Title
ZFPM2 Promoter Methylation in Pre-operative Blood Samples Is Associated with Post-Operative Atrial Fibrillation After Cardiac Surgery

Author
Fischer, Matthew Adam

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

ZFPM2 Promoter Methylation in Pre-operative Blood Samples Is Associated with Post-Operative Atrial Fibrillation After Cardiac Surgery

A thesis submitted in partial satisfaction of the

requirements for the degree Master of Science

in Clinical Research

by

Matthew Adam Fischer

2017

ABSTRACT OF THE THESIS

ZFPM2 Promoter Methylation in Pre-operative Blood Samples
Is Associated with Post-Operative Atrial Fibrillation After Cardiac Surgery

by

Matthew Adam Fischer
Master of Science in Clinical Research
University of California, Los Angeles, 2017
Professor Janet S Sinsheimer, Chair

Post-operative atrial fibrillation (POAF) is a common cause of morbidity and mortality after cardiac surgery. Although clinical risk factors exist, their accuracy is insufficient for widespread clinical use. We enrolled 98 adult patients scheduled for cardiac surgery. Clinical risk factors for POAF were assessed from a history and physical. Occurrence of POAF was observed prior to discharge. DNA methylation data were obtained from whole blood prior to surgery. Statistical significance of POAF was analyzed for every CpG after controlling for smoking, sex, ethnicity and cell type composition with ReFACTor and CEllFi. Four CpGs met genome wide significance for association with POAF after accounting for covariates and cell type composition with both ReFACTor and CEllFi. Two of these CpGs are within the promotor of ZFPM2. Further investigation is needed to validate this finding in a separate validation cohort and to elucidate the mechanism of ZFPM2 and its involvement in atrial arrhythmias.

The thesis of Matthew Adam Fischer is approved.

Aman Mahajan

Thomas Vondriska

David Elashoff

Janet S Sinsheimer, Committee Chair

University of California, Los Angeles

2017

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ACKNOWLEDGEMENTS

I would like to thank Emma Monte, PhD, Todd Kimball, Maximilian Cabaj, Dennis Montoya, PhD, Matteo Pellegrini PhD, David Elashoff PhD, Kimberly Howard-Quijano MD MS, Aman Mahajan MD PhD and Thomas Vondriska PhD for their assistance, guidance and mentorship with this research project. This research is supported by Society of Cardiovascular Anesthesiologists Starter Grant #20164928.

I. INTRODUCTION

Post-operative atrial fibrillation (POAF) is a common and significant cause of morbidity and mortality after cardiac surgery. The incidence of POAF after coronary artery bypass grafting (CABG) is between 27 and 40%. Although there have been many improvements in cardiac surgery over the past 20 years, the incidence of POAF has remained relatively constant. Alarmingly POAF is also associated with devastating complications such as acute renal failure, stroke and congestive heart failure as well as increased hospital resource utilization[1]. Although many clinical risk factors for POAF are known, predictive models of this common perioperative complication have thus far been inadequate for widespread clinical use.

Given the limited prediction of clinical risk factors, we hypothesize that susceptibility to POAF is regulated by genetic and epigenetic risk factors. Kolek et al. investigated several single nucleotide polymorphisms (SNPs) to improve prediction of POAF after CABG but their addition did not improve the accuracy of clinical risk factor models[2]. In the study of epigenomics, DNA methylation has emerged as an important regulator of disease due to its integration of genetic and environmental factors. However, the ability of epigenomic marks to serve as biomarkers for perioperative complications is unknown. Given that several studies have correlated cardiovascular disease with differences in DNA methylation[3,4], we hypothesize that a combination of clinical risk factors and differential DNA methylation reflect individual susceptibility to POAF.

Epigenomics is the study of the effects of chromatin modification on gene expression and ultimately phenotype[5,6]. In this study, we focus on DNA methylation, a subset of epigenomics, at cytosine nucleotides contiguous with guanine (CpG). Differences in DNA methylation alter the function of DNA and can be associated with changes in gene expression and regulation. Each individual's epigenome changes with age and these differences in DNA methylation have been shown to correlate with disease incidence and progression. Epigenomic studies have revealed altered DNA methylation in human cardiomyopathy[7,8] and mouse studies

indicate that DNA methylation can be predictive of susceptibility to cardiac pathology prior to environmental stress[4]. This study seeks to identify DNA methylation sites that are associated with POAF and identify differentially methylated loci to aid in its prediction.

II. METHODS

This research was performed as a prospective cohort study because of the dynamic nature of DNA methylation. Patients are enrolled in this IRB approved study pre-operatively after informed consent. A detailed history and physical is performed before surgery that prospectively records known clinical risk factors for POAF such as age, proposed valve surgery and history of paroxysmal atrial fibrillation.

The primary outcome in this study is POAF, which is defined as any occurrence of atrial fibrillation observed on EKG for at least 30 seconds from the completion of surgery until the patient is discharged from the hospital. Patients in the cardiac ICU have continuous electrocardiogram monitoring to aid in the electronic and clinical diagnosis of atrial arrhythmias. To ensure data validity, the occurrence of POAF is cross-checked with the Society of Thoracic Surgeon (STS) database, which includes outcome data including occurrence of POAF recorded by the surgical team prior to discharge.

The primary predictor of POAF in this study is DNA methylation. For each patient, whole blood samples (approximately 6 ml) are collected from an arterial line after induction of anesthesia but prior to surgical incision. DNA is extracted from the whole blood samples and analyzed by reduced representative bisulfite sequencing (RRBS) which provides single-base resolution of DNA methylation. Analysis was restricted to CpGs for which there were at least 10 reads in an individual patient. CpGs for patients who did not have at least 10 reads were considered missing data. Whole blood was chosen for DNA methylation analysis because it is easy and safe to collect for assessment of perioperative risk prior to surgery.

The clinical covariates included in this study are age, history of paroxysmal atrial fibrillation, proposed valve surgery, proposed MAZE procedure, proposed pulmonary vein

isolation and post-operative prophylactic amiodarone. These covariates were chosen because they are all known clinical risk factors for POAF based on previous studies[1,8] or are related to the prevention of POAF. Methylation covariates included in this study are cell type composition, smoking status, sex and ethnicity. All of these covariates except cell type composition are obtained prospectively from a history and physical prior to surgery.

Correcting for cell type composition was performed via two separate methods and their results were compared to reduce confounding. The ReFACTor algorithm[9] was used to calculate surrogate estimates (ReFACTor components) of the 6 major cell type components of whole blood in an unsupervised manner using CpGs that are present in at least 80% of samples. The beta value (percent methylation) for each CpG was calculated using the RRBS methylated and unmethylated count data. Next, Pearson correlation of each ReFACTor components with every CpG was determined. CpGs that had less than 0.2 correlation with all ReFACTor components were determined to have low cell type association and their count data were later analyzed for association with POAF.

Cell type composition was also calculated using CEllFi[10], a tool developed by the Pelligrini lab at UCLA, that estimates cell type composition in whole blood using methylation sites highly associated with cell type as determined from FACS. Cell type estimates were calculated using CEllFi from the beta values of CpGs present in all samples. These cell type estimates were later used as covariates in modeling differential methylation related to POAF.

The ReFACTor components were used to exclude CpGs associated with cell type composition rather than used as covariates in modeling differential methylation because the statistical significance of individual CpGs were assessed with a count based model. This count-based model is necessary to account for the differences in read count between different samples that occur in RRBS. ReFACTor components are established using PCA on a correlation matrix of beta values from a subset of CpGs and, although highly associated with cell type, are not a measurable phenomenon directly related to methylation counts. The cell

type estimates from CEllFi, however, are analogous to cell composition percentages obtained from FACS and therefore were used as covariates in the count based model to assess for differential methylation.

Statistical significance of CpGs with POAF was determined using DSS general[11] in R. This package models the methylation count data with a beta binomial distribution and incorporates differences in depth of coverage between CpGs. First, every CpG was filtered for prevalence of at least 50% in all patients. For the cell type correction with ReFACTor, differential methylation at CpGs with low ReFACTor correlation (Pearson correlation less than 0.2) was modeled after correcting for sex, smoking, ethnicity, age, history of paroxysmal atrial fibrillation, valve surgery, MAZE procedure, pulmonary vein isolation, and prophylactic amiodarone. For the analysis with cell type correction with CEllFi, differential methylation was modeled at each CpG after correcting for sex, smoking, ethnicity, age, history of paroxysmal atrial fibrillation, valve surgery, MAZE procedure, pulmonary vein isolation, prophylactic amiodarone and the percent composition of the six most common whole blood cell types. Manhattan and Q—Q plots were generated with the QQman package in R.

Simple predictive models were constructed with forward and backward stepwise multiple logistic regression separately with minimum BIC and AIC using JMP. Results for each model were assessed using 5-fold cross validation repeated 10 times using the caret package in R.

III. RESULTS

Ninety-eight patients were prospectively enrolled in this study and nine patients in total were excluded leaving eighty-nine patients for analysis. Six patients were excluded due to complete heart block post-operatively requiring permanent pacemaker placement. Two patients were excluded for not requiring cardiopulmonary bypass. One patient was excluded because they died on post-operative day 2. Among the eighty-nine patients analyzed, thirty-eight experienced POAF for an incidence of 42.7% percent.

Clinical covariates are shown in table 1. Manhattan plots using ReFACTor (Figure 1) and CEllFi (Figure 3) to account for cell type composition demonstrate association of CpGs with POAF. Q-Q plots showing reasonable control of statistical inflation (Figures 2 and 4). The list of CpGs meeting genome-wide significance (p value less than 5×10^{-8}) are listed in tables 2 and 3 for ReFACTor and CEllFi cell type composition correction, respectively. CpGs chr4_38664639, ch8_105670444, ch8_105670450 and chr21_45785189 met genome-wide statistical significance with both CEllFi and ReFACTor cell type composition correction. Ch8_105670444 and Ch8_105670450 are within 6 basepairs of each other in the promotor region of Zinc Finger Protein, FOG Family Member 2 (ZFPM2).

Summary statistics for the POAF prediction models can be found in tables 4 and 5. The AUC, sensitivity and specificity of the predictive model using age, history of paroxysmal atrial fibrillation were 0.77, 0.82 and 0.65, respectively. The AUC, sensitivity and specificity of the predictive model using age, history of paroxysmal atrial fibrillation were 0.77, 0.78 and 0.62, respectively.

Increased methylation at Ch8_105670444 and Ch8_105670450 in the promoter region of ZFPM2 is associated with increased risk of POAF. Average methylation at Ch8_105670444 and Ch8_105670450 for patients who experience POAF are 74.6% and 80% respectively (Table 6). Average methylation at Ch8_105670444 and Ch8_105670450 for patients who do not experience POAF are 60.4% and 65.8% respectively. Histograms for the distribution of percent methylation of Ch8_105670444 and Ch8_105670450 in patients who experience POAF versus those who do not can be found in supplementary figures 1 and 2.

IV. DISCUSSION

This research implicates promoter methylation of ZFPM2 as a risk factor for post-operative atrial fibrillation. Interestingly, ZFPM2 is expressed in the heart but not in bone marrow. These methylation sites discovered in whole blood may reflect cell type non-specific changes in methylation associated with gene expression in the heart. ZFPM2 promoter

methylation has recently been associated with Tetralogy of Fallot[12] though there were no patients in this study with Tetralogy of Fallot (TOF) or any other significant congenital heart disease. Patients with Tetralogy of Fallot had a mean methylation of 80.32% compared to controls with 59.6% methylation[12]. Our results show that patients who experience POAF have 75% and 80% methylation at the studied loci and 60% and 65% methylation in the patients who do not experience POAF. ZFPM2 promoter methylation could reflect myocardial stress or high intracardiac filling pressures that would be present in both patients with TOF and patients who experience post-operative atrial fibrillation. The role of ZFPM2 in atrial arrhythmias is unknown and it is also unknown whether ZFPM2 plays a direct role in atrial fibrillation or is a related risk factor. ZFPM2 methylation is not associated with known clinical risk factors for POAF. Further studies will have to elucidate the mechanism of ZFPM2 related to cardiac arrhythmias. In addition, these findings should be confirmed in a separate validation cohort.

This result suggests genomic risk factors for post-operative complications that can be assessed before surgery. Importantly, DNA methylation samples of whole blood are easy to obtain and confer low risk to the patient. This methodology can be further analyzed to predict other perioperative complications. Prior to surgery patients could obtain whole blood DNA methylation risk for post-operative complications along with their other standard pre-operative labs. For example, patients with low traditional surgical risk who have methylation biomarkers high risk for POAF, could be referred for transcatheter aortic valve replacement (TAVR) instead of open aortic valve replacement (AVR) to minimize the risk of POAF given the risk of POAF is lower after TAVR[13]. In addition, the increased cost of TAVR over AVR may be somewhat offset by the increased length of stay seen in patients who experience POAF after AVR. The care for each patient could be personalized to that patients individual risks and care could be altered to minimize adverse events.

V. FIGURES AND TABLES

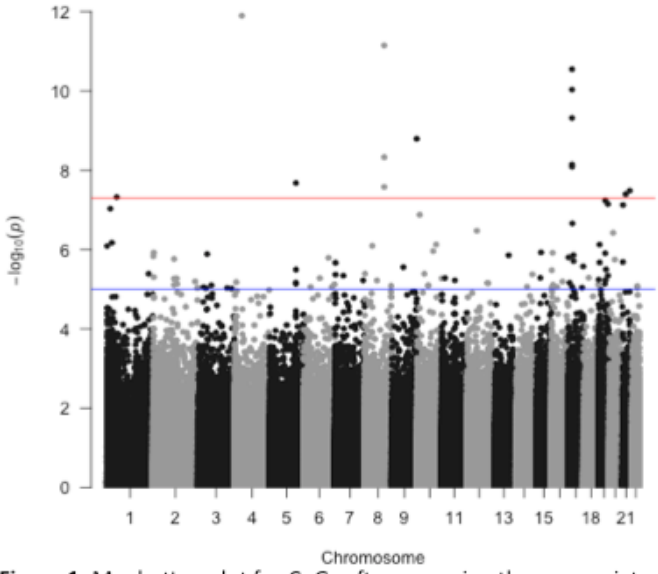


Figure 1: Manhattan plot for CpGs after removing those associated with cell type composition using ReFACTor. The above manhattan plot shows the $-\log_{10}$(p-value) of each CpG with respect to their position on the 22 somatic chromosomes. The blue and red horizontal lines correspond to 10^{-5} and 5×10^{-8}, respectively.

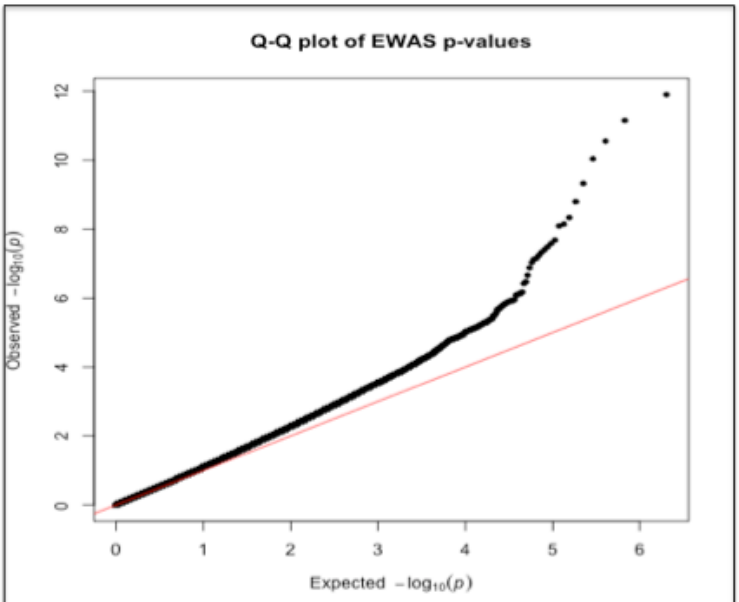


Figure 2: Q-Q plot for CpGs association with POAF after removing those associated with cell type composition using ReFACTor. The horizontal axis shows the expected $-\log_{10}$(p-value) versus the observed $-\log_{10}$(p-values).

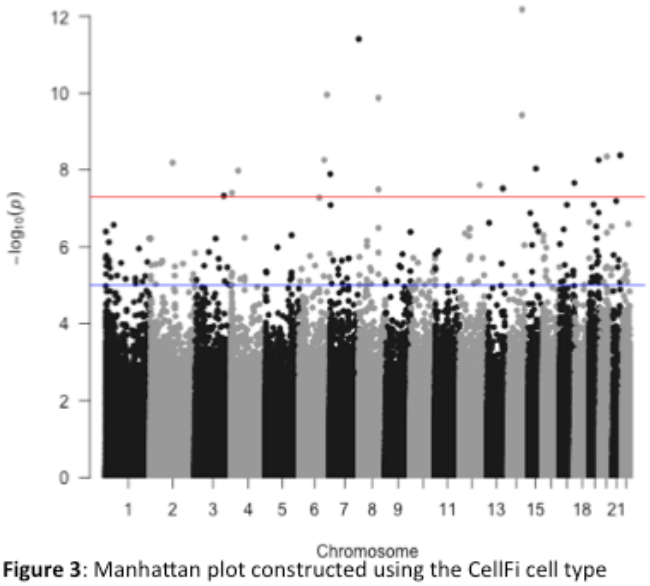


Figure 3: Manhattan plot constructed using the CellFi cell type composition estimates. This manhattan plot shows the $-\log_{10}$(p-value) of each CpG with respect to their position on the 22 somatic chromosomes. The blue and red horizontal lines correspond to 10^{-5} and 5×10^{-8}, respectively.

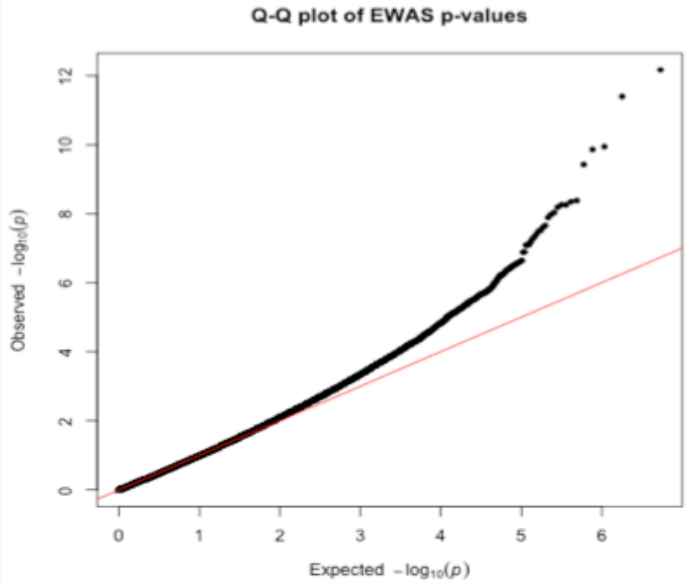


Figure 4: Q-Q plot for CpGs association with POAF after accounting for cell type composition using CellFi. The horizontal axis shows the expected $-\log_{10}$(p-value) versus the observed $-\log_{10}$(p-values).

Characteristic	POAF	No POAF
Total Number	37	52
Age, mean	66.24	60.40
Male	27 (73%)	38 (73.1%)
History of Paroyxsmal Atrial Fibrillation	9 (24.3%)	3 (5.8%)
Valve Surgery	26 (70.3%)	35 (67.3%)
Current Smoker	3 (8.1%)	6 (11.5%)
PVI	2 (5.4%)	1 (1.9%)
MAZE	4 (10.8%)	2 (3.8%)

Table 1: Comparison of clinical covariates between those who experienced POAF and those who did not. These covariates were determined pre-operatively using a history and physical.

CpG	P - Value	Prevalence	CpG	P - Value	Prevalence
chr1_52609391	4.74E-08	0.522	chr17_20376853	7.17E-09	0.543
chr4_38664639	1.27E-12	0.620	chr17_20376867	8.12E-09	0.533
chr5_141123637	2.09E-08	0.609	chr17_20376899	9.26E-11	0.554
chr8_105670444	4.66E-09	0.685	chr17_20376904	2.83E-11	0.554
chr8_105670450	7.14E-12	0.685	chr17_20376906	4.78E-10	0.543
chr8_105670482	2.62E-08	0.685	chr21_23390483	4.01E-08	0.989
chr9_135786264	1.60E-09	0.522	chr21_45785189	3.26E-08	0.630

Table 2: CpGs that meet genome-wide significance for association with POAF after removing those associated with cell type composition from the ReFACTor components. Prevalence refers to how many patients have at least 10 reads of that CpG present.

CpG	P - Value	Prevalence	CpG	P - Value	Prevalence
chr2_119679375	6.47E-09	0.783	chr12_108931782	2.49E-08	0.533
chr3_158104032	4.69E-08	0.804	chr13_102010785	3.03E-08	1.000
chr4_38664639	1.05E-08	0.620	chr14_92121624	3.76E-10	0.707
chr4_4227265	3.97E-08	0.641	chr14_93041538	6.58E-13	0.609
chr6_138107957	5.50E-09	0.663	chr15_62186844	9.23E-09	0.587
chr6_152816684	1.11E-10	0.554	chr17_80731172	2.18E-08	0.793
chr7_157769280	3.93E-12	0.511	chr19_50316280	5.50E-09	0.511
chr7_590837	1.29E-08	0.576	chr20_36539184	4.50E-09	0.522
chr8_105670444	3.21E-08	0.685	chr21_45785189	4.18E-09	0.630
chr8_105670450	1.34E-10	0.685			

Table 3: CpGs that meet genome-wide significance for association with POAF after correcting for cell type composition with CellFi. Prevalence refers to how many patients have at least 10 reads of that CpG present.

A.

Factor	OR	95% CI	P-Value
Age (per 10 years)	1.68	(1.05, 2.71)	0.0322
History of Paroxysmal Atrial Fibrillation	5.40	(1.03, 28.44)	0.0465
chr8_105670450 per 0.10 methylation	2.32	(1.42, 3.79)	0.0007

B.

Statistic	Result
ROC	0.77
Sensitivity	0.82
Specificity	0.65

Table 4: Risk factors (A) and summary statistics (B) for the POAF model incorporating age and chr8_105670450 as predictors of POAF. This model was selected using stepwise regression with minimum BIC and tested with 5-fold, 10x repeated cross validation.

A.

Factor	OR	95% CI	P-Value
Age (per 10 years)	1.74	(1.08, 2.8)	0.0235
History of Paroxysmal Atrial Fibrillation	4.93	(0.9, 26.93)	0.0654
Valve Surgery	2.53	(0.82, 7.76)	0.1046
chr8_105670450 per 0.10 methylation	2.65	(1.55, 4.54)	0.0004

B.

Statistic	Result
ROC	0.77
Sensitivity	0.78
Specificity	0.62

Table 5: Risk factors (A) and summary statistics (B) for the POAF model incorporating age, history of paroxysmal atrial fibrillation, valve surgery and chr8_105670450 as predictors of POAF. This model was selected using stepwise logistic regression with minimum AIC and tested with 5-fold, 10x repeated cross validation.

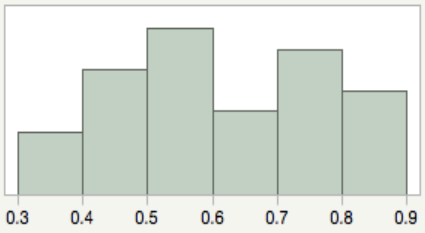
	chr8_105670444	chr8_105670450
POAF	74.60%	80.00%
No POAF	60.40%	65.80%

Table 6: Mean percent methylation at CpGs in the promoter region of ZFPM2 showing increased methylation associated with POAF.

VI. APPENDIX

Distributions POAF=0

chr8_105670444

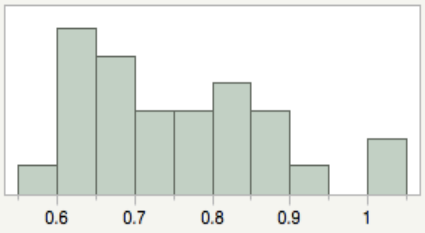


Quantiles		
100.0%	maximum	0.85
99.5%		0.85
97.5%		0.85
90.0%		0.832
75.0%	quartile	0.755
50.0%	median	0.58
25.0%	quartile	0.46
10.0%		0.4
2.5%		0.31
0.5%		0.31
0.0%	minimum	0.31

Summary Statistics	
Mean	0.6036364
Std Dev	0.1569398
Std Err Mean	0.0273197
Upper 95% Mean	0.6592848
Lower 95% Mean	0.5479879
N	33

Distributions POAF=1

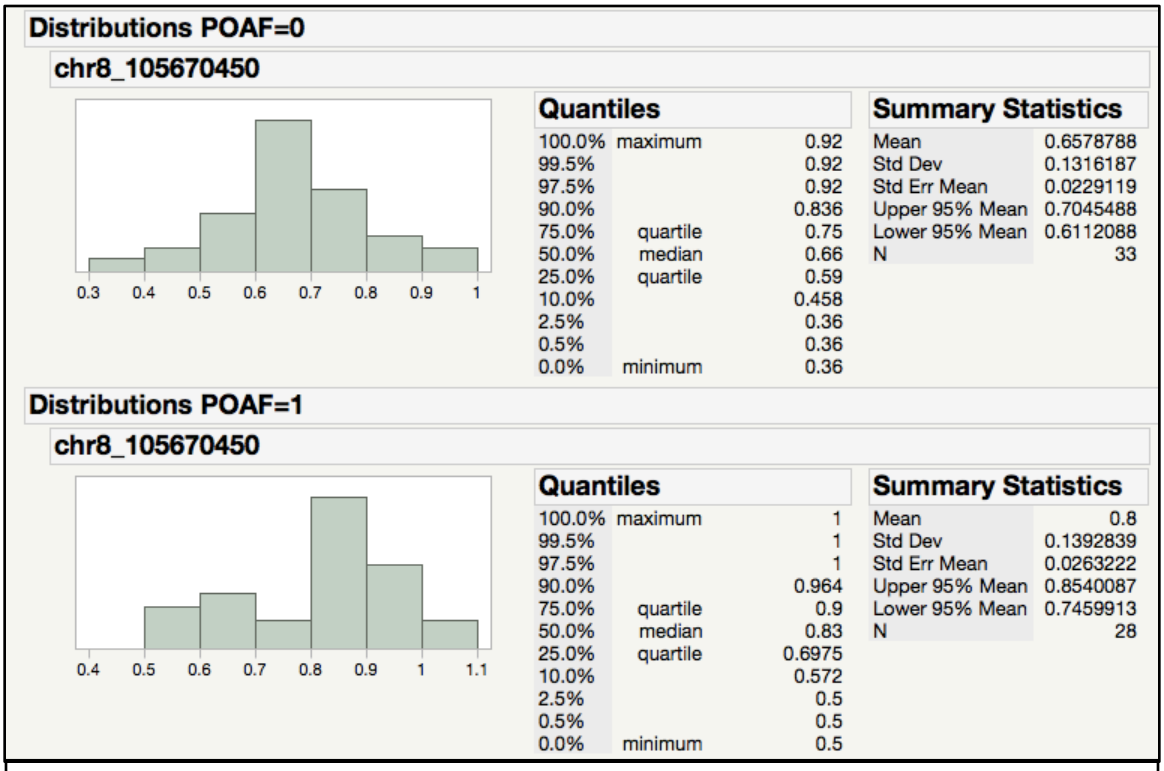
chr8_105670444



Quantiles		
100.0%	maximum	1
99.5%		1
97.5%		1
90.0%		0.928
75.0%	quartile	0.8275
50.0%	median	0.715
25.0%	quartile	0.6425
10.0%		0.6
2.5%		0.58
0.5%		0.58
0.0%	minimum	0.58

Summary Statistics	
Mean	0.7460714
Std Dev	0.1180188
Std Err Mean	0.0223035
Upper 95% Mean	0.7918343
Lower 95% Mean	0.7003085
N	28

Supplementary Figure 1: Histogram showing distribution of percent methylation among patients who do not experience POAF (above) and patients who do experience POAF (below).



Supplementary Figure 2: Histogram showing distribution of percent methylation among patients who do not experience POAF (above) and patients who do experience POAF (below).

VII. REFERENCES:

1. Matthew JP et al. “A Multicenter Risk Index for Atrial Fibrillation After Cardiac Surgery.” *Journal of the American Medical Association*. 2004. 291(14): 1720-1729
2. Kolek, MJ et al. “Genetic and Clinical Risk Prediction Model for Postoperative Atrial Fibrillation.” Circulation Arrhythmia Electrophysiology. 2015. 8(1): 25-31
3. Movassagh M et al. “Distinct Epigenomic Features in End-Stage Failing Human Hearts.” *Circulation*. 2011. 124(22): 2411-2422
4. Chen H et al. “DNA Methylation Indicates Susceptibility to Isoproterenol-Induced Cardiac Pathology and is Associated with Chromatin States.” Circ Res. 2016. In Press
5. Rosa-Garrido, M et al. “Regulation of Chromatin Structure in the Cardiovascular System.” Circ J. 2013; 77(6): 1389-1398
6. Han, P et al. “Chromatin remodeling in cardiovascular development and physiology.” Circ Res. 2011; 108(3): 378-396
7. Orozco L et al. “Epigenome-wide Association of Liver Methylation Patterns and Complex Metabolic Traits in Mice.” *Cell Metabolism*. 2015. 21(6): 905-917
8. Benjamin, E et al. “Independent Risk Factors for Atrial Fibrillation in a Population-Based Cohort.” *JAMA.* 1994; 271(11): 840-844
9. Rahmani, E et al. "Sparse PCA Corrects for Cell Type Heterogeneity in Epigenome-Wide Association Studies". Nature Methods. 2016. 13(5): 443-445
10. Lam, L et al. “CEllFi: Cellular epigenetic fingerprint identification for computational reconstruction of the immune cellular microenvironment”. Selected Oral Presentation. UCLA Quantitative and Computational Biology (QCBio) Annual Symposium. 2017. Los Angeles, CA.
11. Park, Y et al. “Differential Methylation Analysis for BS-seq Data Under General Experimental Design.” Bioinformatics. 2016. 15; 32(10): 1446-1453

12. Sheng, W et al. “CpG Island Shore Methylation of ZFPM2 is Identified in Tetralogy of Fallot Samples.” *Pediatric Research*. 2016. 80(1):151-158

13. Reardon, M et al. “Surgical or Transcatheter Aortic-Valve Replacement in Intermediate-Risk Patients.” *New England Journal of Medicine*. 2017. 376:1321-1331