomagruk, J. Hänze, E. Kolosionek, et al. "Phosphodiesterase-4 Promotes Proliferation and Angiogenesis of Lung Cancer by Crosstalk with HIF." *Oncogene* 32, no. 9 (February 28, 2013): 1121–34.
4. He, Rong-quan, Xiao-jiao Li, Lu Liang, You Xie, Dian-zhong Luo, Jie Ma, Zhi-gang Peng, Xiao-hua Hu, and Gang Chen. "The Suppressive Role of MiR-542-5p in NSCLC: The Evidence from Clinical Data and in Vivo Validation Using a Chick Chorioallantoic Membrane Model." *BMC Cancer* 17 (September 19, 2017).
5. Praveena, K. S. S., Shylaprasad Durgadas, N. Suresh Babu, Surekha Akkenapally, C. Ganesh Kumar, Girdhar Singh Deora, N. Y. S. Murthy, Khagga Mukkanti, and Sarbani Pal. "Synthesis of 2,2,4-Trimethyl-1,2-Dihydroquinolinyl Substituted 1,2,3-Triazole Derivatives: Their Evaluation as Potential PDE 4B Inhibitors Possessing Cytotoxic Properties against Cancer Cells." *Bioorganic Chemistry* 53 (April 1, 2014): 8–14.
6. Garritano, S, A Inga, F Gemignani, and S Landi. "More Targets, More Pathways and More Clues for Mutant P53." *Oncogenesis* 2, no. 7 (July 2013): e54.
7. Qiu, Zhe-Wei, Jia-Hao Bi, Adi F. Gazdar, and Kai Song. "Genome-Wide Copy Number Variation Pattern Analysis and a Classification Signature for Non-Small Cell Lung Cancer." *Genes, Chromosomes and Cancer* 56, no. 7 (2017): 559–69.
8. Nancarrow, Derek J., Herlina Y. Handoko, B. Mark Smithers, David C. Gotley, Paul A. Drew, David I. Watson, Andrew D. Clouston, Nicholas K. Hayward, and David C. Whiteman. "Genome-Wide Copy Number Analysis in Esophageal Adenocarcinoma Using High-Density Single-Nucleotide Polymorphism Arrays." *Cancer Research* 68, no. 11 (June 1, 2008): 4163–72.
9. Vitting-Seerup, Kristoffer, and Albin Sandelin. "The Landscape of Isoform Switches in Human Cancers." *Molecular Cancer Research* 15, no. 9 (September 1, 2017): 1206–20.
10. Wang, Bi-Dar, Kristin Ceniccola, SuJin Hwang, Ramez Andrawis, Anelia Horvath, Jennifer A. Freedman, Jacqueline Olender, et al. "Alternative Splicing Promotes Tumour Aggressiveness and Drug Resistance in African American Prostate Cancer." *Nature Communications* 8, no. 1 (June 30, 2017): 1–15.
11. Aldrich, Melinda C., Sarah F. Mercaldo, Kim L. Sandler, William J. Blot, Eric L. Grogan, and Jeffrey D. Blume. "Evaluation of USPSTF Lung Cancer Screening Guidelines Among African American Adult Smokers." *JAMA Oncology* 5, no. 9 (September 2019): 1318–24.
12. Magee, Rogan, Phillipe Loher, Eric Londin, and Isidore Rigoutsos. "Threshold-Seq: A Tool for Determining the Threshold in Short RNA-Seq Datasets." *Bioinformatics* 33, no. 13 (July 1, 2017): 2034–36.

V. CHAPTER 2: Effect of *PDE4C* Overexpression on the EMT Phenotype in AA and EA Male Lung Cancer Cells

A. Introduction

1. PDE4C overexpression: Potential oncogene

PDE4C overexpression has been associated with a number of disease phenotypes.[1] For example, *PDE4C* expression was higher in autonomous thyroid adenomas (ATA) compared to surrounding normal thyroid tissue, but only significantly so in ATA with TSH receptor mutants.[2] Additionally, Chamseddine and colleagues found that *PDE4C* isoforms are significantly upregulated in patients with myelodysplastic syndrome (MDS), compared to healthy individuals.[3] This upregulation was associated with worse survival outcomes.[3] Lastly, *PDE4C* demonstrated greater protein expression in NSCLC tissues, particularly LUAD, compared to normal tissues.[4] *PDE4C* is a direct target of the tumor suppressing miR-542-5p, which has been shown to reduce in vitro proliferation of the lung cancer cell line A549.[4] miR-542-5p therapies may improve disease phenotypes in individuals with ATA, MDS and NSCLC.

2. PDE4C underexpression: Potential tumor suppressor gene

In contrast, *PDE4C* underexpression has also been linked to a variety of malignancies. *PDE4C* mRNAs were downregulated in stomach adenocarcinoma and associated with worse overall survival.[5] In glioma, *PDE4C* promoter hypermethylation and reduced protein expression were associated with grade progression and overall worse survival.[6] Upon 5-Aza-2'-deoxycytidine treatment and *PDE4C* hypomethylation, *PDE4C* upregulation promoted apoptosis and inhibited migration in glioma cell lines.[6] As such, *PDE4C* has been identified as a candidate glioma biomarker and potential tumor suppressor.[6] Interestingly, *PDE4C* DNA methylation increases with age and may be associated with later age at onset of disease.[7]

Similar to *PDE4C* overexpression, *PDE4C* underexpression can also be regulated by miRNAs. A recent study by Tsai and colleagues on asthma and COPD found elevated levels of miR-10a-5p may contribute to disease pathogenesis by suppressing *PDE4C*.[8] Although the exact mechanism of miR-10a-5p in NSCLC has yet to be elucidated, the miR has also been found to be upregulated in LUAD tumors, relative to normal tissue from the same patients.[9] This would suggest that miR-10a-5p is an oncomiR that downregulates *PDE4C* in LUAD tumors. Anti-miR-10a-5p therapies may improve stomach adenocarcinoma, glioma and LUAD patient outcomes.

3. PDE4C overexpression and underexpression in LUAD patients

Chapter 1 of this study found 30% of EA males had *PDE4C* amplifications and overexpression, compared to 13% of AA males in the TCGA LUAD patient cohort (Ch. 1 Fig. 2). Increased *PDE4C* expression was associated with better survival in both races, suggesting a protective effect (Ch. 1 Table 6-7 and Fig. 6-7). However, the improvement was greater for AA males with high expression (~62 months), compared to their EA counterparts (~38 months). Still, the majority of AA males are low-expressors (87%), suggesting they may benefit from *PDE4C*-targeting interventions.

4. PDE4 genes promote EMT phenotype and EMT markers can be used as indicators of treatment efficacy

EMT is a reversible process by which epithelial cells lose their cell polarity and cell-cell adhesion, and become invasive, migratory mesenchymal cells.[10] This transition enables biological processes such as embryonic development and wound healing, as well as cancer metastasis and progression.[10] In NSCLC, low e-cadherin (*CDH1)* and high vimentin (*VIM)* expression are characteristics of EMT, and correlated with worse prognosis and overall

survival.[11] EMT promotes chemotherapeutic (cisplatin) resistance and EMT markers can be used as a measure of treatment efficacy.[12] If *PDE4C* overexpression treatments mechanistically target EMT, they would be viable alternatives to chemotherapy.

5. Hypothesis

To our knowledge, no study to date has characterized the population-specific effects of *PDE4C* plasmid overexpression on EMT phenotypic markers, *CDH1* and *VIM*. I hypothesize that *PDE4C* plasmid overexpression will result in low *CDH1* and high *VIM* expression in AA and EA LUAD male cell lines.

B. Methods

1. EMT gene comparison by race and PDE4C status in LUAD TCGA cohort

Downloaded mRNA sequencing data for EMT phenotypic markers from Broad GDAC Firehose (https://gdac.broadinstitute.org/) Performed differential analyses for *CDH1* and *VIM* expression according to race and *PDE4C* expression status in the LUAD TCGA cohort. GraphPad Prism 9.0 was used to carry out two-tailed t-tests with Welch corrections.

2. Tissue culture

Immortalized, patient-derived LUAD tumor cell lines were obtained from the National Cancer Institute and American Type Culture Collection. The AA male cell line used was NCI-H1373. The EA male cell line used was A459. Cell lines were cultured in RPMI media supplemented with 10% fetal bovine serum, 1% glutamine, and 1% penicillin-streptomycin. All cells were incubated at 37 ºC and 5% CO_2, and split every three days or until 90% confluent.

3. Bacterial transformation and PDE4C plasmid expansion

Subcloning Efficiency™ DH5α Competent Cells (ThermoFisher Scientific, Cat# 18265017) were transformed with 10 ng *PDE4C* cDNA ORF Clone, Human, N-GFPSpark tag

plasmid (Sino Biological, Cat# HG13977-ANG) and 250 pg pUC19 control plasmid. Briefly, after 30 minutes of chemical transformation on ice, cells were heat shocked for 20 seconds at 42℃, and grown in 950 µL of pre-warmed SOC media (ThermoFisher Scientific, Cat# 15544034) for 1 hour at 225 rpm in 37℃. A total of 200 µL of transformed bacterial cells were spread on pre-warmed 50 µg/mL kanamycin (*PDE4C* plasmid) and 100 µg/mL ampicillin (*pUC19* control plasmid) plates (Sigma Aldrich, Cat# L0543-10EA and L5667-10EA). All plates were incubated for 16-18 hours at 37℃. Individual colonies on PDE4C plates were picked for 3 mL starter cultures in LB Broth with 50 µg/mL kanamycin (Avantor, Cat# 100219-756). Several 50 mL midiprep cultures were inoculated using 3mL starter cultures. *PDE4C* plasmid extraction was performed using a Midiprep Kit (Zymo Research, Cat# D4200) in accordance with manufacturer's protocol.

4. PDE4C plasmid overexpression in AA and EA LUAD cell lines and fluorescent microscopy imaging

Approximately 250K NCI-H1373 (AA) and A549 (EA) cells were seeded in 6-well dishes at least 24 hours pre-transfection.Two biological replicates (BR1 and BR 2) were carried out, each containing technical duplicates of transfected (TF) and untransfected (UTF) cells. Performed overexpression with 2,500 ng of *PDE4C* plasmid via Lipofectamine 3000 Reagent (Cat# L3000008) in accordance with the 6-well, transfection of plasmid protocol. After 48 hours, visualized and captured GFP-images using a Zeiss PrimoVert microscope and Zen Blue 2.6 software.

5. Total RNA isolation and quality

RNA was isolated from TF and UTF cell pellets (BR 1 and BR 2) using the Zymo Direct-zol RNA MiniPrep kit (Cat# 11-330T) in accordance with the associated protocol. The

Qiagen RNase-Free DNase Set (Cat# 79254) was used to perform the optional DNase treatment to degrade contaminating genomic DNA. Total RNA was quantified with an Eppendorf BioSpectrometer, by taking an average of two concentration measurements. A 260/280 ratio was noted to assess RNA quality. Only samples with a 260/280 ratio of ~2.0 were used. All RNA samples were stored at -80 ºC.

6. cDNA synthesis

cDNA synthesis was accomplished via Applied Biosystems High-Capacity RNA-to-cDNA Kit (Cat# 4388950) in accordance with manufacturer's protocol. About 1000 ng total RNA was converted to cDNA for BR1 and BR2, including reverse transcriptase (+RT) and (-RT) controls.

7. qRT-PCR primer design, resuspension, and dilution

Primers were designed for quantitative reverse transcription PCR (qRT-PCR) using NCBI Gene and Primer-BLAST. Candidate genes (*PDE4C*, *CDH1*, and *VIM*) and the housekeeping control gene (*GAPDH*) were searched in NCBI Gene with a *Homo sapiens* filter. The RefSeq transcript accession numbers were obtained. The longest isoforms were selected and the isoform accession numbers were recorded. The “pick primers” feature was linked to Primer-BLAST. Default primer design parameters were selected with the following changes: PCR product size: Min-Max 70-200; Min-Max 50-65; Exon junction span: Primer must span an exon-exon junction. Exon-exon junctions were spanned, when possible, to avoid binding to contaminating genomic DNA which contains introns.

Resuspended primers to make 100 µM stock solutions, and centrifuged lyophilized primers at 14,000 RPM at room temperature for one minute. Added appropriate amounts of RNase-free water to each primer tube based on the amount of nmol in the tube. To create 10 µM

diluted stock solutions for qRT-PCR, 10 µL forward primer [100 µM], 10µL reverse primer [100 µM], and 180 µL RNase free water was added to a 1.5 mL PCR-grade tube for each candidate gene. The same process was repeated for *GAPDH*.

8. qRT-PCR analysis of PDE4C, CDH1 and VIM

Quantitative PCR was performed on cDNA samples using Applied Biosystems SYBR Green PowerUp Master Mix (Cat# 4387406), using a total reaction volume of 20 µL per well set-up in accordance with the company's protocol. All experiments included either *PDE4C* (annealing temperature: 57ºC), *CDH1* (annealing temperature: 58ºC), or *VIM* (annealing temperature: 60ºC), and the following controls: *GAPDH* as the housekeeping gene (internal positive control) and water (no template negative control). Data were analyzed via the delta-delta cycle threshold ($\Delta\Delta$ Ct) method for relative quantification.

C. Results

High PDE4C expression is associated with significant racial differences in VIM expression in AA and EA males with LUAD.

Among *PDE4C* high expressors with LUAD, there was marginally decreased *CDH1* compared to *PDE4C* low expressors (Fig. 1). Though EA high expressors had a wider range for *CDH1*, both AA and EA high expressors shared similar medians of expression for the *CDH1* gene (Fig. 1). There was no significant difference in *CDH1* expression by race (Fig. 1). However, there was a significant difference in *VIM* expression by race (Fig. 2). AA high expressors demonstrated lower *VIM* expression in median and range, compared to their EA high expressor counterparts (Fig. 2).

Fluorescent imaging demonstrated efficient PDE4C transfection.

The fluorescence status of AA (NCI-H1373) and EA (A549) male LUAD cells was assessed 48 hours post-*PDE4C* plasmid transfection. Untransfected AA and EA LUAD cells did not show GFP tag presence and served as a negative control (Fig. 3). Transfected AA and EA LUAD cells both showed GFP tag expression (Fig. 3). This substantiated effective *PDE4C* plasmid transfection and overexpression (Fig. 3).

PDE4C overexpression is associated with increased CDH1 and VIM expression in AA cells.

The expression status of EMT phenotypic markers, *CDH1* and *VIM*, was assessed post-*PDE4C* plasmid transfection. In transfected AA NCI-H1373 cells, both *CDH1* and *VIM* expression was increased relative to *GAPDH* (Fig. 4-5). In transfected EA A549 cells, *CDH1* expression was unchanged and *VIM* expression was decreased relative to *GAPDH* (Fig. 4-5). Since both findings in AA and EA males diverge from the traditional EMT phenotype, *PDE4C* may be acting indirectly.

Six PDE4C protein partners are associated with lung cancer.

Protein-protein interaction analysis revealed PDE4C has eight known direct protein partners in the STRING database (Fig. 6). Of which, six protein partners are associated with lung cancer in Malacards (Table 1). In particular, PRKA and PRKACB play a role in cell differentiation and proliferation, respectively (Table 1).

AA LUAD cell line demonstrates greater genetic admixture compared to EA cell line.

The two human cell lines used were matched by sex and age (Table 2). When comparing genetic ancestry, the AA cell line (NCI-H1373) was 71.79% African, 1.21% Native American, 4.32% Asian and 22.67% European (Table 2). The EA cell line (A549) was 99.28% European

and 0.72% Asian (Table 2). Several studies have shown variation in genetic ancestry drives differences in tumor biology.

D. Discussion

In contrast to the proposed hypothesis of this chapter, high *PDE4C* expression (whether naturally occurring or experimentally induced) is not associated with the EMT phenotype in male patients with LUAD, regardless of race. The LUAD TCGA cohort and cell line revealed seemingly contradictory patterns for AA males. AA *PDE4C* high expressors showed low *CDH1* and low *VIM* gene expression. Meanwhile induced overexpressors showed high *CDH1* and high *VIM* gene expression. This suggests that *PDE4C* is not directly involved in the EMT pathway in AA male patients with LUAD. Instead, PDE4C protein partners, such as PRKACB which plays a role in cell proliferation, may affect lung cancer progression and aggressiveness.

A number of studies have found that differences in genetic ancestry can be linked with more aggressive tumor biology.[12] Due to the Translatlantic Slave Trade, there is greater heterogeneity and genetic admixture in the AA population.[13] This may help explain the discrepancy between AA primary lung tumor data and lung cancer cell lines. AA LUAD patients in the TCGA cohort have West African ancestry that ranges between 33.46% and 94.28%. The AA LUAD cell line NCI-H1373 has 71.79% African ancestry. EA patients have <10% West African ancestry and the EA cell line A549 is 99.28% European.

It is important to note the limitations of this study. One of which is the use of only one cell line of each race. Mechanistic studies using multiple cell lines of varying genetic ancestry would provide more robust results and more generalizable conclusions. Another limitation is the lack of selection for successfully transfected H1373 and A549 cells with Geneticin (G418). As a result, qRT-PCR results for the *PDE4C* OE experimental group may reflect expression levels in both successfully and unsuccessfully transfected cells. Future investigations should transfect

multiple cell lines with a *PDE4C* plasmid with a GFP tag, and select for both kanamycin and G418 resistance.

E. Tables and Figures

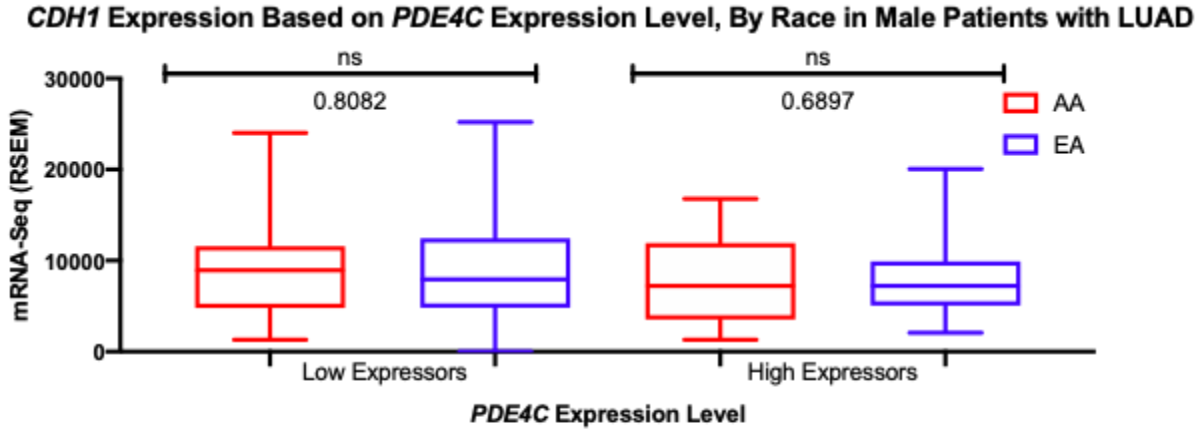

Figure 1. *CDH1* expression based on *PDE4C* expression level, by race in male TCGA patients with LUAD. Box and whisker plot showing unpaired two-tailed t-test with a Welch's correction showing mRNA-Seq RSEM values of *CDH1* for AA (red) and EA (blue) male LUAD patients, stratified by low and high *PDE4C* expression status. (ns>0.05)

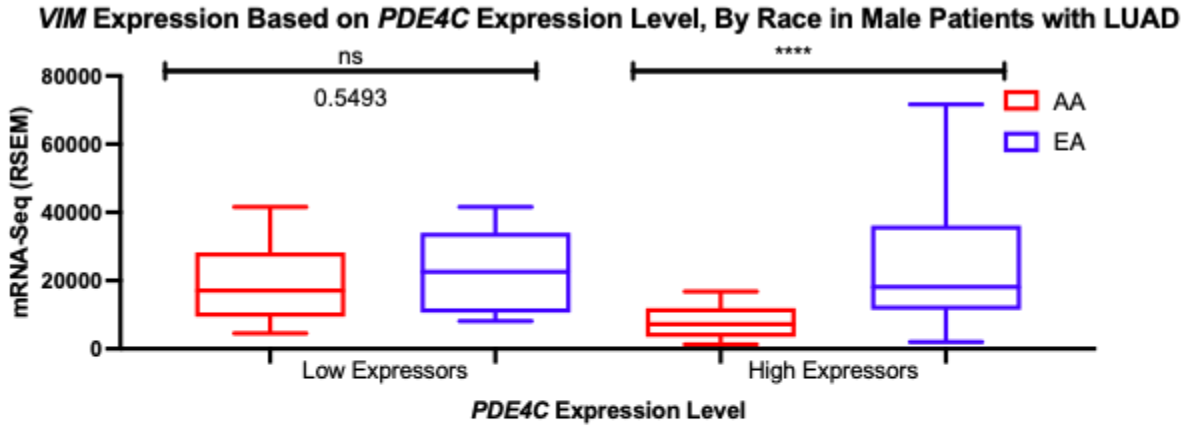

Figure 2. *VIM* expression based on *PDE4C* expression level, by race in male TCGA patients with LUAD. Box and whisker plot showing unpaired two-tailed t-test with a Welch's correction showing mRNA-Seq RSEM values of *VIM* for AA (red) and EA (blue) male LUAD patients, stratified by low and high *PDE4C* expression status. (unpaired two-tailed t test with Welch correction ns>0.05, ****<0.0001)

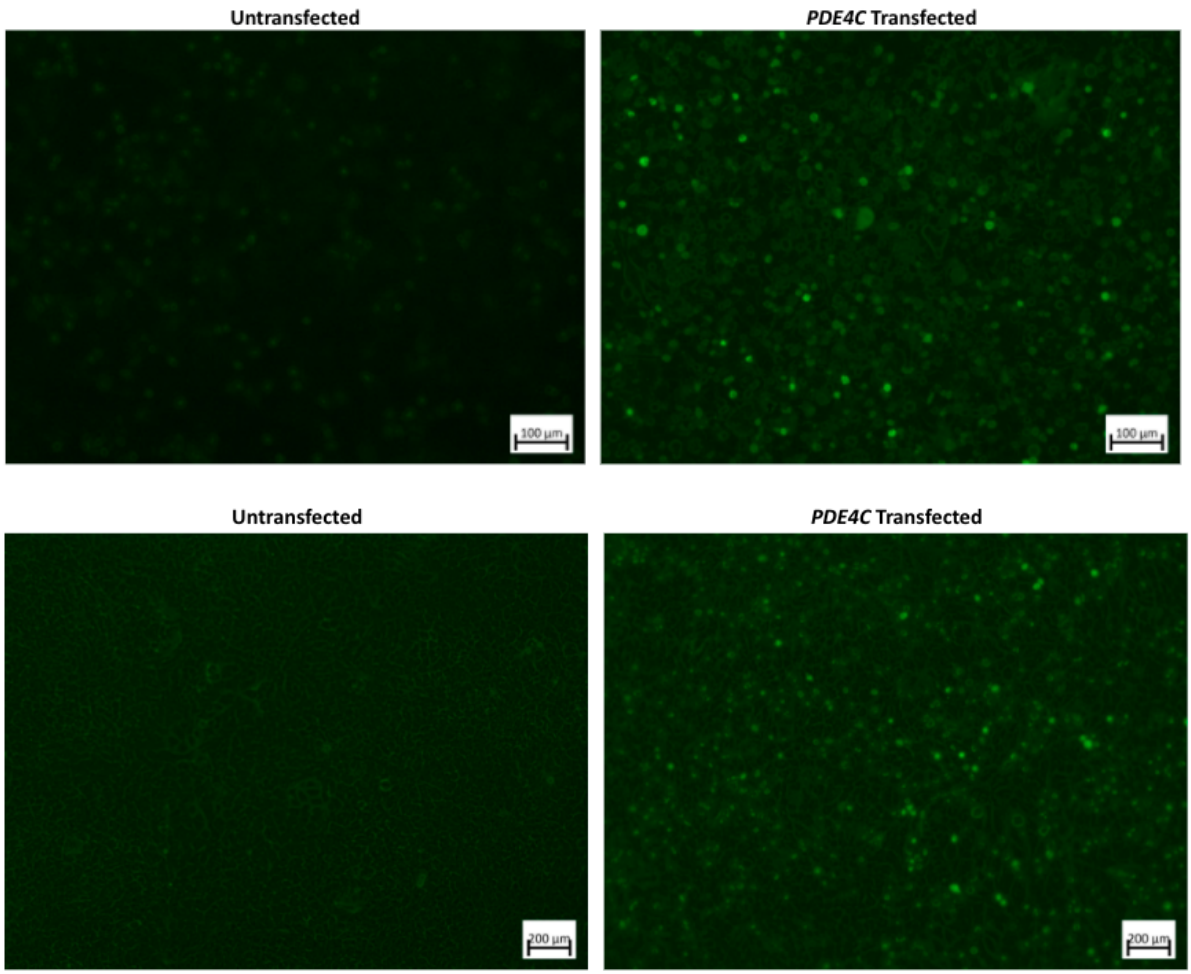

Figure 3. *PDE4C* overexpression in male LUAD cell lines. Fluorescent imaging showing *PDE4C* plasmid expressing GFP-tag in untransfected and transfected a) AA and b) EA male LUAD cell lines.

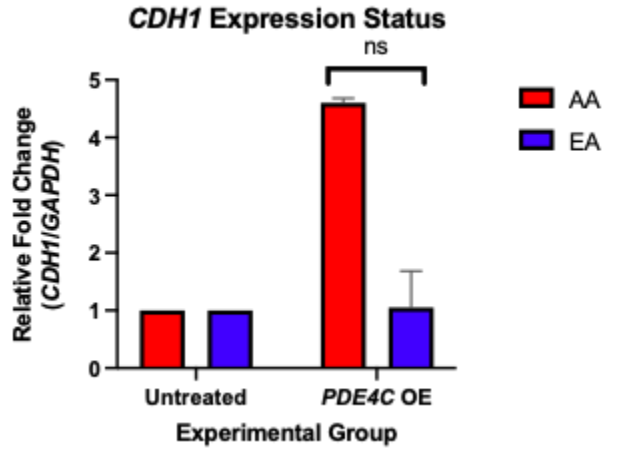

Figure 4. *CDH1* expression status after *PDE4C* overexpression. qRT-PCR analysis showing relative *CDH1* gene expression levels in untreated and 48 hour *PDE4C* transfected (OE) AA (NCI-H1373) and EA (A549) LUAD cell lines. All fold changes were calculated using the Delta Delta CT method. (repeated measures one-way ANOVA ns>0.05, error bars represent technical and biological replicates)

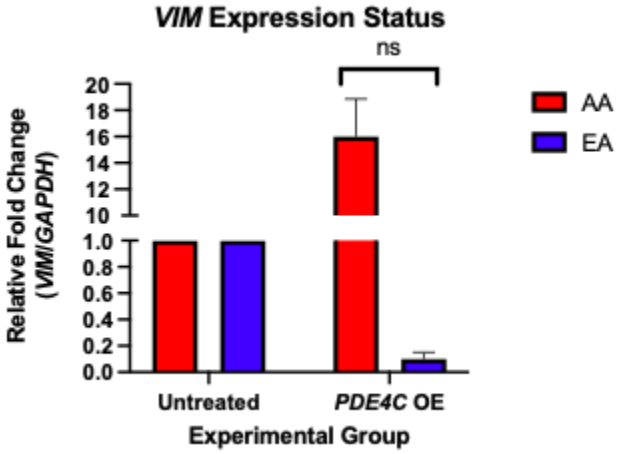

Figure 5. *VIM* expression status after *PDE4C* overexpression. qRT-PCR analysis showing relative *VIM* gene expression levels in untreated and 48 hour *PDE4C* transfected (OE) AA (NCI-H1373) and EA (A549) LUAD cell lines. All fold changes were calculated using the Delta Delta CT method. (repeated measures one-way ANOVA ns>0.05, error bars represent technical and biological replicates)

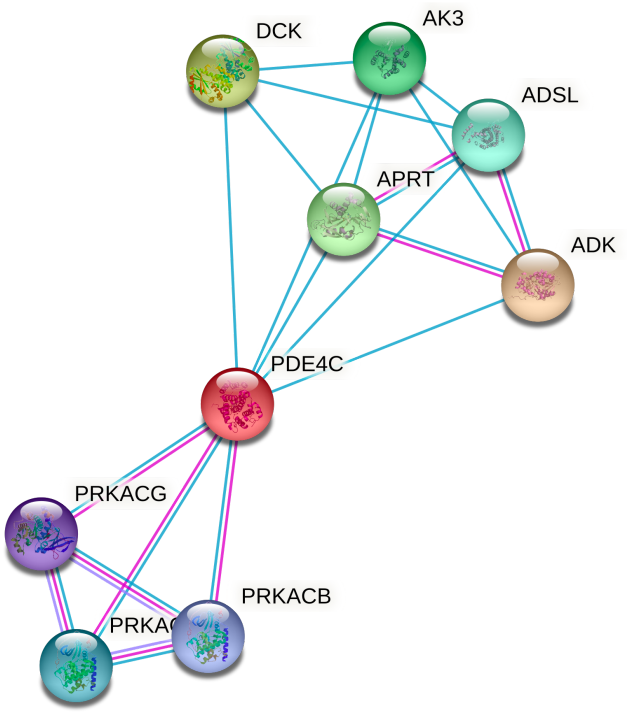

Figure 6. PDE4C protein-protein interactions. STRING network of high confidence (score≥0.700) PDE4C protein-protein interactions from curated databases (blue) and experiments (pink).

Table 1: Functions of PDE4C protein partners associated with lung cancer. Lung cancer association based on Malacards.

Protein partner	Solr Relevance Score	Function
DCK	21.432	Required for the phosphorylation of the deoxyribonucleosides deoxycytidine (dC), deoxyguanosine (dG) and deoxyadenosine (dA).
PRKACB	2.485	Mediates cAMP-dependent signaling, thus regulating cellular processes like cell proliferation.
PRKA	1.081	Phosphorylates substrates in the cytoplasm and the nucleus, regulates some cell differentiation.
PRKACG	0.504	Phosphorylates substrates in the cytoplasm and the nucleus.
ADK	0.192	ATP dependent phosphorylation of adenosine and other nucleosides, potential regulator of adenine and adenosine concentrations.
APRT	0.141	Catalyzes a salvage reaction to form AMP.

Table 2: LUAD Cell Line Patient Characteristics, matched by age and sex.

				Genetic Ancestry			
Cell Line	**Sex**	**Age**	**Race**	***African***	***Native American***	***Asian***	***European***
NCI-H1373	Male	56	AA	71.79	1.21	4.32	22.67
A549	Male	58	EA	0	0	0.72	99.28

F. References

1. Hsien Lai, Samuel, Guston Zervoudakis, Jesse Chou, Mark E. Gurney, and Kelly M. Quesnelle. "PDE4 Subtypes in Cancer." *Oncogene* 39, no. 19 (May 2020): 3791–3802.
2. Persani, Luca, Andrea Lania, Luisella Alberti, Roberto Romoli, Giovanna Mantovani, Sebastiano Filetti, Anna Spada, and Marco Conti. "Induction of Specific Phosphodiesterase Isoforms by Constitutive Activation of the CAMP Pathway in Autonomous Thyroid Adenomas." *The Journal of Clinical Endocrinology & Metabolism* 85, no. 8 (August 1, 2000): 2872–78.
3. Chamseddine, Ali N., Monica Cabrero, Yue Wei, Irene Ganan-Gomez, Simona Colla, Koichi Takahashi, Hui Yang, Zachary S. Bohannan, and Guillermo Garcia-Manero. "PDE4 Differential Expression Is a Potential Prognostic Factor and Therapeutic Target in Patients With Myelodysplastic Syndrome and Chronic Myelomonocytic Leukemia." *Clinical Lymphoma, Myeloma and Leukemia* 16 (August 1, 2016): S67–73.
4. He, Rong-quan, Xiao-jiao Li, Lu Liang, You Xie, Dian-zhong Luo, Jie Ma, Zhi-gang Peng, Xiao-hua Hu, and Gang Chen. "The Suppressive Role of MiR-542-5p in NSCLC: The Evidence from Clinical Data and in Vivo Validation Using a Chick Chorioallantoic Membrane Model." *BMC Cancer* 17 (September 19, 2017).
5. Zhang, Xiulei, Peiming Zheng, Zhen Li, Shanjun Gao, and Guangzhi Liu. "The Somatic Mutation Landscape and RNA Prognostic Markers in Stomach Adenocarcinoma." *OncoTargets and Therapy* 13 (August 4, 2020): 7735–46.
6. Bao, Zhaoshi, Ying Feng, Hongjun Wang, Chuanbao Zhang, Lihua Sun, Zhuohong Yan, Qingyang Liu, et al. "Integrated Analysis Using Methylation and Gene Expression Microarrays Reveals PDE4C as a Prognostic Biomarker in Human Glioma." *Oncology Reports* 32, no. 1 (July 1, 2014): 250–60.
7. Márquez-Ruiz, Ana Belén, Lucas González-Herrera, Juan de Dios Luna, and Aurora Valenzuela. "DNA Methylation Levels and Telomere Length in Human Teeth: Usefulness for Age Estimation." *International Journal of Legal Medicine* 134, no. 2 (March 1, 2020): 451–59.
8. Tsai, Ming-Ju, Yu-Chen Tsai, Wei-An Chang, Yi-Shiuan Lin, Pei-Hsun Tsai, Chau-Chyun Sheu, Po-Lin Kuo, and Ya-Ling Hsu. "Deducting MicroRNA-Mediated Changes Common in Bronchial Epithelial Cells of Asthma and Chronic Obstructive Pulmonary Disease—A Next-Generation Sequencing-Guided Bioinformatic Approach." *International Journal of Molecular Sciences* 20, no. 3 (January 28, 2019).
9. Cinegaglia, Naiara C., Sonia Cristina S. Andrade, Tomas Tokar, Maísa Pinheiro, Fábio E. Severino, Rogério A. Oliveira, Erica N. Hasimoto, et al. "Integrative Transcriptome Analysis Identifies Deregulated MicroRNA-Transcription Factor Networks in Lung Adenocarcinoma." *Oncotarget* 7, no. 20 (April 12, 2016): 28920–34.
10. Otsuki, Yuji, Hideyuki Saya, and Yoshimi Arima. "Prospects for New Lung Cancer Treatments That Target EMT Signaling." Developmental Dynamics 247, no. 3 (2018): 462–72.
11. Tsoukalas, Nikolaos, Eleni Aravantinou-Fatorou, Maria Tolia, Constantinos Giaginis, Michail Galanopoulos, Maria Kiakou, Ioannis D. Kostakis, et al.
12. Ramesh, Vignesh, Thomas Brabletz, and Paolo Ceppi. "Targeting EMT in Cancer with Repurposed Metabolic Inhibitors." *Trends in Cancer* 6, no. 11 (November 1, 2020): 942–50.

13. “Epithelial–Mesenchymal Transition in Non Small-Cell Lung Cancer.” *Anticancer Research* 37, no. 4 (April 1, 2017): 1773–78.
14. Davis, Melissa, Rachel Martini, Lisa Newman, Olivier Elemento, Jason White, Akanksha Verma, Indrani Datta, et al. “Identification of Distinct Heterogenic Subtypes and Molecular Signatures Associated with African Ancestry in Triple Negative Breast Cancer Using Quantified Genetic Ancestry Models in Admixed Race Populations.” *Cancers* 12, no. 5 (May 2020): 1220.
15. Mitchell, K. A. *et al.* Relationship between West African ancestry with lung cancer risk and survival in African Americans. *Cancer Causes Control CCC* **30**, 1259–1268 (2019).

VI. CONCLUSIONS

Chapter 1 revealed *PDE4C* SCNV profiles varied significantly between AA and EA males with lung cancer (Ch. 1 Tables 2-5). In the LUAD histology specifically, AA males demonstrated less *PDE4C* amplification and lower expression compared to their European counterparts (Ch. 1 Fig. 2). This low expression was associated with shorter disease free survival in both populations (Ch. 1 Fig. 7, Ch. 1 Table 7).

The mechanistic study in Chapter 2 was inconclusive as to the direct effect of *PDE4C* overexpression on the EMT phenotype in LUAD cells from AA males (Ch. 2 Fig. 4-5). AA genetic heterogeneity may contribute to the conflicting *CDH1* and *VIM* expression results of the TCGA cohort and cell lines (Ch. 2 Table 2). AA patients with high *PDE4C* expression and longer survival may have lower West African ancestry. They may also depend on protein partners like PRKACB instead of CDH1 and VIM (Ch. 2 Fig. 6, Ch. 2 Table 1). Future investigations should explore the *in vivo* effects of PDE4C and its six interacting protein partners on cancer progression pathways and survival in transgenic mouse or fish models.

VII. ACKNOWLEDGEMENTS

I would like to thank the Lafayette College Department of Biology and the TriBeta Research Foundation for funding and nurturing my honors thesis project. I am especially grateful to Dr. Eric Ho and Dr. Robert Kurt in the biology department, and Dr. Zoe Boekelheide in the physics department for their contributions. Finally I would like to thank my mentor, Dr. Khadijah A. Mitchell, for her endless guidance and support throughout my undergraduate research career.