

Examining relationships between long non-coding RNA and warfarin
pharmacogenes

Zylee Liddy

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2023

Committee:

Edward J. Kelly, Chair

Kenneth E. Thummel

Program Authorized to Offer Degree:

Public Health Genetics

©Copyright 2023

Zylee Liddy

University of Washington

Abstract

Examining relationships between long non-coding RNA and warfarin pharmacogenes

Zylee Liddy

Chair of the Supervisory Committee:

Edward J. Kelly

Department of Pharmaceutics

Inter-individual variability in warfarin metabolism and response is influenced by multiple factors, including genetic variations in important pharmacogenes such as *VKORC1* and *CYP2C9*. LncRNA (long non-coding RNA) plays roles in diverse cellular processes including gene regulation. This study explored the relationship between the hepatic expression of lncRNA genes and warfarin pharmacogenes through correlation analysis and penalized linear regression modeling. Significant correlations emerged between specific lncRNA genes and warfarin pharmacogenes. The Lasso penalized linear regression model also highlighted the lncRNA genes impacting the variability in warfarin pharmacogene expression. Despite study limitations, these novel results provide a basis for further exploration of these relationships and their underlying biological mechanisms. Understanding these connections could potentially lead to improved warfarin dosing for patients on the medication, enhancing both patient safety and drug efficacy.

Integrating pharmacogenomics and lncRNA research holds promise for refining drug dosing strategies and advancing personalized medicine.

ACKNOWLEDGEMENTS

I would like to first thank Dr. Edward Kelly and Dr. Kenneth Thummel for affording me the opportunity to contribute on this project. Special thanks go to Dr. Katrina Claw for generously providing the data used in this endeavor.

My heartfelt gratitude extends to my committee chair, Dr. Kelly, whose valuable guidance, patience, and dedication to my academic and personal growth have been instrumental in shaping the trajectory of this work.

I am also sincerely grateful to the biostatistician, James MacDonald, for his advice and support was instrumental in overcoming the challenges encountered during my data analysis process. His expertise not only challenged me to think critically but it also helped me further develop my data analysis skills.

A sincere acknowledgment goes out to my peers and colleagues in the Kelly Lab and the Institute of Public Health Genetics. Your encouragement and collaborative spirit have made this process both enjoyable and intellectually enriching.

Finally, to my family and friends, I am deeply thankful for their unwavering support and steadfast belief in my abilities. Your encouragement through every step of the process has been a constant motivation.

Completing this project would not have been possible without the combined efforts of all those mentioned above. Thank you for your invaluable contributions.

TABLE OF CONTENTS

List of Figures.....	iv
List of Tables.....	x
Gene Abbreviations.....	xvi
Chapter 1. Introduction to Warfarin Pharmacogenes.....	1
1.1 Pharmacogenomics.....	1
1.2 Warfarin.....	2
1.3 Warfarin Pharmacogenes.....	4
1.3.1 <i>VKORC1</i>	4
1.3.2 <i>CYP2C9</i>	5
1.3.3 Other Warfarin Pharmacogenes.....	6
Chapter 2. Introduction to Long Non-Coding RNA.....	8
Chapter 3. Aim and Objectives.....	9
Chapter 4. Methods.....	10
4.1 RNA Sequencing Data.....	10
4.2 Warfarin Pharmacogene Data.....	10
4.3 Long Non-Coding RNA Data.....	10
4.4 Statistical Analyses.....	11
Chapter 5. Results.....	14
Chapter 6. Discussion.....	26
6.1 Limitations.....	30
6.2 Future Directions.....	32
References.....	35

Appendix A.....	42
Appendix B.....	51
Appendix C.....	126

LIST OF FIGURES

Figure 1.2.1. Chemical structure of (S)-warfarin on the left and (R)-warfarin on the right. Images obtained from PubChem. ^{1,2}	2
Figure 1.3.3.1. <i>VKORC1</i> and several other known or suspected genes associated with warfarin pharmacodynamics. Image obtained from PharmGKB. ³	7
Figure 1.3.3.2. <i>CYP2C9</i> and several other known or suspected genes associated with warfarin pharmacokinetics. Image obtained from PharmGKB. ³	7
Figure 5.1. Q-Q plot of original warfarin pharmacogenes FPKM values.....	15
Figure 5.2. Q-Q plot of original lncRNA FPKM values.....	16
Figure 5.3. Q-Q plot of log ₂ -transformed warfarin pharmacogene FPKM values.....	17
Figure 5.4. Q-Q plot of log ₂ -transformed lncRNA FPKM values with outliers included.....	18
Figure 5.5. Q-Q plot of log ₂ -transformed lncRNA FPKM values with outliers removed.....	18
Figure 5.6. Q-Q plot of log ₂ -transformed and scaled warfarin pharmacogene FPKM values.....	19
Figure 5.7. Q-Q plot of log ₂ -transformed and scaled lncRNA FPKM values with outliers included.....	20
Figure 5.8. Q-Q plot of log ₂ -transformed and scaled lncRNA FPKM values with outliers removed.....	20
Figure 5.9. Heatmap of all Pearson correlation coefficients.....	21
Figure 5.10. Heatmap of significant Pearson correlation coefficients (Bonferroni corrected $\alpha = 2.817e-05$).....	23
Figure B.1. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene <i>ABCB1</i>	53

Figure B.2. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *APOB*.....56

Figure B.3. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *APOCI*.....59

Figure B.4. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *APOE*.....62

Figure B.5. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *ASPH*.....65

Figure B.6. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CALU*.....68

Figure B.7. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CRP*.....71

Figure B.8. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP1A1*.....74

Figure B.9. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP1A2*.....77

Figure B.10. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP2A6*.....80

Figure B.11. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP2C18*.....83

Figure B.12. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP2C19*.....86

Figure B.13. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP2C9*.....89

Figure B.14. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP3A4*.....92

Figure B.15. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP4F11*.....95

Figure B.16. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP4F2*.....98

Figure B.17. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *DDHD1*.....101

Figure B.18. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *DNMT3A*.....104

Figure B.19. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *EPHX1*.....107

Figure B.20. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *VKORC1*.....110

Figure B.21. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F11*.....113

Figure B.22. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F13A1*.....116

Figure B.23. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F2*.....119

Figure B.24. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F5*.....122

Figure B.25. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F7*.....125

Figure C.1. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against all warfarin pharmacogenes as one group.....128

Figure C.2. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against 12 warfarin pharmacogenes (*APOC1*, *APOE*, *CYP1A2*, *CYP2A6*, *CYP2C19*, *CYP2C9*, *CYP3A4*, *CYP4F2*, *EPHX1*, *F5*, *F7*, and *F11*).....131

Figure C.3. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against 11 warfarin pharmacogenes (*ABCB1*, *APOB*, *ASPH*, *CYP1A1*, *CYP2C18*, *CYP4F11*, *DDHD1*, *DNMT3A*, *F13A1*, *F2*, and *VKORC1*).....134

Figure C.4. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against two warfarin pharmacogenes (*CALU* and *CRP*).....137

Figure C.5. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against nine warfarin pharmacogenes (*CYP1A2*, *CYP2A6*, *CYP2C9*, *CYP2C19*, *CYP3A4*, *CYP4F2*, *EPHX1*, *F7*, and *F11*).....140

Figure C.6. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against three warfarin pharmacogenes (*APOC1*, *APOE*, and *F5*).....143

Figure C.7. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against five warfarin pharmacogenes (*APOB*, *ASPH*, *CYP2C18*, *CYP4F11*, and *F2*).....146

Figure C.8. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against two warfarin pharmacogenes (*ABCB1* and *CYP1A1*).....149

Figure C.9. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against four warfarin pharmacogenes (*DDHD1*, *DNMT3A*, *F13A1*, and *VKORC1*).....152

Figure C.10. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against two warfarin pharmacogenes (*CALU* and *CRP*).....155

LIST OF TABLES

Table 5.1. Liver Bank Donor Demographics.....	14
Table 5.2 The groups of warfarin pharmacogenes tested in the Lasso, Elastic Net, and Ridge penalized linear regression models as informed by hierarchical clustering in the correlation heatmap.....	25
Table A.1. Summary statistics and standard deviations for the 25 warfarin pharmacogenes remaining after quality control.....	42
Table A.2. Summary statistics and standard deviations for the 71 lncRNA genes remaining after quality control.....	43
Table A.3. Summary statistics and standard deviations of 25 log ₂ -transformed warfarin pharmacogenes.....	45
Table A.4. Summary statistics and standard deviations of 71 log ₂ -transformation lncRNA genes.....	46
Table A.5. Summary statistics and standard deviations of 25 log ₂ -transformed and scaled warfarin pharmacogenes.....	48
Table A.6. Summary statistics and standard deviations of 71 log ₂ -transformed and scaled lncRNA genes.....	49
Table B.1. LncRNA genes selected as predictors to explain the expression of warfarin pharmacogene <i>ABCB1</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	51
Table B.2. LncRNA genes selected as predictors to explain the expression of warfarin pharmacogenes <i>APOB</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	54

Table B.3. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *APOC1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....57

Table B.4. LncRNA genes selected as predictors to explain the expression of warfarin pharmacogene *APOE* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....60

Table B.5. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *ASPH* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....63

Table B.6. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CALU* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....66

Table B.7. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CRP* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....69

Table B.8. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP1A1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....72

Table B.9. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP1A2* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....75

Table B.10. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene <i>CYP2A6</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	78
Table B.11. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene <i>CYP2C18</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	81
Table B.12. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene <i>CYP2C19</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	84
Table B.13. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene <i>CYP2C9</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	87
Table B.14. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene <i>CYP3A4</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	90
Table B.15. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene <i>CYP4F11</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	93
Table B.16. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene <i>CYP4F2</i> in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....	96

Table B.17. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *DDHDI* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....99

Table B.18. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *DNMT3A* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....102

Table B.19. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *EPHX1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....105

Table B.20. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *VKORC1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....108

Table B.21. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F11* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....111

Table B.22. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F13A1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....114

Table B.23. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F2* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....117

Table B.24. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F5* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....120

Table B.25. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F7* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....123

Table C.1. LncRNA genes selected as predictors to explain the expression of all 25 warfarin pharmacogenes as one group in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....126

Table C.2. LncRNA genes selected as predictors to explain the expression of 12 warfarin pharmacogenes (*APOC1*, *APOE*, *CYP1A2*, *CYP2A6*, *CYP2C19*, *CYP2C9*, *CYP3A4*, *CYP4F2*, *EPHX1*, *F5*, *F7*, and *F11*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....129

Table C.3. LncRNA genes selected as predictors to explain the expression of 11 warfarin pharmacogenes (*ABCB1*, *APOB*, *ASPH*, *CYP1A1*, *CYP2C18*, *CYP4F11*, *DDHD1*, *DNMT3A*, *F13A1*, *F2*, and *VKORC1*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....132

Table C.4. LncRNA genes selected as predictors to explain the expression of two warfarin pharmacogenes (*CALU* and *CRP*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....135

Table C.5. LncRNA genes selected as predictors to explain the expression of nine warfarin pharmacogenes (*CYP1A2*, *CYP2A6*, *CYP2C9*, *CYP2C19*, *CYP3A4*, *CYP4F2*, *EPHX1*, *F7*, and *F11*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....138

Table C.6. LncRNA genes selected as predictors to explain the expression of three warfarin pharmacogenes (*APOC1*, *APOE*, and *F5*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....141

Table C.7. LncRNA genes selected as predictors to explain the expression of five warfarin pharmacogenes (*APOB*, *ASPH*, *CYP2C18*, *CYP4F11*, and *F2*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....144

Table C.8. LncRNA genes selected as predictors to explain the expression of two warfarin pharmacogenes (*ABCB1* and *CYP1A1*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....147

Table C.9. LncRNA genes selected as predictors to explain the expression of four warfarin pharmacogenes (*DDHD1*, *DNMT3A*, *F13A1*, and *VKORC1*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....150

Table C.10. LncRNA genes selected as predictors to explain the expression of two warfarin pharmacogenes (*CALU* and *CRP*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.....153

GENE ABBREVIATIONS

ABCB1 (MDR1) – P-glycoprotein 1 (Multidrug resistance protein 1)

APOB – Apolipoprotein B

APOC1 – Apolipoprotein C1

APOE – Apolipoprotein E

ASPH – Aspartate Beta-Hydroxylase

CALU – Calumenin

CRP – C-Reactive Protein

CYP1A1 – Cytochrome P450, Family 1, Subfamily A, Member 1

CYP1A2 – Cytochrome P450, Family 1, Subfamily A, Member 2

CYP2A6 – Cytochrome P450, Family 2, Subfamily A, Member 6

CYP2C18 – Cytochrome P450, Family 2, Subfamily C, Member 18

CYP2C19 – Cytochrome P450, Family 2, Subfamily C, Member 19

CYP2C9 – Cytochrome P450, Family 2, Subfamily C, Member 9

CYP3A4 – Cytochrome P450, Family 3, Subfamily A, Member 4

CYP4F11 – Cytochrome P450, Family 4, Subfamily F, Member 11

CYP4F2 – Cytochrome P450, Family 4, Subfamily F, Member

DDHD1 – DDHD Domain Containing 1

DNMT3A – DNA Methyltransferase 3 Alpha

EPHX1 – Epoxide Hydrolase 1

F11 – Coagulation Factor XI

F13A1 – Coagulation Factor XIII A Chain

F2 – Coagulation Factor II, Thrombin

F5 – Coagulation Factor V

F7 – Coagulation Factor VII

VKORC1 – Vitamin K Epoxide Reductase Complex Subunit 1

Chapter 1. INTRODUCTION TO WARFARIN PHARMACOGENES

1.1 PHARMACOGENOMICS

There are many factors contributing to inter-individual variability in drug response that has been frequently observed. However, it is the recent advancements in science and technology that have revealed the contributions of genetics to this noted variability, successively leading to the expansion of pharmacogenomics. Pharmacogenomics is a field that aims to utilize genomics to better understand how genes and genetic variants affect drug metabolism and response. It is also a field that uses genomics to improve drug efficacy and safety.^{4,5}

Pharmacogenomics encompasses the concepts of pharmacodynamics and pharmacokinetics. Pharmacodynamics examines the drug dose-response relationship and seeks to describe a drug's mechanism of action, response, effectiveness, and safety, as well as drug-drug interactions.⁵ Pharmacokinetics, on the other hand, describes how drugs are absorbed, distributed, metabolized, and eliminated by the body.⁵ Different genes and genetic variants can affect a drug's pharmacodynamics, pharmacokinetics, or a combination of both. These can subsequently lead to a spectrum of metabolic phenotypes that include poor, intermediate, normal, fast, or ultrarapid drug metabolizers. Metabolic phenotypes in turn can create a range of drug responses that are clinically observable in patients.^{5,6}

Numerous studies have been conducted highlighting the usefulness of genetics in developing new drugs and improving the effectiveness and safety of existing drugs. For example, multiple RCTs (randomized clinical trials) have shown that warfarin therapy guided by pharmacogenomics significantly improves patient outcomes by increasing effectiveness and decreasing severe adverse events.^{5,7} These and similar studies with other drugs have increased the use and acceptance of pharmacogenomics. It has also encouraged entities like CPIC (the Clinical

Pharmacogenetics Implementation Consortium) to develop guidelines and recommendations for using genetic information to guide drug selection and dosing in clinical care.⁶ Additionally, it has prompted the FDA (U.S. Food and Drug Administration) to approve drug labels that provide recommendations on how to dose certain drugs according to a patient's genotype.^{6,8}

1.2 WARFARIN

Anticoagulants are drugs commonly used to treat and prevent thrombotic events associated with diverse etiologies.⁹ Warfarin is one of many examples of an anticoagulant that is mainly administered as an oral pill that is a racemic mixture of two enantiomers, (R)-warfarin and (S)-warfarin, as seen in Figure 1.2.1.^{1,2,9,10}

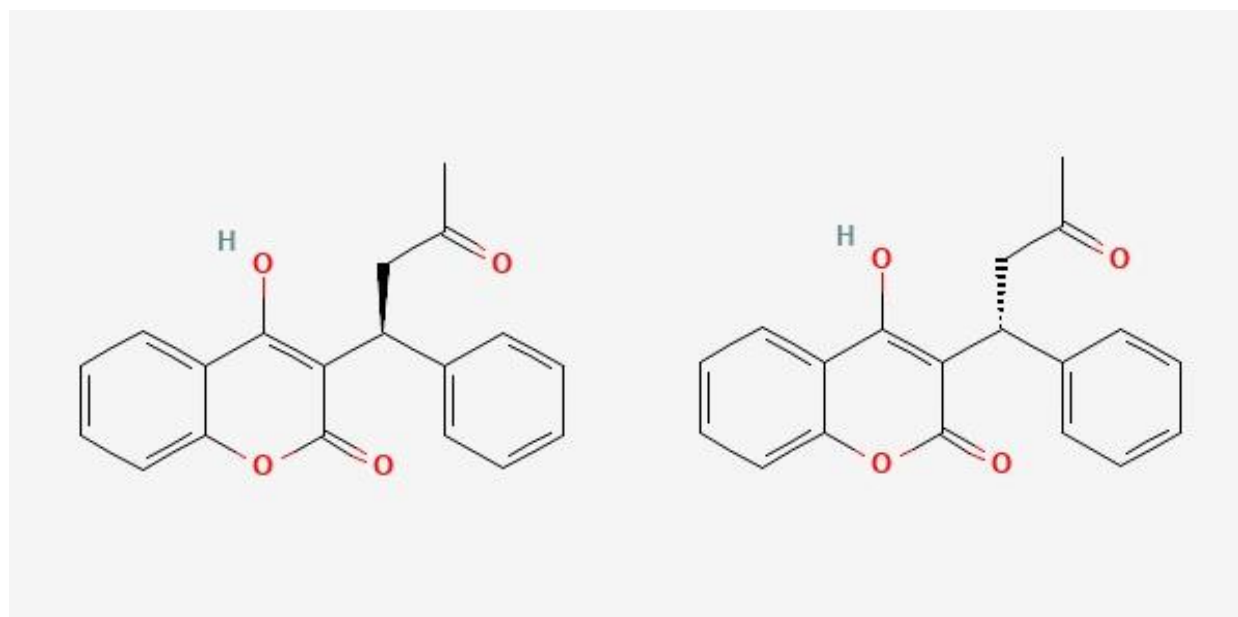


Figure 1.2.1. Chemical structure of (S)-warfarin on the left and (R)-warfarin on the right. Images obtained from PubChem.^{1,2}

Once warfarin enters a patient's bloodstream approximately 98-99% of it is bound to plasma proteins, primarily albumin.^{7,9,11} This binding allows warfarin to be easily transported through the bloodstream to target sites such as the liver where it can exert its anticoagulant effects. These effects typically occur within 24 hours after administration of the drug though peak

anticoagulation effect could be delayed up to 72 to 96 hours.^{9,11} The primary mechanism of action of warfarin is to inhibit the enzyme VKORC1 (Vitamin K Epoxide Reductase Complex Subunit 1) from recycling an oxidized form of vitamin K (epoxide) back to the enzymatically active reduced hydroquinone form of the hormone. Reduced vitamin K is an essential cofactor for the synthesis of vitamin K-dependent clotting factors and proteins.^{6,9} Most of this inhibitory drug action comes from the (S)-warfarin stereoisomer which accounts for approximately 70% of the therapeutic effects of warfarin.^{5,6,8,10,12} Conversely, levels of (R)-warfarin in patients have been observed to be approximately twice as high as (S)-warfarin, but its inhibitory effect is notably reduced by a factor of three to five.¹³

In addition to being the primary target site of action, the liver is also the site of warfarin metabolic clearance.^{6,8} Each warfarin enantiomer, (S)-warfarin and (R)-warfarin, is metabolized by the body through different pathways that utilize different enzymes which will be discussed in more detail in following sections.^{6,8,10,12} Once in a patient's body, warfarin can continue to exert its effects for nearly two to five days before it is fully metabolized and excreted.⁹

It is also important to note that warfarin is a drug with a narrow therapeutic range and inappropriate dosing can cause over- or under-anticoagulation, which could lead to severe adverse side effects such as excessive bleeding, thromboembolism, and death.^{5,6,9} In fact, warfarin is a drug that has been implicated in over 30% of hospitalizations that are due to drug-related adverse events.¹⁴⁻¹⁶ Due to this, patients who take warfarin require frequent monitoring which includes measuring their INR (International Normalized Ratio) on a regular basis. INR is a standardized measurement that compares a patient's PT (prothrombin time), which is how long it takes for their blood to clot, to a reference value.^{6,12} This INR value is used to evaluate the efficacy and safety of the patient's warfarin therapy and helps clinicians make adjustments to a patient's medication dose

as necessary for their condition.⁹ Unfortunately, many patients have to go through a long period of dose adjustments to find an adequate and safe dose. Situations like this is where pharmacogenomics can become quite useful.

1.3 WARFARIN PHARMACOGENES

A number of factors such as age, height, weight, smoking history, comorbidities, and drug interactions have been estimated to make up 30-50% of the variation seen in warfarin metabolism and pharmacological response.^{5,7,10,17,18} However, there have been an increase in the number of studies revealing various genetic factors associated with warfarin.^{5,17,18} The two major contributing genes consistently identified in these studies are *VKORC1* and *CYP2C9*. These two warfarin pharmacogenes have been estimated to explain up to 35-45% of the variability seen in warfarin dosing needed to achieve a therapeutic INR.^{7,14}

1.3.1 *VKORC1*

VKORC1 (Vitamin K Epoxide Reductase Complex Subunit 1) is a gene primarily expressed in the liver and it encodes an enzyme system that spans the membrane of the endoplasmic reticulum.^{7,12} The main responsibility of the VKOR enzyme is to convert vitamin K epoxide back into its reduced form, vitamin K hydroquinone, which is a necessary cofactor for the synthesis of coagulation factors II, VI, IX and X, and anticoagulation proteins C, S and Z.^{6,7,10,12,19} As briefly mentioned in the previous section, warfarin acts by targeting and inhibiting the activity of the VKOR enzyme which subsequently affects the synthesis of important clotting factors and proteins.^{6,10}

Previous studies have identified variants of *VKORC1* that have different effects on warfarin pharmacodynamics. A common *VKORC1* variant is c.-1639G>A which has been associated with warfarin sensitivity.^{6,20} Patients with this variant require lower starting and maintenance doses of warfarin due to increased sensitivity to the drug.⁶ There are also less common LOF (loss-of-

function) variants in *VKORC1* that have been associated with warfarin resistance.⁵ Patients with LOF variants require higher starting and maintenance doses of warfarin.⁶ Figure 1.3.3.1 depicts *VKORC1* and a few other known or suspected genes associated with warfarin pharmacodynamics.

1.3.2 *CYP2C9*

CYP2C9 (Cytochrome P450, Family 2, Subfamily C, Member 9) encodes the corresponding *CYP2C9* enzyme, which is a well-known CYP450 (cytochrome P450) involved in the metabolism of up to 15-30% of clinically used drugs.^{7,18,21} CYP450 is a superfamily of enzymes that exist in a variety of forms and can serve numerous functions. However, the CYP450 enzymes that are known to be specifically associated with warfarin are mostly involved in warfarin metabolism. As noted in a previous section warfarin is administered as an oral pill that is a racemic mixture of (R)- and (S)-warfarin enantiomers that are metabolized by the body through different pathways. *CYP2C9* is the enzyme involved in the metabolism of the more potent (S)-warfarin while (R)-warfarin is metabolized by other CYP450 enzymes: *CYP3A4*, *CYP1A1* and *CYP1A2*.^{5,6,8-10,18}

Similar to *VKOR*, *CYP2C9* is an enzyme that is abundantly expressed in the liver and can be found in the membrane of the endoplasmic reticulum.²² The *CYP2C9* enzyme is also responsible for converting 80-90% of the more potent (S)-warfarin to (S)-7-hydroxywarfarin and (S)-6-hydroxywarfarin, with majority of the inactive metabolite being (S)-7-hydroxywarfarin.^{10,12,21,22} After the formation of these relatively inactive metabolites, they continue to be further metabolized and ultimately excreted from the body.²¹

Prior studies suggest that *CYP2C9* variants account for up to 10-20% of the inter-individual variability seen in warfarin pharmacokinetics.^{12,17,19,23} There are also over 60 identified *CYP2C9* variants that vary in frequency across different populations.^{6,8,17,19} A majority of the known *CYP2C9* variants have been characterized as single nonsynonymous mutations in exons that lead

to a reduced ability to metabolism (S)-warfarin, with many found to be associated with a two to three-fold increase in the risk of bleeding from warfarin.^{6,8,14,17} Two commonly identified variants that are examples of this are *CYP2C9*2* and *CYP2C9*3*. These two variants are associated with warfarin sensitivity, therefore, patients with these variants require lower starting and maintenance doses of warfarin.^{6,10,18-20} Figure 1.3.3.2 depicts *CYP2C9* and a few other known or suspected genes associated with warfarin pharmacokinetics.

1.3.3 OTHER WARFARIN PHARMACOGENES

Although *VKORC1* and *CYP2C9* are considered the major genes contributing to warfarin pharmacodynamics and pharmacokinetics, there are a number of other genes that have also been linked with warfarin dose-response.^{8,12,15,17,24,25} The contributions and mechanisms of action for many of these known and suspected warfarin pharmacogenes are not fully understood, but any information on them may be found through PharmGKB.

PharmGKB is one of many interactive databases and data mining tools that continues to incorporate information for a growing range of pharmacogenes.³ This information includes variant annotations, pharmacodynamic and pharmacokinetic pathways, drug-pharmacogene interactions, drug response summaries, clinical annotations, drug labels, dosing guidelines, and more.³⁻⁵ With regards to warfarin, PharmGKB currently has information on 58 known or suspected warfarin pharmacogenes.³ Some of these genes and the pathways they are involved in can be visualized in Figure 1.3.3.1 and Figure 1.3.3.2.

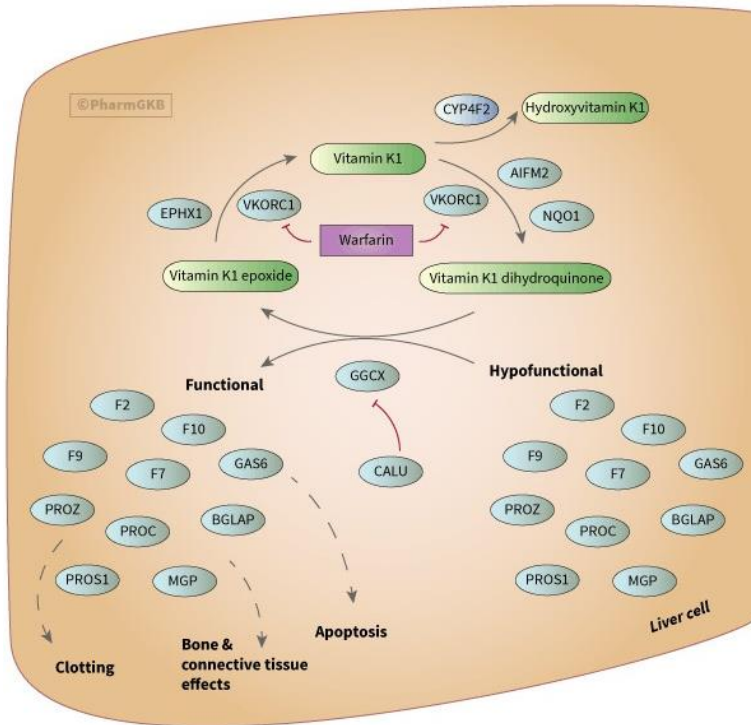


Figure 1.3.3.1. *VKORC1* and several other known or suspected genes associated with warfarin pharmacodynamics. Image obtained from PharmGKB.³

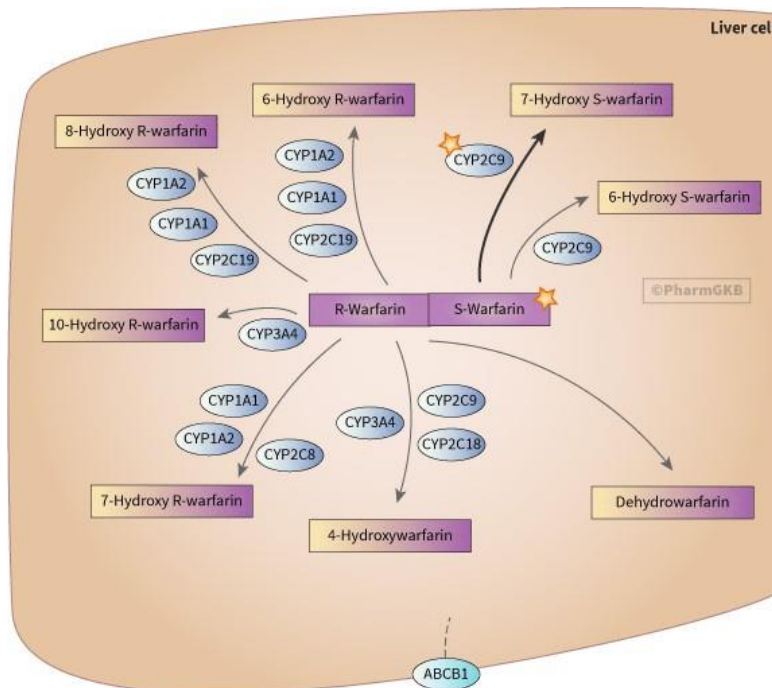


Figure 1.3.3.2. *CYP2C9* and several other known or suspected genes associated with warfarin pharmacokinetics. Image obtained from PharmGKB.³

Chapter 2. INTRODUCTION TO LONG NON-CODING RNA

The majority of the human genome is non-coding, and large portions of these non-coding regions contain genes that are transcribed into RNA (ribonucleic acid) molecules that do not encode for proteins.²⁶⁻²⁹ It is estimated that roughly 80% of these non-coding RNA genes are specifically lncRNA (long non-coding RNA) genes.^{26,28} LncRNAs make up a diverse class of non-coding RNA molecules that are longer than 200 nucleotides in length.²⁶⁻²⁸ They are also comprised of RNA polymerase II and III transcripts that may or may not be modified post-transcriptionally by splicing, capping, and polyadenylation.^{28,29}

LncRNAs have been found in many species but their functions are not well understood, which is in part due to poor sequence conservation.³⁰ However, lncRNAs have been implicated in a variety of cellular processes such as epigenetic, transcriptional, post-transcriptional, and translational regulation, post-translational modification, cellular differentiation, cell cycle regulation, protein relocation, and more.²⁶⁻²⁸ LncRNAs can even act as miRNA (micro RNA) sponges or precursors for miRNA and siRNA (small interfering RNA), which are other types of non-coding RNA molecules.^{26,27}

LncRNAs have a wide range of capabilities and their complex nature has become increasingly apparent as the field of lncRNA research rapidly expands. For example, studies have suggested that lncRNAs often do not function alone but rather interact with other genes and gene products.²⁶ This further increases the difficulty of characterizing the functions and mechanisms of lncRNAs as they cannot be classified as independently acting molecules.²⁶ This is another reason why previously identified or suspected lncRNAs have not been well annotated even though countless of them have been implicated in an assortment of disease processes.²⁸

In attempts to remedy this gap in knowledge, several projects have been started with the goal of identifying and annotating lncRNAs and other non-coding RNAs present in various species. This has subsequently led to the creation of numerous databases comprising of information on a range of non-coding RNAs. One such database is GENCODE.^{28,29,31,32} The GENCODE project is part of the larger ENCODE (Encyclopedia of DNA Elements) project, and it aims to “identify and classify all gene features in the human and mouse genomes with high accuracy based on biological evidence, and to release these annotations for the benefit of biomedical research and genome interpretation”.^{31,32} The GENCODE database includes information on genomic location, nucleotide sequence, and alternative splice isoforms of various known or suspected lncRNAs. GENCODE also includes some functional annotations, however these are minimal as this continues to be an active and challenging area of research.^{31,32} With regards to human lncRNA, GENCODE currently has 19,928 known or suspected human lncRNA genes.³¹

Chapter 3. AIM AND OBJECTIVES

Previous studies have uncovered numerous factors contributing to warfarin pharmacodynamics and pharmacokinetics. However, there is still a large portion of inter-individual variability seen in warfarin metabolism and response that is due to unknown factors.^{5,17,23} The aim of this thesis is to shed light on these elusive factors, which could be genetic or non-genetic in origin, and gain a deeper understanding of warfarin’s complex pharmacogenetic landscape. The primary objective is to conduct an initial exploratory analysis investigating the relationship between the expression of hepatic lncRNA genes and warfarin pharmacogenes, and identify lncRNA genes that may explain or predict the variability observed in warfarin pharmacogene expression.

Chapter 4. METHODS

4.1 RNA SEQUENCING DATA

An anonymized dataset consisting of FPKM (fragments per kilobase of transcript per million mapped reads) values for genes captured with RNA sequencing from previous studies was obtained. Prior studies conducted through the University of Washington utilized human liver tissue samples sourced from both the University of Washington Human Liver Bank and the St. Jude Liver Resource at St. Jude Children's Research Hospital.^{33,34} Studies like Tanner et al. (2017) and Wong et al. (2018) performed RNA sequencing on the human liver samples used in their study in order to quantify RNA expression levels. FPKM serves as a unit measure for gene expression, and the methods used to obtain and quantify the RNA sequencing data are thoroughly outlined in each of these studies.^{33,34}

4.2 WARFARIN PHARMACOGENE DATA

A list of known and suspected warfarin pharmacogenes was obtained from the PharmGKB database.^{3,35} This list was cross-referenced with the RNA sequencing FPKM output data, and a new data frame was created containing the FPKM values for only the warfarin pharmacogenes captured by RNA sequencing. The resulting warfarin pharmacogene data frame was utilized for the statistical analyses.

4.3 LONG NON-CODING RNA DATA

A list of known and suspected human lncRNA genes was obtained from the GENCODE database.^{31,32} LncRNAs identified in GENCODE by Ensembl gene IDs were converted to their corresponding gene symbols, provided they were designated, using the HGNC (HUGO Gene Nomenclature Committee) Database.^{36,37} The list of lncRNA genes was cross-referenced with the RNA sequencing FPKM output data to create a new data frame containing only the FPKM values

for lncRNA genes captured by RNA sequencing. This lncRNA data frame was utilized for the statistical analyses outlined in the following section.

4.4 STATISTICAL ANALYSES

All statistical analyses were performed in R (version 4.1.1), using the packages tidyverse (v.2.0.0), dplyr (v.1.1.2), gplots (v.3.1.3), ggplot2 (v.3.4.2), glmnet (v.4.1.7), and png (v.0.1.8).³⁸

FPKM values for both warfarin pharmacogenes and lncRNA genes were subject to further quality control. An FPKM threshold value of 0.1 was applied, and only the warfarin pharmacogenes and lncRNA genes in which 95% or more of the samples met the threshold were retained.^{30,39–41} This 0.1 FPKM threshold value was applied to ensure only robustly expressed genes were utilized in the statistical analyses.^{30,39–41} Summary statistics and standard deviations were then calculated in order to gain insight on the distribution and variability of the data at this stage. Additionally, Q-Q plots were constructed to assess whether the warfarin pharmacogenes and lncRNA genes follow a normal distribution, identify skewness and tails, and visualize outliers or extreme values.

Following quality control and an initial pass through the data, log₂-transformation was performed on both the warfarin pharmacogene and lncRNA gene FPKM values.^{39,42,43} To handle FPKM values of zero, a pseudocount of one was exclusively applied before the log₂-transformation to zero values. All other values remained unaltered for the log₂-transformation. Log₂-transformation was done to help stabilize variance, approximate a more normal distribution of data, and linearize multiplicative measurements. To assess the log₂-transformed data, new summary statistics and standard deviations were calculated and new Q-Q plots were created. Any outliers or extreme values were removed and new Q-Q plots were reconstructed as appropriate.

Next, the log₂-transformed warfarin pharmacogene and lncRNA gene FPKM values were standardized using the scale function.⁴⁴ This function centers the log₂-transformed FPKM values around their mean and scales them by their standard deviation so that they have a standard deviation of one. This standardization puts the warfarin pharmacogene FPKM values on the same scale as the lncRNA gene FPKM values, which further helps in approximating a normal distribution. New summary statistics and standard deviations were calculated and new Q-Q plots were constructed with the log₂-transformed and scaled FPKM values. After evaluating and removing any outliers or extreme values, additional Q-Q plots were reconstructed as appropriate.

Using the log₂-transformed and scaled FPKM values, Pearson correlation coefficients and their associated p-values were calculated. This was done to quantify and assess the strength and significance of the linear relationship between warfarin pharmacogenes and lncRNA genes.^{30,42,43,45} A Bonferroni correction was then performed to account for multiple testing and only correlation p-values less than the new significance threshold ($\alpha = 2.817e-05$) were identified as significant. New data frames were then created with only the log₂-transformed and scaled FPKM values for those warfarin pharmacogenes and lncRNA genes that had significant correlations.

Two Pearson correlation coefficient heatmaps with hierarchical clustering were generated using the heatmap.2 function.^{30,46} One heatmap was generated using all of the correlation coefficients obtained from the correlation test regardless of their significance level. A second heatmap was generated by assigning the insignificant correlations a value of zero, allowing visualization of only the significant correlations.

The significantly correlated warfarin pharmacogenes and lncRNA genes were then tested in various penalized linear regression models using the glmnet function.⁴⁷ The purpose of using penalized linear regressions was to address issues of overfitting and multicollinearity by

incorporating a penalty term to improve the model's generalization performance.⁴⁷ The subset of relevant predictors that are then retained in the model have a significant impact in explaining or predicting the response variable. One model, Lasso (Least absolute shrinkage and selection operator), adds a penalty term that is proportional to the absolute values of the coefficients to the sum of squared residuals.⁴⁷ Some coefficients in the Lasso model end up getting reduced down to a value of zero which results in them being excluded from the model. A second model, Ridge, adds a penalty term proportional to the squared values of the coefficients to the sum of squared residuals.⁴⁷ The Ridge model reduces and regularizes the coefficients so that the impact of large coefficients and multicollinearity is decreased. Coefficients in the Ridge model may get diminished but they do not necessarily get excluded. A third model, Elastic Net, is a combination of Lasso and Ridge and adds both penalty terms so that it is a linear combination of absolute and squared values of the coefficients.⁴⁷

Lasso, Ridge, and Elastic Net penalized linear regression models were employed in several scenarios with significantly correlated lncRNA genes as the predictor variables and significantly correlated warfarin pharmacogenes as the response variables. One scenario tested all significant lncRNA genes against each of the significant warfarin pharmacogenes independently. In contrast, three other scenarios tested all significant lncRNA genes against different groups of warfarin pharmacogenes. One of these scenarios tested all significant warfarin pharmacogenes together as one group. The other two scenarios used different groups of warfarin pharmacogenes determined by the branching pattern and arrangement of internal nodes in the dendrogram resulting from hierarchical clustering generated as part of the correlation heatmap.

For each penalized linear regression model, cross-validation was performed to evaluate model performance and generate cross-validation plots to be compared. Cross-validation plots are

graphical representations of model performance that plots MSE (mean-squared error) against a range of lambda values. MSE is the average squared difference between theoretical and actual values, and lower MSE values generally indicate better model performance.^{47,48} On the other hand, lambda is a regularization parameter and the lambda value associated with the best performing model is considered the “optimal lambda”.^{47,48} The optimal lambda value was used to identify and extract the subset of lncRNA genes that significantly explain the variability in warfarin pharmacogene expression.

Chapter 5. RESULTS

The anonymized RNA sequencing FPKM output data used for this thesis encompassed a total of 360 samples and 16,384 genes. The demographic breakdown of the liver bank donors in which the data was obtained can be seen in Table 5.1.

Table 5.1. Liver Bank Donor Demographics.

Characteristic	Parameters	Percentage (%)
Sex	Female	41.5
	Male	58.5
Age (years)	0-9	12.1
	10-19	11.1
	20-29	10.6
	30-39	9.09
	40-49	17.2
	50-59	18.8
	60-69	15.7
	70-79	4.29
Race	80+	1.26
	Asian	0.48
	Black	2.73
	Hispanic	0.80
	White	96.0

With regards to warfarin pharmacogenes, the PharmGKB database reports 58 known and suspected warfarin pharmacogenes. Of these 58, only 28 warfarin pharmacogenes were among the

16,384 genes captured in the RNA sequencing data. As for lncRNA genes, the GENCODE database reports 19,928 known and suspected human lncRNA genes. Among these, only 454 human lncRNA genes were captured in the RNA sequencing data. After applying quality control measures to ensure robust gene expression, there were only 25 remaining warfarin pharmacogenes and 71 remaining lncRNA genes that met the FPKM threshold.

Summary statistics including standard deviation for the 25 warfarin pharmacogenes that passed quality control are listed in Appendix A, Table A.1. The data ranges from a minimum FPKM value of 0.0233 for *F13A1* and a maximum FPKM value of 23,667 for *CALU*. *DDHDI* had the lowest mean and standard deviation at 1.0761 and 0.471, respectively. *CRP* had the highest mean and standard deviation at 8826 and 4780, respectively. Figure 5.1 displays a Q-Q plot of the original warfarin pharmacogene FPKM values before preprocessing techniques were employed.

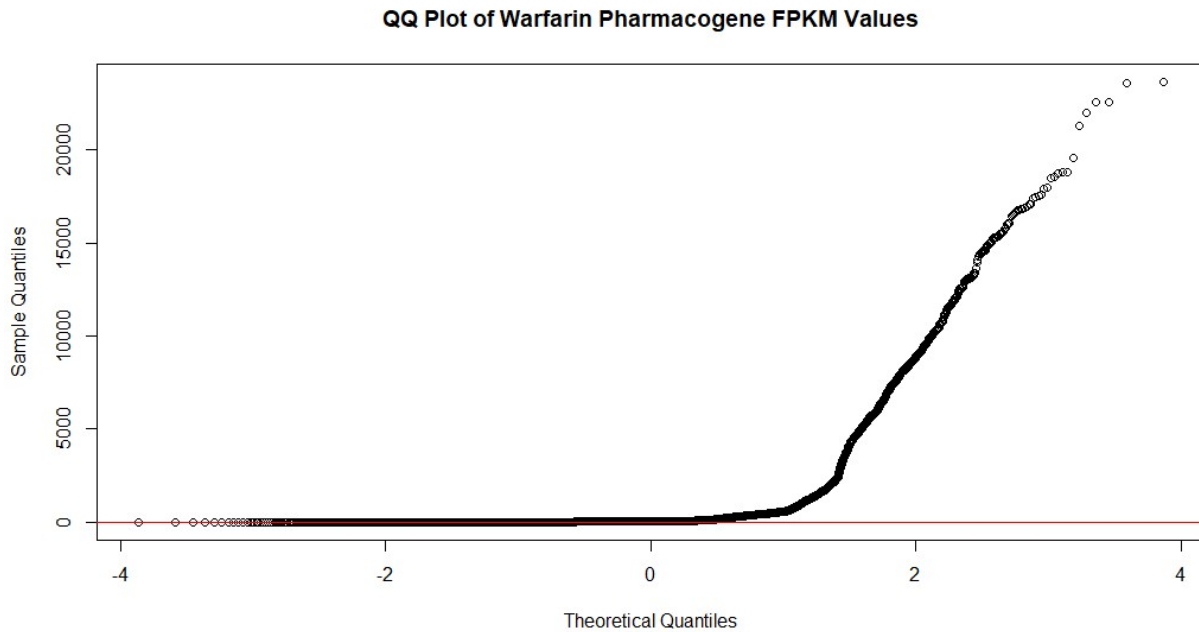


Figure 5.1. Q-Q plot of original warfarin pharmacogenes FPKM values.

Table A.2 in Appendix A lists the summary statistics including standard deviation for the 71 lncRNA genes that passed quality control. There were 12 lncRNA genes with a minimum FPKM value of zero. *CESIP1* had the highest maximum FPKM value, mean, and standard deviation at values of 1490, 335, and 169, respectively. *DPY19L2P4* had the lowest mean and standard deviation at 0.310 and 0.141, respectively. Figure 5.2 illustrates a Q-Q plot of the original lncRNA gene FPKM. values.

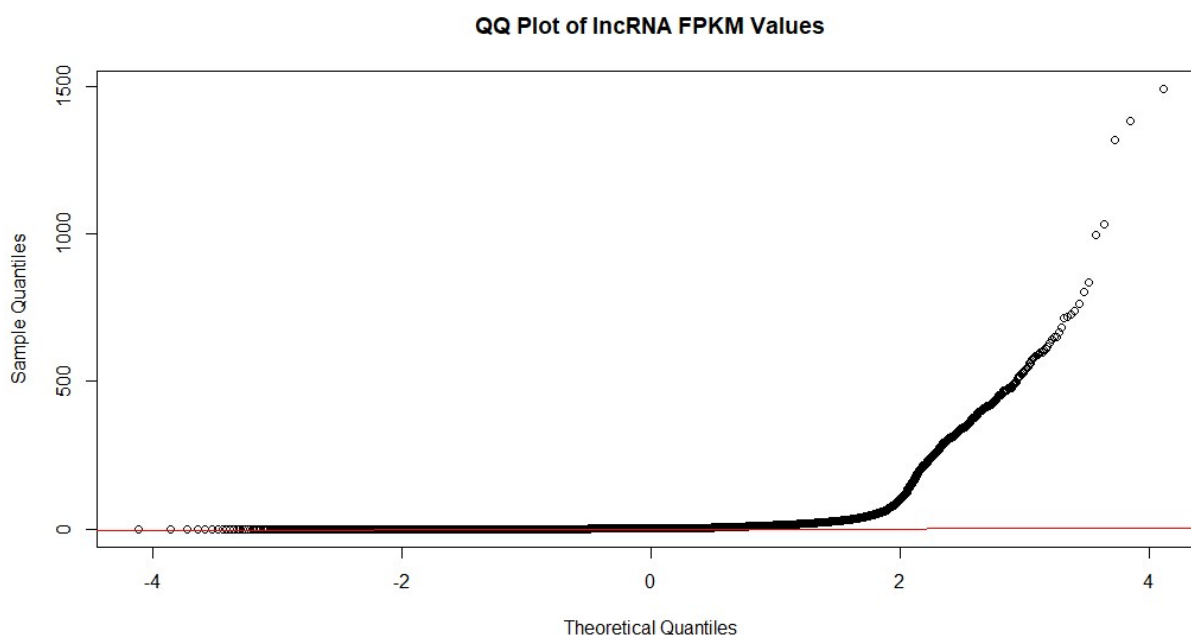


Figure 5.2. Q-Q plot of original lncRNA FPKM values.

The 25 warfarin pharmacogenes that passed quality control were log₂-transformed and the corresponding summary statistics including standard deviation are listed in Appendix A, Table A.3. The data ranges from a minimum log₂-transformed FPKM value of -5.291 for *F13A1* and a maximum log₂-transformed FPKM value of 14.531 for *CRP*. *F13A1* also had the lowest mean of -0.0418 while *CRP* had the highest mean of 12.796. For standard deviation, *VKORC1* had the

lowest standard deviation of 0.349 and *CYP3A4* had the highest standard deviation of 2.458. A Q-Q plot of the log₂-transformed warfarin pharmacogene FPKM values is displayed in Figure 5.3.

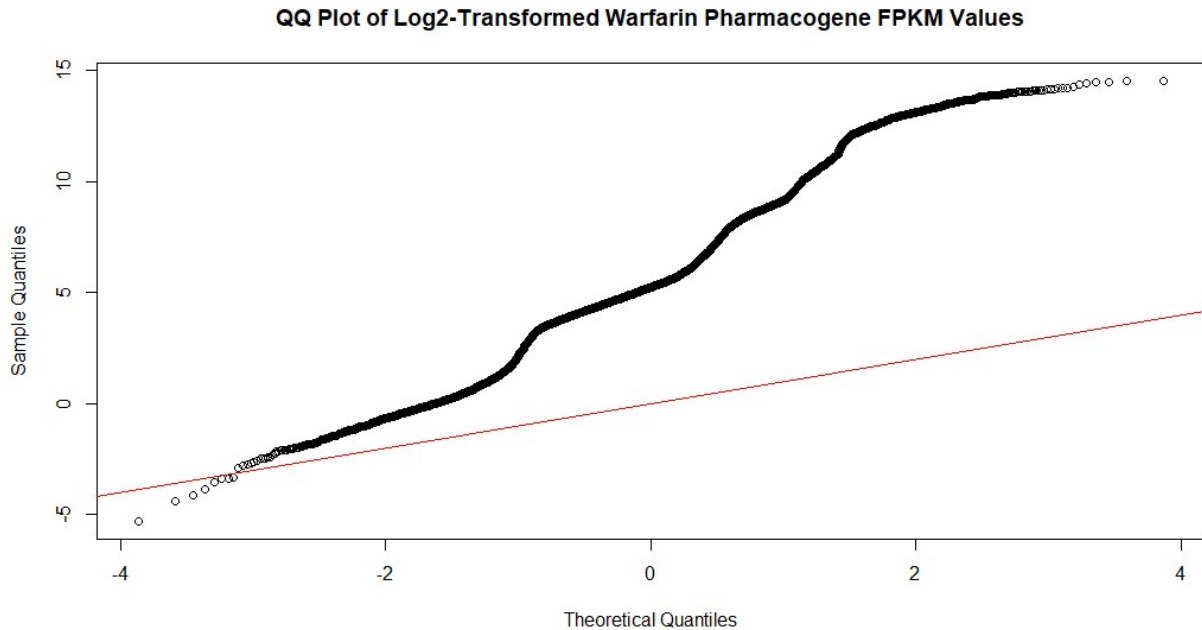


Figure 5.3. Q-Q plot of log₂-transformed warfarin pharmacogene FPKM values.

The summary statistics including standard deviations for the 71 log₂-transformed lncRNA genes are listed in Appendix A, Table A.4. It is important to note that these values were calculated with the outliers included. Regarding the range of the data, *CTC.338M12.4* had the lowest minimum log₂-transformed FPKM value of -211 and *CESIP1* had the highest maximum log₂-transformed FPKM value of 10.541. *DPY19L2P4* had the lowest mean of -1.845 while *CESIP1* had the highest mean of 8.250. With regards to standard deviation, *FGD5.AS1* had the lowest standard deviation of 0.380 and *CTC.338M12.4* had the highest standard deviation of 11.125. Figure 5.4 shows the Q-Q plot of the log₂-transformed lncRNA gene FPKM values with outliers included and Figure 5.5 displays the updated Q-Q plot after outliers were removed.

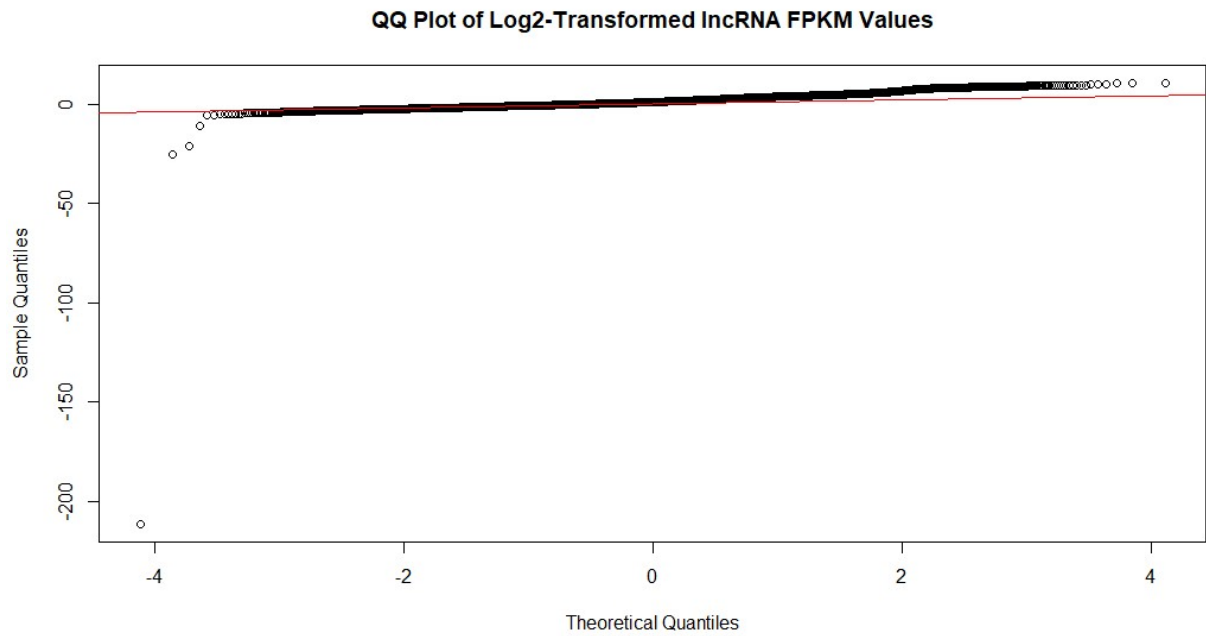


Figure 5.4. Q-Q plot of log₂-transformed lncRNA FPKM values with outliers included.

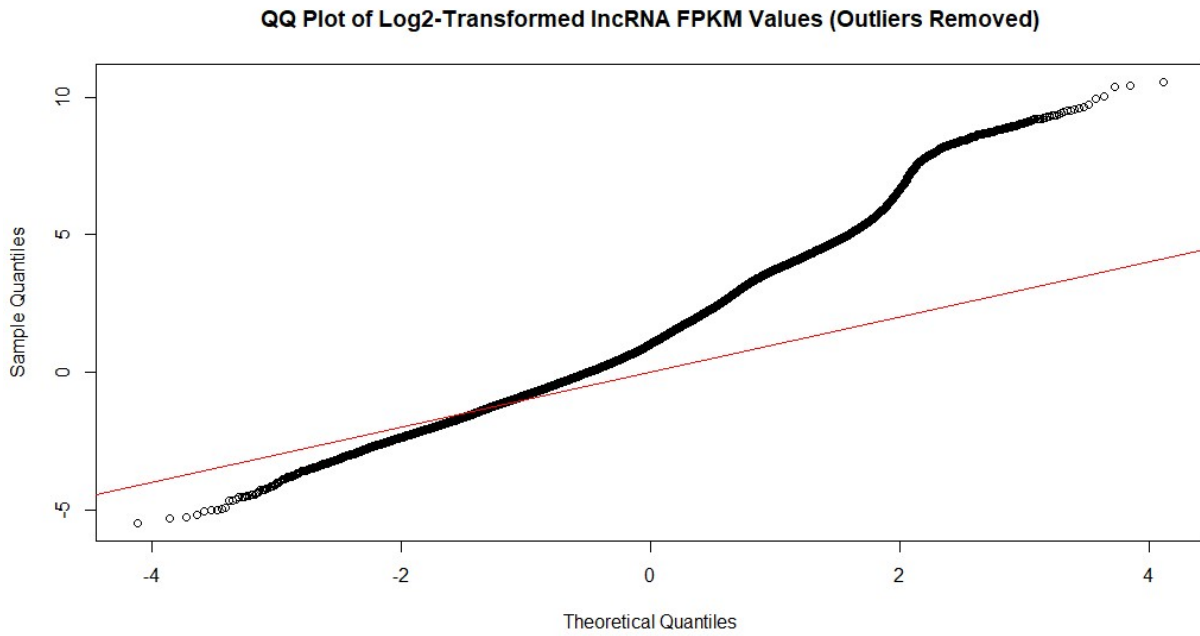


Figure 5.5. Q-Q plot of log₂-transformed lncRNA FPKM values with outliers removed.

Appendix A, Table A.5 contains summary statistics including standard deviation for the 25 log₂-transformed and scaled warfarin pharmacogenes. Due to scaling, all warfarin pharmacogenes had a standard deviation of one. Regarding the mean, *CYP4F11* has the lowest mean of -8.312e-16 and *APOE* has the highest mean of 1.520e-15. The log₂-transformed and scaled FPKM values range from a minimum of -5.832 for *CYP4F2* to a maximum of 3.928 for *VKORC1*. Figure 5.6 exhibits a Q-Q plot of the log₂-transformed and scaled warfarin pharmacogene FPKM values.

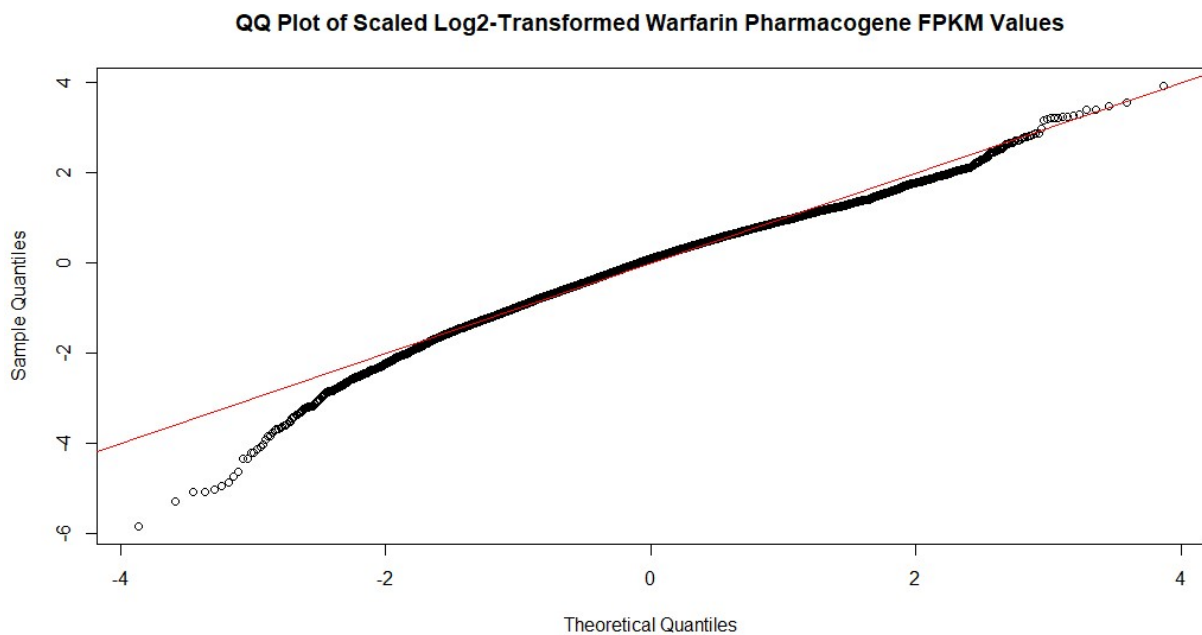


Figure 5.6. Q-Q plot of log₂-transformed and scaled warfarin pharmacogene FPKM values.

Summary statistics including standard deviation for the 71 log₂-transformed and scaled lncRNA genes are listed in Appendix A, Table A.6. These values were calculated with the outliers included. The log₂-transformed and scaled FPKM values range from a minimum of -18.869 for *CTC.338MI2.4* to a maximum of 4.690 for *DNAJC9.AS1*. For the mean, *DANCR* had the lowest mean of -3.215e-16 and *CROCCP2* had the highest mean of 5.429e-16. All lncRNA genes had a standard deviation of one due to scaling. Q-Q plots of the log₂-transformed and scaled lncRNA

gene FPKM values with and without outliers are exhibited in Figure 5.7 and Figure 5.8, respectively.

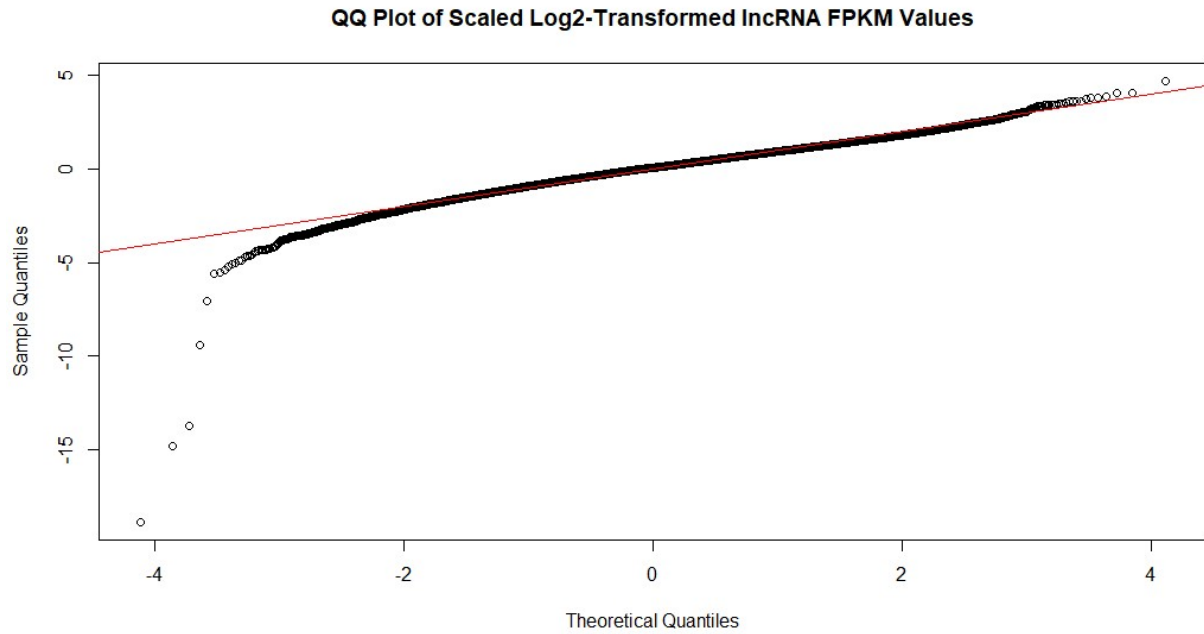


Figure 5.7. Q-Q plot of log2-transformed and scaled lncRNA FPKM values with outliers included.

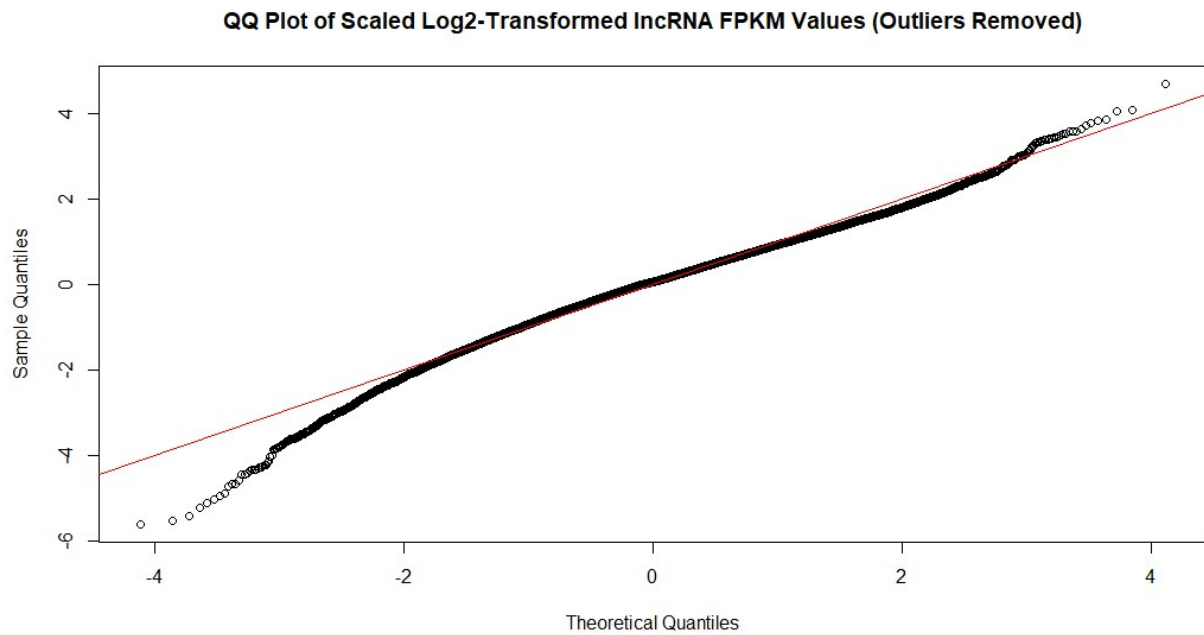


Figure 5.8. Q-Q plot of log2-transformed and scaled lncRNA FPKM values with outliers removed.

Figure 5.9 illustrates a Pearson correlation coefficient heatmap obtained using the log2-transformed and scaled lncRNA gene and warfarin pharmacogene FPKM values. All correlation coefficients are included regardless of significance. While clusters of strongly correlated variables can be visualized in the upper right and upper left corners of the heatmap, any further clustering is difficult to ascertain due to overlap with regions containing a mix of positive and negative correlations.

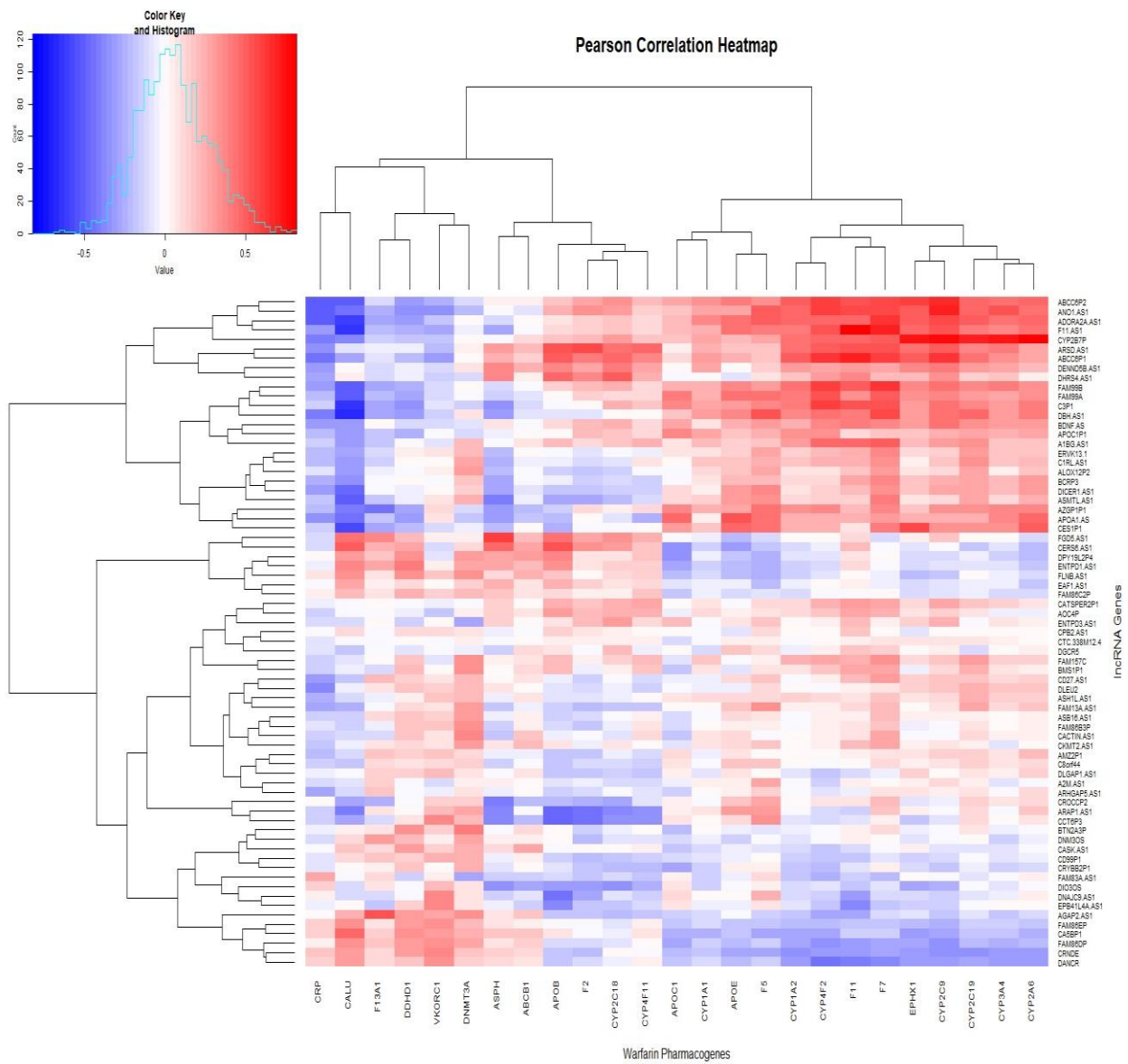


Figure 5.9. Heatmap of all Pearson correlation coefficients.

Figure 5.10 is a Pearson correlation coefficient heatmap displaying only significant correlations after Bonferroni correction (Bonferroni corrected $\alpha = 2.817e-05$). Of the 71 lncRNA genes included in the analysis, 65 of them had significant correlations. In contrast, all 25 warfarin pharmacogenes were significantly correlated with multiple lncRNA genes. The number of significant correlations ranged from six lncRNA genes correlated with the warfarin pharmacogene *ABCB1* to 38 lncRNA genes correlated with the warfarin pharmacogene *CALU*. Additionally, more than half of the significant correlations were positive.

As visualized on the correlation heatmap, a block of 23 lncRNA genes in the upper right corner are positively correlated with nine warfarin pharmacogenes: *F11*, *F7*, *CYP2A6*, *CYP2C19*, *CYP4F2*, *CYP2C9*, *EPHX1*, *CYP3A4*, and *CYP1A2*. These nine warfarin pharmacogenes are also negatively correlated with five other lncRNA genes in the lower right of the heatmap. There is another cluster in the lower right of the heatmap consisting of eight lncRNA genes positively correlated with five warfarin pharmacogenes: *F5*, *F11*, *F7*, *CYP2A6*, and *CYP2C19*. Additionally, there is a smaller cluster of 14 lncRNA genes that are positively correlated with three warfarin pharmacogenes (*APOC1*, *APOE*, and *F5*) in the upper middle along with an adjacent cluster of 14 lncRNA genes positively correlated with five other warfarin pharmacogenes (*ASPH*, *APOB*, *CYP4F11*, *F2*, and *CYP2C18*). These five warfarin pharmacogenes are also negatively correlated with 10 other lncRNA genes in the lower middle of the heatmap. Lastly, in the upper left corner of the heatmap there is a cluster of 15 lncRNA genes that are negatively correlated with four warfarin pharmacogenes: *CRP*, *CALU*, *F13A1*, and *DDHDI*.

In addition to the clusters of positive and negative correlations in Figure 5.10 that differ from those in Figure 5.9, the hierarchical clustering and resulting dendrogram also differ

slightly between the two heatmaps. The results of the hierarchical clustering were used to inform the warfarin pharmacogenes groups created for the penalized linear regressions.

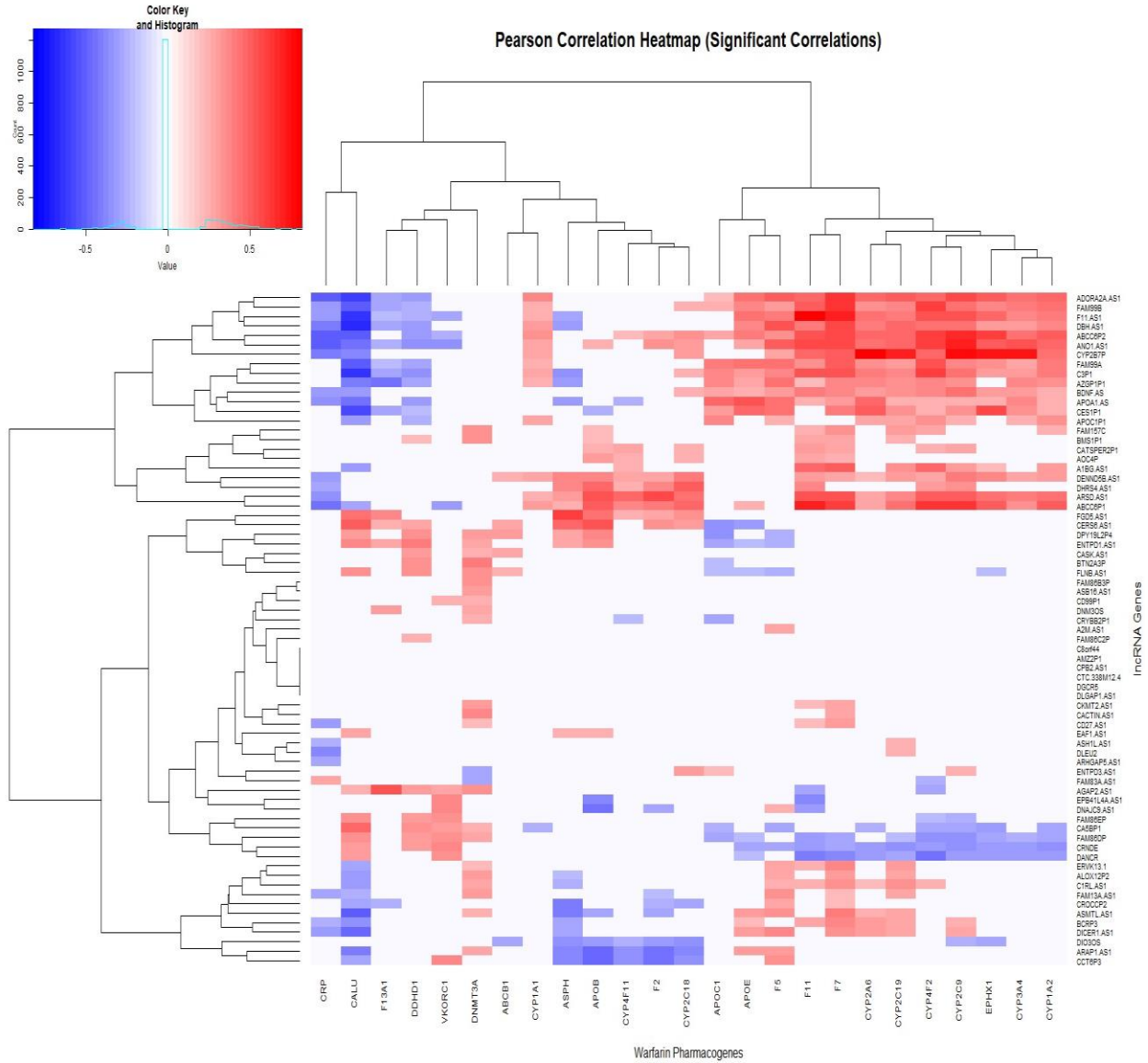


Figure 5.10. Heatmap of significant Pearson correlation coefficients (Bonferroni corrected $\alpha = 2.817e-05$).

Appendix B and C contain both penalized linear regression cross-validation plots as well as the lists of lncRNA genes contributing the most to the variability in warfarin pharmacogene expression under the Lasso, Elastic Net, and Ridge models. Data for the models using individual

warfarin pharmacogenes are in Appendix B, and data for the models using various groups of warfarin pharmacogenes are in Appendix C.

For the testing of individual warfarin pharmacogenes, the Lasso and Elastic Net models performed better than the Ridge model across all warfarin pharmacogenes. When comparing the Lasso and Elastic Net models to each other, they performed similarly although the Lasso model slightly outperformed the Elastic Net model with regards to model complexity and model accuracy. The number of lncRNA genes found to significantly contribute to the variability in warfarin pharmacogene expression ranged from 29 to 55 with the Lasso model and 29 to 54 with the Elastic Net model. In contrast, the Ridge model identified all 65 lncRNA genes to significantly contribute regardless of the warfarin pharmacogene tested in the model.

With regards to testing groups of warfarin pharmacogenes, the Lasso, Elastic Net, and Ridge models found all 65 lncRNA genes to be significant contributors when testing all 25 warfarin pharmacogenes as a single group. However, when comparing the cross-validation plots the Lasso model performed better than the Elastic Net and Ridge models in this setting. In addition to testing all 25 warfarin pharmacogenes together as a single group, there were two other scenarios considered in which the warfarin pharmacogenes were divided into different groups based on hierarchical clustering in the correlation heatmap. One scenario divided the warfarin pharmacogenes into three groups and the other divided them into six groups. These groupings can be seen in Table 5.2 on the following page.

When testing the Lasso, Elastic Net, and Ridge models with the three groups of warfarin pharmacogenes, the Lasso and Elastic Net models slightly outperformed the Ridge model across all three groups. Additionally, there was not much variability across the three models or the three groups with regards to the number of lncRNA genes identified as significant contributors to the

variability in warfarin pharmacogene expression. For the first group consisting of 12 warfarin pharmacogenes, the number of lncRNA genes found to be significant contributors by the Lasso, Elastic Net, and Ridge models were 64, 65, and 65, respectively. The Lasso, Elastic Net, and Ridge models for the second group consisting of 11 warfarin pharmacogenes identified 63, 65, and 65 significantly contributing lncRNA genes, respectively. Finally, the third group consisting of two warfarin pharmacogenes, the Lasso, Elastic Net, and Ridge models found 53, 56, and 65 lncRNA genes to be significant contributors.

Table 5.2 The groups of warfarin pharmacogenes tested in the Lasso, Elastic Net, and Ridge penalized linear regression models as informed by hierarchical clustering in the correlation heatmap.

Scenario	Group	Warfarin pharmacogenes
1	1	ABCB1, APOB, APOC1, APOE, ASPH, CALU, CRP, CYP1A1, CYP1A2, CYP2A6, CYP2C18, CYP2C19, CYP2C9, CYP3A4, CYP4F11, CYP4F2, DDHD1, DNMT3A, EPHX1, VKORC1, F11, F13A1, F2, F5, F7
	1	APOC1, APOE, CYP1A2, CYP2A6, CYP2C19, CYP2C9, CYP3A4, CYP4F2, EPHX1, F5, F7, F11
2	2	ABCB1, APOB, ASPH, CYP1A1, CYP2C18, CYP4F11, DDHD1, DNMT3A, F13A1, F2, VKORC1
	3	CALU, CRP
3	1	CYP1A2, CYP2A6, CYP2C9, CYP2C19, CYP3A4, CYP4F2, EPHX1, F7, F11
	2	APOC1, APOE, F5
	3	APOB, ASPH, CYP2C18, CYP4F11, F2
	4	ABCB1, CYP1A1
	5	DDHD1, DNMT3A, F13A1, VKORC1
	6	CALU, CRP

For the scenario testing six groups of warfarin pharmacogenes, the Lasso model generally outperformed the Elastic Net and Ridge models across all six groups. However, there were some groups in which the Lasso and Elastic Net models performed similarly and were only slightly better than the Ridge model, as evident by the cross-validation plots. In addition, the Ridge model identified all 65 lncRNA genes as significantly influencing the variability of warfarin

pharmacogene expression across all six groups. On the other hand, the number of lncRNA genes identified as significant contributors ranged from 39 to 64 with the Lasso model and 40 to 64 with the Elastic Net model.

Regardless of whether individual or groups of warfarin pharmacogenes were tested, the Lasso penalized linear regression model generally outperformed the Elastic Net and Ridge penalized linear regression models with regards to model complexity and accuracy. When comparing the results of the Lasso model across individual and groups of warfarin pharmacogenes, the models testing individual warfarin pharmacogenes performed better than those testing groups of warfarin pharmacogenes. Furthermore, the scenario in which six groups of warfarin pharmacogenes were tested outperformed the scenario in which three groups of warfarin pharmacogenes were tested using the Lasso penalized linear regression model.

Chapter 6. DISCUSSION

The purpose of this thesis was to perform exploratory analysis, investigating the relationship between lncRNA genes and warfarin pharmacogenes as well as identifying lncRNA genes that help explain or predict the variability observed in warfarin pharmacogene expression. This was to gain insight on the other potential genetic factors contributing to variability in warfarin metabolism and response. Calculating the Pearson correlation coefficient assessed the relationship between lncRNA genes and warfarin pharmacogenes. Subsequent testing of various penalized linear regression models identified subsets of lncRNA genes that may explain or predict variability in warfarin pharmacogenes.

Data received in the form of FPKM values were first subjected to quality control measures and preprocessing techniques. This was necessary to ensure the use of lncRNA genes and warfarin

pharmacogenes that were robustly expressed, and to ensure the data met the assumptions for certain statistical analyses. Pearson correlation coefficients were then calculated with the transformed data. There were 65 lncRNA genes significantly correlated with at least one warfarin pharmacogene and all 25 warfarin pharmacogenes were correlated with multiple lncRNA genes. Moreover, greater than half of the significant correlations were positive. These positive correlations indicate that higher expression of one gene tends to be associated with higher expression of the other gene, or lower expression of one gene tends to be associated with lower expression of the other gene. On the other hand, those that were negatively correlated indicates that higher expression of one gene tends to be associated with lower expression of the other gene, and vice versa.

Regarding clustering in the correlation heatmap, genes exhibiting similar expression profiles were grouped together. However, these expression profiles do not necessarily correspond to similarities in gene function, location, or involvement in certain pathways. For example, *CYP2C18* and *F2* were clustered together but *CYP2C18* is involved in the minor pathway for (S)- and (R)-warfarin metabolism whereas *F2* is involved in the blood clotting cascade pathway and plays a crucial role in activating a number of coagulation factors and proteins.⁴⁹ Such differences among some of the warfarin pharmacogenes are only known thanks to prior knowledge. Unfortunately, many warfarin pharmacogenes and virtually all lncRNA genes lack comprehensive annotations regarding their function or the biological pathways they are involved in. Consequently, it is challenging to ascertain whether clustering based on gene expression profile also signifies similarities in these areas between warfarin pharmacogenes and lncRNA genes.

The existing body of research concerning associations between lncRNA genes and warfarin pharmacogenes is notably limited. The significant correlations revealed in this analysis suggest

associations that have not been identified in the past. While these correlation coefficients quantify the linear relationship between lncRNA genes and warfarin pharmacogenes, they do not provide information about their temporal or spatial relationship. The significant results from this study are anticipated to offer a foundation for further exploration into these relationships. It may also encourage characterization of lncRNA genes and warfarin pharmacogenes which will aid in better understanding the associations and interactions between the two.

Now regarding the penalized linear regressions, three different models (Lasso, Elastic Net, and Ridge) were used to test all 65 significantly correlated lncRNA genes against individual or groups of warfarin pharmacogenes. There were a few instances in which the Elastic Net penalty resulted in the best fitting model; however, the Lasso penalized linear regression model generally outperformed the Elastic Net and Ridge models across all warfarin pharmacogene groupings. It was also observed that model performance improved as the groups of warfarin pharmacogenes tested in the model became smaller in size. Furthermore, the Lasso penalized linear regression models tested with individual warfarin pharmacogenes retained less lncRNA genes in the cross-validation compared to those tested with various groups of warfarin pharmacogenes.

First, it is important to reiterate that the groups of warfarin pharmacogenes were informed by hierarchical clustering which arranged genes together based on similar gene expression profiles which may not correspond with similarities in their gene function or the biological pathways they are involved in. Nevertheless, given that several warfarin pharmacogenes share analogous functions or participate in common pathways, it was anticipated that testing the warfarin pharmacogenes in three or six groups would yield more accurate and generalizable models than testing all warfarin pharmacogenes individually or together as one group. By evaluating larger groups of warfarin pharmacogenes with broader similarities such as the scenario of dividing them

into three groups, the anticipation was to capture overarching gene expression patterns and gain insight to the common regulatory mechanisms influenced by groups of lncRNA genes. This type of model would offer better interpretability when visualizing larger scale relationships. In contrast, testing smaller groups of warfarin pharmacogenes with more specific functions such as the scenario that divided them into six groups, was expected to reveal intricate interactions and specific regulatory mechanisms that would otherwise be missed when analyzing larger groups. This type of model would offer greater sensitivity for detecting effects.

Results from this analysis do capture smaller scale relationships between lncRNA genes and individual warfarin pharmacogenes. However, it does not adequately capture larger scale relationships or insights that could be obtained by testing warfarin pharmacogenes in various groups. This suggests that as the number of warfarin pharmacogenes increases, the stability and generalizability of the model decreases. One potential contributing factor could be the method used to group warfarin pharmacogenes by their expression profiles. This method might not account for any underlying heterogeneity, including variations in gene function and involvement in biological pathways. Individual gene effects could also be masked by this grouping method. Other potential factors affecting model performance could also include model overfitting, increased model complexity, high dimensionality, the presence of many strong correlations, and nonlinearity.

While testing individual warfarin pharmacogenes in the model may provide valuable insight, it may not be the most ideal approach. By using individual warfarin pharmacogenes, gene interactions could be overlooked along with some biological functions that may only be evident when considering groups of genes that work together. Additionally, the computational intensity and increased risk of false positives due to multiple testing could become issues.

Furthermore, while the number of lncRNA genes retained in a model does not inherently dictate its performance, it can serve as an indicator of certain model characteristics that could result in a model that exhibits suboptimal performance. These characteristics include model overfitting, increased model complexity, the presence of numerous strong correlations, high variability in gene expression, inclusion of genes that are involved in complex biological processes, or the presence of group effects which occur when genes share similar regulatory mechanisms or functional relationships. It is important to recognize that while a model may retain more lncRNA genes, this does not automatically suggest poor performance. Conversely, it could be an indication that those lncRNA genes have a consistent relationship with warfarin pharmacogenes across different samples. If this is the case it could lead to enhanced generalizability and predictive capacity.

As highlighted earlier, the existing literature examining the associations between lncRNA genes and warfarin pharmacogenes is minimal, but the outcome of these penalized linear regression models suggests previously undescribed associations between the two. However, the exact nature of these relationships still requires comprehensive understanding. Moreover, the specific lncRNA genes responsible for explaining the variability in warfarin pharmacogene expression were not definitively identified. The findings from this study encourage additional investigation of the association between lncRNA genes and warfarin pharmacogenes, along with investigating the potential of one's presence to predict the presence of the other.

6.1 LIMITATIONS

This study is subject to several limitations that could have influenced the results and subsequent interpretations. One limitation is the sample of 360 donors which was obtained from a previous study. With this sample size, there may not have been sufficient statistical power to detect

significant gene expression levels, correlations, or associations. Another constraint is the low sample diversity, as over 95% of the donors were of White European ancestry. This limited diversity might have led to an incomplete representation of the population's variability.

The data obtained for the study was also in the form of FPKM values, and the quality control and preprocessing techniques that were employed could have influenced the results. The FPKM values were subjected to a threshold of 0.1 and a requirement that 95% of the samples should meet this threshold. These measures aimed to ensure the use of robustly expressed genes in the analyses. However, there is a consideration regarding the threshold which was chosen based off previous studies that worked with lncRNA.^{30,39-41} Depending on cell type, developmental state, and environmental conditions, lncRNA may be expressed at lower levels which would support an argument that a lower threshold should be utilized to ensure lncRNA genes with low expression are captured.⁵⁰ On the other hand, there is the potential for random variability in gene expression measurements that are not due to true biological signals that could be excluded by using a higher threshold. All this to say, there is a trade-off between capturing biologically relevant data and avoiding background noise.

Log₂-transformed and scaled FPKM values were utilized for the Pearson correlation and penalized linear regression analyses, following the rationale of previous studies that utilized these techniques to help normalize lncRNA data.^{39,42,43} Although the Q-Q plot generated for the log₂-transformed and scaled data in this study aligned more closely with a normal distribution compared to the original and log₂-transformed data, it still exhibited a heavy tail. This deviation from the normal distribution could have disproportionately influenced the results of the Pearson correlations and subsequent penalized linear regressions, providing results that may not accurately represent the underlying relationship between lncRNA genes and warfarin pharmacogenes.

Another limitation to the study is the assumption that the relationship between lncRNA genes and warfarin pharmacogenes is linear. This assumption was made to simplify the analysis and allow utilization of parametric methods such as Pearson correlations and penalized linear regressions. Nevertheless, this could have led to associations that inadequately capture the dynamics underlying the interactions between lncRNA genes and warfarin pharmacogenes.

Lastly, the use of hierarchical clustering to group warfarin pharmacogenes by gene expression profiles might have introduced limitations to the subsequent penalized linear regression analysis. As previously mentioned, groups of genes that share similar gene expression profiles may not share similarities in their gene function or the biological pathways they are involved in. This grouping strategy possibly contributed to the vague and inconclusive outcomes of the penalized linear regressions that utilized these gene groups.

6.2 FUTURE DIRECTIONS

Prior studies investigating the factors affecting warfarin pharmacodynamics and pharmacokinetics have predominantly included participants of White European ancestry. However, historically excluded groups face an elevated risk of experiencing adverse events during warfarin therapy due to a number of reasons.^{14,15} This underscores the importance of conducting additional studies that include these underrepresented individuals. Increasing sample diversity not only addresses this gap but it also ensures a more comprehensive representation of population variability. Augmenting sample sizes for these studies would also enhance statistical power, enabling the detection of significant relationships.

Another consideration for future studies involves evaluating the trade-offs associated with different FPKM threshold values. This is because the balance created between capturing genes

with low expression and minimizing background noise could have profound changes in the results and interpretation. Additionally, alternative preprocessing methods could be explored for more robust and accurate results. These approaches might encompass different normalization techniques or assess the impact of varying threshold values.

Considering the analysis using log₂-transformed and scaled FPKM values, it may be appropriate to explore other transformation or scaling techniques that may better align the data with the assumptions of the statistical methods being used. Investigating these other techniques could mitigate the impact of using data that does not already meet the necessary assumptions. This would also enhance the accurate modeling of relationships between lncRNA genes and warfarin pharmacogenes.

Gene regulation is a complex process and multifaceted interactions between lncRNA genes and warfarin pharmacogenes may vary depending on numerous factors. These could be anything from the specific drug in question, cell types involved, disease state, individual genetic background, and more. Acknowledging this complexity, it may not be appropriate to assume a linear relationship between lncRNA genes and warfarin pharmacogenes. It may be essential to consider methods that allow more flexibility in the assumptions made about underlying relationships. Exploring non-linear modeling techniques may also be considered to capture this complex dynamic.

To further enhance our understanding of associations between lncRNA genes and warfarin pharmacogenes, studies should also be conducted to uncover the biological mechanisms that could explain any associations observed. This could involve methods such as gene knockout studies, functional assays, expression profiling, pathway analysis, functional enrichment analysis, network

analysis, pathway modeling, and more. These approaches might help uncover patterns in the data that may not be revealed by traditional methods.

Validation and replication studies should also be considered as they are important components of rigorous scientific research. Validation studies assess the accuracy of the analyses used, while replication studies assess the generalizability of results. These principles apply not only to the analysis and results obtained from this study, but also to any future research endeavors. Additionally, independent data obtained from different populations should be employed in these studies.

Ultimately, exploring the functions and biological pathways of lncRNA genes and warfarin pharmacogenes, along with understanding their interactions, would provide a deeper comprehension of the pharmacokinetics and pharmacodynamics of warfarin. This knowledge could then be used to investigate specific associations between certain lncRNA genes and individual patient responses to warfarin dosage to shed light on personalized treatment strategies. For instance, the *CYP2C9**3 variant allele has been previously associated with warfarin sensitivity, requiring patients who carry this variant to be prescribed lower starting and maintenance doses of warfarin.^{22,51–53} In this study, *CYP2C9* was found to be positively correlated with lncRNA gene *ABCC6P1*. If *ABCC6P1* is linked specifically with variant allele *CYP2C9**3, it could serve as a bases for determining the lower warfarin dosage requirements in patients expressing this specific lncRNA gene. By identifying these associations between lncRNA genes and warfarin pharmacogenes, clinicians could tailor warfarin dosages to individual patients, ensuring more effective anticoagulation while simultaneously mitigating the risk of adverse side effects. This targeted approach has the potential to revolutionize patient care in the realm of anticoagulation therapy.

REFERENCES

1. National Center for Biotechnology Information. (S)-Warfarin. PubChem Compound Summary for CID 54688261, (S)-Warfarin. Published 2023. Accessed June 5, 2023. <https://pubchem.ncbi.nlm.nih.gov/compound/54688261>
2. National Center for Biotechnology Information. (R)-Warfarin. PubChem Compound Summary for CID 54684598, (R)-Warfarin. Published 2023. Accessed June 5, 2023. https://pubchem.ncbi.nlm.nih.gov/compound/R_-Warfarin
3. Whirl-Carrillo M, Huddart, R., Gong L, et al. An evidence-based framework for evaluating pharmacogenomics knowledge for personalized medicine. *Clinical Pharmacology & Therapeutics*. 2021;0(0):1-10. doi:10.1002/cpt.2350
4. Katara P, Yadav A. Pharmacogenes (PGx-genes): Current understanding and future directions. *Gene*. 2019;718:144050. doi:10.1016/j.gene.2019.144050
5. Roden DM, McLeod HL, Relling MV, et al. Pharmacogenomics. *Lancet*. 2019;394(10197):521-532. doi:10.1016/S0140-6736(19)31276-0
6. Dean L. Warfarin Therapy and VKORC1 and CYP Genotype. In: Pratt VM, Scott SA, Pirmohamed M, Esquivel B, Kattman BL, Malheiro AJ, eds. *Medical Genetics Summaries*. National Center for Biotechnology Information (US); 2012. Accessed May 27, 2023. <http://www.ncbi.nlm.nih.gov/books/NBK84174/>
7. Jonas DE, McLeod HL. Genetic and clinical factors relating to warfarin dosing. *Trends in Pharmacological Sciences*. 2009;30(7):375-386. doi:10.1016/j.tips.2009.05.001
8. Gage BF, Lesko LJ. Pharmacogenetics of warfarin: regulatory, scientific, and clinical issues. *J Thromb Thrombolysis*. 2008;25(1):45-51. doi:10.1007/s11239-007-0104-y

9. Patel S, Singh R, Preuss CV, Patel N. Warfarin. In: *StatPearls*. StatPearls Publishing; 2023. Accessed July 10, 2023. <http://www.ncbi.nlm.nih.gov/books/NBK470313/>
10. Rettie AE, Tai G. The Pharmacogenomics of Warfarin. *Molecular Interventions*. 2006;6(4):223. doi:10.1124/mi.6.4.8
11. Kawai M, Harada M, Motoike Y, et al. Impact of serum albumin levels on supratherapeutic PT-INR control and bleeding risk in atrial fibrillation patients on warfarin: A prospective cohort study. *Int J Cardiol Heart Vasc*. 2019;22:111-116. doi:10.1016/j.ijcha.2019.01.002
12. Lal S, Jada SR, Xiang X, Lim WT, Lee EJD, Chowbay B. Pharmacogenetics of target genes across the warfarin pharmacological pathway. *Clinical Pharmacokinetics*. 2006;45(12):1189-1201.
13. Pouncey DL, Barnette DA, Sinnott RW, et al. Discovery of Novel Reductive Elimination Pathway for 10-Hydroxywarfarin. *Frontiers in Pharmacology*. 2022;12. Accessed August 24, 2023. <https://www.frontiersin.org/articles/10.3389/fphar.2021.805133>
14. Johnson JA, Cavallari LH. Warfarin pharmacogenetics. *Trends in Cardiovascular Medicine*. 2015;25(1):33-41. doi:10.1016/j.tcm.2014.09.001
15. Kaye JB, Schultz LE, Steiner HE, Kittles RA, Cavallari LH, Karnes JH. WARFARIN PHARMACOGENOMICS IN DIVERSE POPULATIONS. *Pharmacotherapy*. 2017;37(9):1150-1163. doi:10.1002/phar.1982
16. Lee MTM, Klein TE. Pharmacogenetics of warfarin: challenges and opportunities. *J Hum Genet*. 2013;58(6):334-338. doi:10.1038/jhg.2013.40
17. Verhoef TI, Redekop WK, Daly AK, van Schie RMF, de Boer A, Maitland-van der Zee AH. Pharmacogenetic-guided dosing of coumarin anticoagulants: algorithms for warfarin,

- acenocoumarol and phenprocoumon. *Br J Clin Pharmacol*. 2014;77(4):626-641.
doi:10.1111/bcp.12220
18. Thummel KE, Lin YS. Sources of Interindividual Variability. In: Nagar S, Argikar UA, Tweedie DJ, eds. *Enzyme Kinetics in Drug Metabolism: Fundamentals and Applications*. Methods in Molecular Biology. Humana Press; 2014:363-415. doi:10.1007/978-1-62703-758-7_17
19. Johnson J, Caudle K, Gong L, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. *Clinical Pharmacology & Therapeutics*. 2017;102(3):397-404. doi:10.1002/cpt.668
20. Pavani A, Naushad SM, Rupasree Y, et al. Optimization of warfarin dose by population-specific pharmacogenomic algorithm. *The Pharmacogenomics Journal*. 2012;12(4):306-311. doi:10.1038/tpj.2011.4
21. Flora DR, Rettie AE, Brundage RC, Tracy TS. CYP2C9 Genotype-Dependent Warfarin Pharmacokinetics: Impact of CYP2C9 Genotype on R- and S-Warfarin and Their Oxidative Metabolites. *The Journal of Clinical Pharmacology*. 2017;57(3):382-393. doi:10.1002/jcph.813
22. Fekete F, Mangó K, Déri M, Incze E, Minus A, Monostory K. Impact of genetic and non-genetic factors on hepatic CYP2C9 expression and activity in Hungarian subjects. *Sci Rep*. 2021;11(1):17081. doi:10.1038/s41598-021-96590-3
23. Tavares LC, Marcatto LR, Santos PC. Genotype-guided warfarin therapy: current status. *Pharmacogenomics*. 2018;19(7):667-685. doi:10.2217/pgs-2017-0207

24. Bader LA, Elewa H. The Impact of Genetic and Non-Genetic Factors on Warfarin Dose Prediction in MENA Region: A Systematic Review. *PLoS One*. 2016;11(12):e0168732. doi:10.1371/journal.pone.0168732
25. Fohner AE, Robinson R, Yracheta J, et al. Variation in Genes Controlling Warfarin Disposition and Response in American Indian and Alaska Native People: CYP2C9, VKORC1, CYP4F2, CYP4F11, GGCX. *Pharmacogenet Genomics*. 2015;25(7):343-353. doi:10.1097/FPC.000000000000143
26. Zhang X, Wang W, Zhu W, et al. Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels. *Int J Mol Sci*. 2019;20(22):5573. doi:10.3390/ijms20225573
27. Ponting CP, Oliver PL, Reik W. Evolution and Functions of Long Noncoding RNAs. *Cell*. 2009;136(4):629-641. doi:10.1016/j.cell.2009.02.006
28. Pinkney HR, Wright BM, Diermeier SD. The lncRNA Toolkit: Databases and In Silico Tools for lncRNA Analysis. *Noncoding RNA*. 2020;6(4):49. doi:10.3390/ncrna6040049
29. Fritah S, Niclou SP, Azuaje F. Databases for lncRNAs: a comparative evaluation of emerging tools. *RNA*. 2014;20(11):1655-1665. doi:10.1261/rna.044040.113
30. Lim LJ, Jin Y, Yang H, et al. Network of clinically-relevant lncRNAs-mRNAs associated with prognosis of hepatocellular carcinoma patients. *Sci Rep*. 2020;10(1):11124. doi:10.1038/s41598-020-67742-8
31. GENCODE - Human Release 43. Accessed May 13, 2023. <https://www.encodegenes.org/human/>
32. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res*. 2012;22(9):1760-1774. doi:10.1101/gr.135350.111

33. Wong T, Wang Z, Chapron BD, et al. Polymorphic Human Sulfotransferase 2A1 Mediates the Formation of 25-Hydroxyvitamin D3-3-O-Sulfate, a Major Circulating Vitamin D Metabolite in Humans. *Drug Metab Dispos.* 2018;46(4):367-379.
doi:10.1124/dmd.117.078428
34. Tanner JA, Prasad B, Claw KG, et al. Predictors of Variation in CYP2A6 mRNA, Protein, and Enzyme Activity in a Human Liver Bank: Influence of Genetic and Nongenetic Factors. *J Pharmacol Exp Ther.* 2017;360(1):129-139. doi:10.1124/jpet.116.237594
35. warfarin. PharmGKB. Accessed May 13, 2023.
<https://www.pharmgkb.org/chemical/PA451906/overview>
36. HUGO Gene Nomenclature Committee (HGNC), European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, United Kingdom. HGNC Database. www.genenames.org
37. Seal RL, Braschi B, Gray K, et al. Genenames.org: the HGNC resources in 2023. *Nucleic Acids Research.* 2023;51(D1):D1003-D1009. doi:10.1093/nar/gkac888
38. R Core Team. R: A Language and Environment for Statistical Computing. Published online 2021. <https://www.R-project.org/>
39. Verma A, Jiang Y, Du W, Fairchild L, Melnick A, Elemento O. Transcriptome sequencing reveals thousands of novel long non-coding RNAs in B cell lymphoma. *Genome Medicine.* 2015;7(1):110. doi:10.1186/s13073-015-0230-7
40. Liu W, Yuan X, Yuan S, et al. Optimal reference genes for gene expression analysis in polyploid of *Cyprinus carpio* and *Carassius auratus*. *BMC Genetics.* 2020;21(1):107.
doi:10.1186/s12863-020-00915-6

41. Hart T, Komori HK, LaMere S, Podshivalova K, Salomon DR. Finding the active genes in deep RNA-seq gene expression studies. *BMC Genomics*. 2013;14:778. doi:10.1186/1471-2164-14-778
42. Sun X, Liu Y, Gao X, et al. Analysis of LncRNA-mRNA Co-Expression Profiles in Patients With Polycystic Ovary Syndrome: A Pilot Study. *Front Immunol*. 2021;12:669819. doi:10.3389/fimmu.2021.669819
43. Sarropoulos I, Marin R, Cardoso-Moreira M, Kaessmann H. Developmental dynamics of lncRNAs across mammalian organs and species. *Nature*. 2019;571(7766):510-514. doi:10.1038/s41586-019-1341-x
44. scale function - RDocumentation. Accessed August 14, 2023. <https://www.rdocumentation.org/packages/base/versions/3.6.2/topics/scale>
45. Chen W, Lin C, Gong L, et al. Comprehensive Analysis of the mRNA-lncRNA Co-expression Profile and ceRNA Networks Patterns in Chronic Hepatitis B. *Curr Genomics*. 2019;20(4):231-245. doi:10.2174/1389202920666190820122126
46. heatmap.2 function - RDocumentation. Accessed August 14, 2023. <https://www.rdocumentation.org/packages/gplots/versions/3.1.3/topics/heatmap.2>
47. An Introduction to `glmnet`. Accessed July 23, 2023. <https://glmnet.stanford.edu/articles/glmnet.html>
48. Goeman J, Meijer R, Chaturvedi N. L1 and L2 Penalized Regression Models.
49. Wadelius M, Pirmohamed M. Pharmacogenetics of warfarin: current status and future challenges. *Pharmacogenomics J*. 2007;7(2):99-111. doi:10.1038/sj.tpj.6500417

50. Liu SJ, Nowakowski TJ, Pollen AA, et al. Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biology*. 2016;17(1):67. doi:10.1186/s13059-016-0932-1
51. Gaikwad T, Ghosh K, Avery P, Kamali F, Shetty S. Warfarin Dose Model for the Prediction of Stable Maintenance Dose in Indian Patients. *Clin Appl Thromb Hemost*. 2018;24(2):353-359. doi:10.1177/1076029616683046
52. Cooper GM, Johnson JA, Langae TY, et al. A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood*. 2008;112(4):1022-1027. doi:10.1182/blood-2008-01-134247
53. Chen J, Shao L, Gong L, et al. A Pharmacogenetics-Based Warfarin Maintenance Dosing Algorithm from Northern Chinese Patients. *PLoS One*. 2014;9(8):e105250. doi:10.1371/journal.pone.0105250

APPENDIX A

Table A.1. Summary statistics and standard deviations for the 25 warfarin pharmacogenes remaining after quality control.

Gene	Min	1st.Qu.	Median	Mean	3rd.Qu.	Max	SD
ABCB1	3.47283	11.39528	14.0895	15.25248	18.47135	45.939	5.586434
APOB	56.6022	236.1805	305.4875	311.6172	389.8828	638.626	108.6237
APOC1	858.271	4575.938	5679.38	6143.544	7400.513	19586.5	2468.444
APOE	315.499	1207.81	1508.83	1565.678	1849.505	4354.54	489.8797
ASPH	7.44598	20.23743	26.19115	27.28762	32.2582	70.5811	9.923605
CALU	4.77255	17.74413	23.16385	25.07554	30.676	92.2947	10.89785
CRP	111.815	5343.465	8365.31	8825.518	11847.25	23666.7	4780.248
CYP1A1	0.04812	1.468225	4.264965	14.06402	13.3799	364.011	29.89541
CYP1A2	0.057568	6.831493	25.12685	48.86818	58.86633	476.125	66.70895
CYP2A6	0.178129	45.6917	196.2315	332.181	472.6578	3226.85	418.312
CYP2C18	3.26676	26.21328	35.7127	36.71177	45.05948	113.407	14.97425
CYP2C19	0.281154	3.645005	8.762135	24.50781	26.84888	397.613	44.2911
CYP2C9	3.78548	155.7345	322.1445	363.4997	493.5328	1949.01	275.4752
CYP3A4	0.47718	77.52435	343.347	615.2998	832.2583	5295.39	818.1293
CYP4F11	5.01366	17.53633	21.5661	22.83022	26.70455	55.2294	7.473433
CYP4F2	2.24293	34.26535	48.2255	48.02643	61.18753	125.118	19.85032
DDHD1	0.301714	0.768309	0.998556	1.076139	1.297278	3.95283	0.471109
DNMT3A	0.626033	1.533103	2.03757	2.111	2.611473	4.84271	0.73638
EPHX1	11.2531	140.8068	270.799	332.7301	443.304	2315.74	279.6908
VKORC1	7.19159	12.02393	14.0672	14.53997	16.37193	36.4982	3.744865
F11	3.82818	26.88383	37.1654	36.74977	44.9277	81.4363	13.91755
F13A1	0.025539	0.49266	0.995269	1.66421	1.76234	27.6858	2.728628
F2	134.136	374	435.606	436.8736	511.902	778.68	104.3457
F5	14.592	67.37645	89.25715	92.43826	110.8443	281.909	33.51051
F7	5.96837	25.2339	34.2263	33.00444	41.44248	74.5187	11.77459

Table A.2. Summary statistics and standard deviations for the 71 lncRNA genes remaining after quality control.

Gene	Min	1st.Qu.	Median	Mean	3rd.Qu.	Max	SD
A1BG.AS1	0	0.51509	0.841945	1.002505	1.30119	4.20856	0.700376
A2M.AS1	0	0.25044	0.371718	0.434024	0.545236	1.90143	0.278775
ABCC6P1	0.468686	6.922845	9.49967	9.980353	12.93365	27.6853	4.437036
ABCC6P2	0.124616	2.20707	3.142215	3.451828	4.262555	23.9518	2.097093
ADORA2A.AS1	0.170268	2.15133	3.393105	3.65055	4.899303	11.7548	2.00904
AGAP2.AS1	0.043101	0.327878	0.475396	0.597912	0.741187	3.47364	0.443155
ALOX12P2	0.358038	2.39311	3.431265	3.950678	4.940973	16.7405	2.343204
AMZ2P1	0.41976	1.084495	1.4842	1.651132	2.03252	6.18228	0.804779
ANO1.AS1	0.128192	3.779415	8.65795	10.87384	14.58883	59.3417	9.661188
AOC4P	0.131213	1.84	2.856445	3.387439	4.509373	21.8755	2.219452
APOA1.AS	1.07588	14.72378	23.80135	27.2752	35.98755	135.309	17.60253
APOC1P1	4.98006	40.0004	89.38525	126.6911	174.8638	764.656	120.1548
ARAP1.AS1	2.96673	11.9548	17.0306	20.60511	26.57395	84.5969	12.54499
ARHGAP5.AS1	0	0.262345	0.401579	0.432156	0.550048	1.7934	0.243814
ARSD.AS1	0.819138	5.590943	7.63905	8.67276	11.41773	25.0595	3.974538
ASB16.AS1	0.032185	0.734621	0.958736	1.058863	1.303383	3.49344	0.486481
ASH1L.AS1	0	0.410575	0.60269	0.64427	0.828256	2.74426	0.353763
ASMTL.AS1	0.450469	3.52311	6.27497	7.911859	10.38745	50.1852	6.639841
AZGP1P1	3.61264	16.21223	21.7836	23.18067	28.43098	106.824	9.989375
BCRP3	0	0.25907	0.44586	0.504182	0.664024	2.41215	0.341731
BDNF.AS	0.085268	0.45209	0.705023	0.739507	0.961433	2.40984	0.386072
BMS1P1	0.179045	0.977021	1.36693	1.427652	1.73941	4.92621	0.631919
BTN2A3P	0.039904	0.252131	0.385024	0.426655	0.539007	1.44115	0.245549
C1RL.AS1	0.661195	2.862318	4.243385	4.646011	5.947588	14.2382	2.404263
C3P1	2.02598	34.1459	47.92825	56.7829	72.90463	311.757	36.39183
C8orf44	0	0.400547	0.634796	0.717547	0.928957	4.08229	0.466129
CA5BP1	1.22362	3.077943	3.77293	3.962248	4.636805	11.4768	1.329751
CACTIN.AS1	0.214109	0.628436	0.793318	0.832032	1.006205	2.43231	0.305603
CASK.AS1	0	0.857702	1.22065	1.347457	1.67815	6.61308	0.723245
CATSPER2P1	0.241029	0.919307	1.269605	1.359557	1.762548	3.73912	0.603634
CCT6P3	1.12477	2.171983	2.776095	3.10549	3.697825	11.3229	1.37179
CD27.AS1	1.26996	5.440908	7.031335	7.465412	8.827688	27.1811	3.06744
CD99P1	0.074283	0.74051	0.95474	1.053943	1.288193	3.56625	0.501969
CERS6.AS1	0.87768	3.26378	4.22575	4.325461	5.179115	20.0052	1.821604
CES1P1	65.4861	230.0788	307.713	335.1701	396.983	1490.23	169.1292
CKMT2.AS1	0.177461	1.636085	2.19939	2.280111	2.810373	5.27555	0.918186
CPB2.AS1	0.05661	0.585237	0.897671	0.982088	1.251853	3.56248	0.560224
CRNDE	0	0.549562	0.877107	1.152935	1.430248	19.3123	1.336866
CROCCP2	6.99699	15.00915	19.8493	21.17816	25.40713	72.1778	8.473022
CRYBB2P1	0.619219	2.348095	3.120455	3.873373	4.504075	20.8819	2.51396

CTC.338M12.4	0	0.408804	0.575209	0.627788	0.794705	3.14679	0.350062
CYP2B7P	0.033768	1.636535	6.60501	20.30215	22.3264	450.449	39.70348
DANCR	6.62002	13.808	18.95175	21.53735	26.10568	67.8668	11.1083
DBH.AS1	0.177165	5.72643	11.151	14.47363	20.10378	74.21	11.84375
DENND5B.AS1	0.274711	1.041575	1.531505	1.63452	2.097003	5.30386	0.813345
DGCR5	0.058223	0.853452	1.2285	1.447792	1.80045	5.80897	0.913071
DHRS4.AS1	0.623214	3.65192	5.7134	6.556252	8.6196	24.485	3.893668
DICER1.AS1	0.080228	0.544966	0.859942	1.024681	1.308225	4.24778	0.683222
DIO3OS	0.098957	1.383803	2.477485	3.764759	4.851845	36.3323	4.087627
DLEU2	0.043475	0.459889	0.64816	0.728996	0.861422	4.86361	0.446359
DLGAP1.AS1	0.086954	0.996878	1.55832	2.907808	2.570725	49.0196	5.367976
DNAJC9.AS1	4.56631	10.41655	12.58125	13.84313	15.57258	68.6972	5.926717
DNM3OS	0	0.571076	0.894843	1.046923	1.31002	6.99936	0.710275
DPY19L2P4	0.025979	0.207373	0.288345	0.309651	0.390584	0.989653	0.140612
EAFL.AS1	0	0.368449	0.60671	0.715125	0.890494	8.4192	0.65637
ENTPD1.AS1	0.252316	0.899266	1.242815	1.266288	1.59	3.09344	0.516706
ENTPD3.AS1	0.832143	3.104023	3.77987	4.110983	4.834088	16.2125	1.804741
EPB41L4A.AS1	1.44504	3.733905	5.1752	5.887723	6.9599	36.3613	3.769785
ERVK13.1	0.633561	1.83731	2.69336	2.995219	3.935553	8.99098	1.515124
F11.AS1	2.26983	12.37598	15.5759	15.92713	19.62028	43.877	6.064397
FAM13A.AS1	0.048328	0.242195	0.35668	0.402103	0.511002	1.30185	0.225375
FAM157C	0.238802	1.54683	2.27152	2.562704	3.18046	10.6377	1.576864
FAM83A.AS1	0.078714	2.364185	6.73806	14.95362	14.92945	254.439	27.40961
FAM86B3P	0.136577	0.780998	1.130415	1.289283	1.573653	5.11285	0.736551
FAM86C2P	0.248856	0.90554	1.23634	1.275773	1.5455	3.78221	0.513994
FAM86DP	0.14734	0.909296	1.279165	1.44737	1.689778	8.01708	0.952178
FAM86EP	0.214001	0.866901	1.18538	1.306362	1.633803	4.2926	0.634446
FAM99A	3.96E-07	7.054168	10.64375	13.20278	17.24345	70.13	9.862955
FAM99B	0	4.724005	8.480125	9.408519	12.26415	36.8168	6.250469
FGD5.AS1	3.15999	11.7631	13.7889	14.27287	16.46215	33.0617	3.640687
FLNB.AS1	0.072248	0.364357	0.545055	0.624212	0.77293	3.11662	0.390316

Table A.3. Summary statistics and standard deviations of 25 log₂-transformed warfarin pharmacogenes.

Gene	Min	1st.Qu.	Median	Mean	3rd.Qu.	Max	SD
ABCB1	1.796112	3.510364	3.816547	3.839047	4.207215	5.521648	0.520248
APOB	5.822786	7.883725	8.254969	8.184809	8.606896	9.318827	0.560759
APOC1	9.745289	12.15985	12.47152	12.48356	12.85341	14.25757	0.539042
APOE	8.301492	10.23818	10.55921	10.54377	10.85292	12.0883	0.452709
ASPH	2.896462	4.338954	4.711006	4.676571	5.011581	6.14121	0.527514
CALU	2.25476	4.14927	4.533802	4.517355	4.939038	6.528176	0.627428
CRP	6.80497	12.3835	13.0302	12.79558	13.53226	14.53057	1.135866
CYP1A1	-4.37721	0.554061	2.092521	2.17409	3.741994	8.507838	2.199044
CYP1A2	-4.1186	2.772198	4.651158	4.309796	5.879303	8.895197	2.248972
CYP2A6	-2.48901	5.51357	7.616409	7.104296	8.884629	11.65591	2.324833
CYP2C18	1.70786	4.712222	5.158365	5.065211	5.493759	6.825366	0.662007
CYP2C19	-1.83057	1.865915	3.131278	3.323198	4.746643	8.635221	1.892065
CYP2C9	1.920476	7.282945	8.331564	8.032717	8.946995	10.92853	1.322423
CYP3A4	-1.06739	6.276444	8.423522	7.860195	9.700886	12.37052	2.457711
CYP4F11	2.325864	4.132274	4.430693	4.43508	4.739007	5.787365	0.484927
CYP4F2	1.165385	5.098677	5.591724	5.432225	5.935165	6.967146	0.731579
DDHD1	-1.72875	-0.38024	-0.00209	-0.00669	0.375487	1.982886	0.560138
DNMT3A	-0.67569	0.616454	1.026847	0.988908	1.384861	2.275815	0.516219
EPHX1	3.492251	7.137571	8.081078	7.92757	8.792145	11.17726	1.203412
VKORC1	2.846311	3.58783	3.814259	3.818683	4.033152	5.189753	0.349014
F11	1.936659	4.748652	5.215886	5.071263	5.489529	6.3476	0.66381
F13A1	-5.29115	-1.02134	-0.00684	-0.04183	0.817492	4.791074	1.415251
F2	7.067553	8.546894	8.76688	8.724721	8.999724	9.604887	0.381867
F5	3.867106	6.074172	6.479894	6.437741	6.79239	8.139086	0.527529
F7	2.577337	4.657291	5.097033	4.931531	5.373038	6.219531	0.615058

Table A.4. Summary statistics and standard deviations of 71 log2-transformation lncRNA genes.

Gene	Min	1st.Qu.	Median	Mean	3rd.Qu.	Max	SD
A1BG.AS1	-5.31292	-0.94313	-0.24058	-0.35772	0.379831	2.073327	1.119547
A2M.AS1	-5.45799	-1.99493	-1.41394	-1.4627	-0.8743	0.927085	0.916882
ABCC6P1	-1.09331	2.791322	3.247875	3.149462	3.693057	4.791048	0.767946
ABCC6P2	-3.00444	1.142132	1.65178	1.546539	2.091718	4.582062	0.893911
ADORA2A.AS1	-2.55412	1.105227	1.7626	1.601824	2.292576	3.555178	0.961162
AGAP2.AS1	-4.53615	-1.60877	-1.0728	-1.0501	-0.43212	1.796448	0.953561
ALOX12P2	-1.48182	1.258884	1.77874	1.75074	2.304795	4.065271	0.843101
AMZ2P1	-1.25236	0.117023	0.569682	0.569104	1.023265	2.628139	0.668878
ANO1.AS1	-2.96362	1.918125	3.113865	2.823468	3.866785	5.890974	1.500091
AOC4P	-2.93002	0.879706	1.51422	1.478647	2.172925	4.451244	0.946926
APOA1.AS	0.105517	3.880075	4.572971	4.461516	5.169424	7.080114	1.01184
APOC1P1	2.316163	5.321927	6.481962	6.353165	7.450087	9.578667	1.434736
ARAP1.AS1	1.568874	3.579518	4.090057	4.129593	4.73194	6.402533	0.830968
ARHGAP5.AS1	-4.4519	-1.8862	-1.27078	-1.39082	-0.84109	0.842697	0.849347
ARSD.AS1	-0.28782	2.48309	2.933391	2.96156	3.513201	4.647286	0.690614
ASB16.AS1	-4.95749	-0.44493	-0.0608	-0.06795	0.382246	1.804648	0.693776
ASH1L.AS1	-3.99289	-1.27669	-0.7221	-0.83253	-0.26042	1.456417	0.825731
ASMTL.AS1	-1.1505	1.816843	2.649606	2.519645	3.376769	5.64919	1.230289
AZGP1P1	1.853053	4.019004	4.44517	4.411842	4.829389	6.739092	0.607497
BCRP3	-5.15542	-1.9231	-1.15291	-1.30162	-0.56856	1.27032	1.058045
BDNF.AS	-3.55185	-1.14532	-0.50426	-0.64937	-0.05675	1.268937	0.830184
BMS1P1	-2.48161	-0.03355	0.450937	0.372204	0.798597	2.300478	0.659134
BTN2A3P	-4.64732	-1.98775	-1.37698	-1.45723	-0.89162	0.527221	0.837655
C1RL.AS1	-0.59685	1.517183	2.085213	2.015095	2.572291	3.831695	0.793477
C3P1	1.01862	5.093632	5.582804	5.566816	6.187918	8.284278	0.907199
C8orf44	-11.1249	-1.29616	-0.64961	-0.78286	-0.09944	2.029379	1.094865
CA5BP1	0.291156	1.621966	1.915684	1.909706	2.21313	3.520649	0.473103
CACTIN.AS1	-2.22358	-0.67016	-0.33403	-0.36092	0.008924	1.282327	0.535423
CASK.AS1	-3.02562	-0.2185	0.287641	0.24135	0.746868	2.725322	0.769849
CATSPER2P1	-2.05272	-0.12138	0.344374	0.288286	0.817653	1.902699	0.701734
CCT6P3	0.16963	1.119012	1.473045	1.519168	1.886652	3.501172	0.563649
CD27.AS1	0.344783	2.443847	2.813799	2.786058	3.142036	4.764532	0.58816
CD99P1	-3.75082	-0.43341	-0.06682	-0.08296	0.365347	1.834408	0.704251
CERS6.AS1	-0.18823	1.706542	2.079206	1.993833	2.372705	4.322303	0.603633
CES1P1	6.033117	7.845982	8.265441	8.249771	8.632933	10.54132	0.617879
CKMT2.AS1	-2.49443	0.71022	1.137098	1.057175	1.490754	2.399322	0.658338
CPB2.AS1	-4.1428	-0.77291	-0.15574	-0.27728	0.324061	1.832882	0.91902
CRNDE	-3.30097	-0.8399	-0.11687	-0.20462	0.516264	4.271448	1.095125
CROCCP2	2.806734	3.90777	4.311015	4.300506	4.667161	6.173483	0.5453
CRYBB2P1	-0.69148	1.231491	1.64175	1.723242	2.171218	4.384181	0.795609
CTC.338M12.4	-211.361	-1.26743	-0.77701	-1.44505	-0.31055	1.653881	11.12505

CYP2B7P	-4.88822	0.710377	2.723561	2.550206	4.480679	8.81522	2.523968
DANCR	2.726836	3.78743	4.244253	4.263993	4.706289	6.084634	0.678986
DBH.AS1	-2.49683	2.517634	3.479098	3.336296	4.329394	6.213542	1.352718
DENND5B.AS1	-1.86401	0.058751	0.614944	0.521816	1.068326	2.407043	0.767021
DGCR5	-4.10226	-0.22862	0.296897	0.245813	0.848352	2.538282	0.980778
DHRS4.AS1	-0.6822	1.868655	2.514349	2.468968	3.107621	4.613826	0.860901
DICER1.AS1	-3.63975	-0.87577	-0.21769	-0.26679	0.387611	2.086709	0.967327
DIO3OS	-3.33705	0.46863	1.308866	1.309789	2.278522	5.183181	1.333645
DLEU2	-4.52368	-1.12064	-0.62559	-0.67768	-0.21521	2.282028	0.824889
DLGAP1.AS1	-3.52361	-0.00452	0.63999	0.706232	1.362175	5.615287	1.368736
DNAJC9.AS1	2.191029	3.380806	3.653203	3.692789	3.960935	6.102179	0.513735
DNM3OS	-3.00311	-0.75211	-0.13345	-0.18873	0.389588	2.807223	0.921256
DPY19L2P4	-5.26653	-2.2697	-1.79413	-1.84532	-1.35632	-0.01501	0.700629
EAF1.AS1	-25.2374	-1.39418	-0.68649	-0.86205	-0.12754	3.073683	1.645232
ENTPD1.AS1	-1.9867	-0.15319	0.313611	0.204401	0.669026	1.629212	0.664277
ENTPD3.AS1	-0.2651	1.63413	1.918337	1.923554	2.273244	4.019035	0.578249
EPB41L4A.AS1	0.531109	1.900678	2.371612	2.371824	2.799066	5.184332	0.693099
ERVK13.1	-0.65844	0.877593	1.429405	1.396655	1.976562	3.168478	0.751521
F11.AS1	1.182584	3.629464	3.961243	3.873406	4.294273	5.455393	0.631173
FAM13A.AS1	-4.37101	-2.04576	-1.48733	-1.54694	-0.9686	0.380563	0.861708
FAM157C	-2.06611	0.629314	1.183651	1.093413	1.669226	3.411114	0.919432
FAM83A.AS1	-3.66723	1.241296	2.752333	2.537489	3.900011	7.991176	2.084306
FAM86B3P	-2.87221	-0.35663	0.176846	0.152215	0.654117	2.354128	0.804146
FAM86C2P	-2.00662	-0.14315	0.306074	0.233739	0.628073	1.919229	0.599911
FAM86DP	-2.76278	-0.1372	0.355198	0.321006	0.756829	3.003077	0.768085
FAM86EP	-2.22431	-0.20606	0.24535	0.219756	0.708233	2.101852	0.713493
FAM99A	-21.2691	2.818475	3.411933	3.237999	4.107957	6.13196	1.781041
FAM99B	-0.86927	2.240008	3.084083	2.883979	3.616375	5.202292	1.093197
FGD5.AS1	1.65992	3.556196	3.785435	3.787388	4.041081	5.047089	0.380282
FLNB.AS1	-3.7909	-1.45658	-0.87554	-0.92954	-0.37159	1.639982	0.86483

Table A.5. Summary statistics and standard deviations of 25 log₂-transformed and scaled warfarin pharmacogenes.

Gene	Min	1st.Qu.	Median	Mean	3rd.Qu.	Max	SD
ABCB1	-3.92685	-0.63178	-0.04325	-1.55E-16	0.707677	3.234229	1
APOB	-4.21219	-0.53692	0.125116	1.23E-15	0.752707	2.022292	1
APOC1	-5.07989	-0.60053	-0.02234	6.15E-16	0.686121	3.291044	1
APOE	-4.95303	-0.67504	0.034106	1.52E-15	0.682886	3.411754	1
ASPH	-3.37452	-0.64001	0.065278	-7.82E-17	0.635073	2.776491	1
CALU	-3.60614	-0.58666	0.026213	-5.10E-16	0.672081	3.204861	1
CRP	-5.27404	-0.36278	0.206562	-6.36E-16	0.648569	1.527464	1
CYP1A1	-2.97916	-0.7367	-0.03709	-7.78E-17	0.712993	2.880228	1
CYP1A2	-3.74767	-0.68369	0.151786	-1.65E-16	0.697877	2.038887	1
CYP2A6	-4.12645	-0.68423	0.220279	-1.72E-16	0.765789	1.957824	1
CYP2C18	-5.07147	-0.53321	0.140715	-1.08E-16	0.647346	2.658815	1
CYP2C19	-2.72388	-0.77021	-0.10143	-1.85E-17	0.752324	2.807526	1
CYP2C9	-4.622	-0.56697	0.225984	4.17E-16	0.691366	2.189775	1
CYP3A4	-3.63248	-0.6444	0.229208	-5.95E-17	0.748945	1.835173	1
CYP4F11	-4.34955	-0.62443	-0.00905	-8.31E-16	0.626749	2.788636	1
CYP4F2	-5.83237	-0.45593	0.21802	9.51E-17	0.687472	2.098092	1
DDHD1	-3.07434	-0.66689	0.008217	-8.57E-18	0.682299	3.551948	1
DNMT3A	-3.2246	-0.7215	0.073495	1.08E-16	0.767026	2.49295	1
EPHX1	-3.68562	-0.65647	0.127561	1.28E-16	0.718436	2.700394	1
VKORC1	-2.78606	-0.66144	-0.01268	-6.15E-16	0.614501	3.928416	1
F11	-4.72214	-0.486	0.217869	2.94E-16	0.630099	1.922746	1
F13A1	-3.70911	-0.69211	0.024725	2.38E-17	0.60719	3.414878	1
F2	-4.33965	-0.46568	0.110402	1.43E-15	0.720152	2.304899	1
F5	-4.87297	-0.68919	0.079906	-6.19E-16	0.672283	3.22512	1
F7	-3.8276	-0.44588	0.269083	-5.65E-16	0.71783	2.094112	1

Table A.6. Summary statistics and standard deviations of 71 log2-transformed and scaled lncRNA

genes.

Gene	Min	1st.Qu.	Median	Mean	3rd.Qu.	Max	SD
A1BG.AS1	-4.42607	-0.52289	0.104635	4.75E-18	0.658797	2.171458	1
A2M.AS1	-4.35748	-0.58048	0.053181	8.88E-17	0.641742	2.606422	1
ABCC6P1	-5.52482	-0.46636	0.128151	2.47E-16	0.707856	2.137632	1
ABCC6P2	-5.09109	-0.4524	0.117732	1.15E-17	0.609881	3.395778	1
ADORA2A.AS1	-4.32388	-0.51666	0.167273	6.19E-17	0.718664	2.032285	1
AGAP2.AS1	-3.65583	-0.58588	-0.02381	-7.34E-18	0.648072	2.985174	1
ALOX12P2	-3.83413	-0.58339	0.033211	-1.81E-17	0.657164	2.745261	1
AMZ2P1	-2.72317	-0.67588	0.000864	-3.38E-17	0.678989	3.078344	1
ANO1.AS1	-3.85783	-0.60353	0.193586	5.24E-17	0.695502	2.04488	1
AOC4P	-4.65577	-0.63251	0.037566	-8.08E-17	0.733191	3.139208	1
APOA1.AS	-4.30503	-0.57464	0.110151	1.34E-16	0.699625	2.587956	1
APOC1P1	-2.81376	-0.71877	0.089771	-2.05E-16	0.764546	2.24815	1
ARAP1.AS1	-3.08161	-0.66197	-0.04758	4.62E-16	0.724874	2.735293	1
ARHGAP5.AS1	-3.60404	-0.58325	0.141331	-1.23E-18	0.647242	2.62969	1
ARSD.AS1	-4.70506	-0.69282	-0.04079	-1.91E-16	0.798769	2.44091	1
ASB16.AS1	-7.04771	-0.54337	0.010315	9.31E-19	0.648913	2.699144	1
ASH1L.AS1	-3.82735	-0.5379	0.133739	1.75E-17	0.692845	2.772022	1
ASMTL.AS1	-2.98316	-0.57125	0.105634	8.49E-17	0.696685	2.543749	1
AZGP1P1	-4.21202	-0.64665	0.054862	1.31E-17	0.687325	3.830884	1
BCRP3	-3.64238	-0.58738	0.140556	-3.33E-17	0.692845	2.430842	1
BDNF.AS	-3.49619	-0.5974	0.174799	5.99E-17	0.713848	2.310706	1
BMS1P1	-4.32963	-0.61558	0.11945	-3.99E-17	0.646899	2.925465	1
BTN2A3P	-3.80835	-0.63334	0.095806	-1.17E-16	0.675231	2.36906	1
C1RL.AS1	-3.29177	-0.62751	0.088369	2.81E-16	0.702222	2.289417	1
C3P1	-5.01345	-0.52159	0.017624	2.15E-16	0.684637	2.995441	1
C8orf44	-9.44599	-0.46883	0.121704	1.99E-18	0.624203	2.56857	1
CA5BP1	-3.42113	-0.6082	0.012635	-1.06E-16	0.641349	3.405054	1
CACTIN.AS1	-3.47886	-0.57757	0.050219	-4.96E-17	0.690749	3.069061	1
CASK.AS1	-4.24364	-0.59732	0.06013	6.42E-18	0.656646	3.226569	1
CATSPER2P1	-3.33603	-0.58379	0.079927	-1.62E-17	0.754369	2.300604	1
CCT6P3	-2.39429	-0.70994	-0.08183	3.01E-17	0.651974	3.516382	1
CD27.AS1	-4.1507	-0.58183	0.047164	2.75E-16	0.605238	3.363834	1
CD99P1	-5.20817	-0.49762	0.02292	-1.58E-17	0.636575	2.722565	1
CERS6.AS1	-3.61489	-0.47594	0.141432	1.64E-16	0.627653	3.857424	1
CES1P1	-3.58752	-0.65351	0.025362	1.59E-17	0.620126	3.708736	1
CKMT2.AS1	-5.3948	-0.52702	0.121402	1.67E-16	0.658596	2.038689	1
CPB2.AS1	-4.20614	-0.5393	0.132244	-1.15E-17	0.654328	2.296099	1
CRNDE	-2.8274	-0.5801	0.080133	1.29E-17	0.658269	4.087272	1
CROCCP2	-2.73936	-0.72022	0.019272	5.43E-16	0.67239	3.434762	1
CRYBB2P1	-3.03506	-0.61808	-0.10243	1.03E-16	0.56306	3.344532	1

CTC.338M12.4	-18.8687	0.015966	0.060048	3.28E-18	0.101977	0.278554	1
CYP2B7P	-2.94711	-0.72894	0.068683	6.14E-17	0.764856	2.482208	1
DANCR	-2.2639	-0.70187	-0.02907	-3.22E-16	0.651406	2.681411	1
DBH.AS1	-4.31216	-0.6052	0.105566	-1.29E-16	0.73415	2.127011	1
DENND5B.AS1	-3.11051	-0.60372	0.121415	2.15E-17	0.712509	2.457854	1
DGCR5	-4.43329	-0.48373	0.052086	-3.50E-18	0.614348	2.337399	1
DHRS4.AS1	-3.66031	-0.69731	0.052714	8.66E-17	0.741842	2.49141	1
DICER1.AS1	-3.48689	-0.62955	0.050753	8.89E-18	0.676499	2.432987	1
DIO3OS	-3.48432	-0.63072	-0.00069	-5.29E-18	0.72638	2.904365	1
DLEU2	-4.66244	-0.53699	0.063156	1.96E-17	0.560651	3.58801	1
DLGAP1.AS1	-3.09032	-0.51927	-0.0484	4.02E-17	0.479232	3.58656	1
DNAJC9.AS1	-2.92322	-0.60728	-0.07705	4.21E-16	0.521955	4.689947	1
DNM3OS	-3.05494	-0.61154	0.060001	3.89E-18	0.627749	3.252031	1
DPY19L2P4	-4.88305	-0.60572	0.07306	-3.71E-17	0.697948	2.612387	1
EAF1.AS1	-14.8158	-0.32344	0.106705	-3.46E-17	0.446445	2.392205	1
ENTPD1.AS1	-3.29847	-0.53831	0.164404	1.85E-17	0.699444	2.144906	1
ENTPD3.AS1	-3.78496	-0.50052	-0.00902	-1.57E-16	0.604738	3.623836	1
EPB41L4A.AS1	-2.65577	-0.67977	-0.00031	-2.82E-16	0.616424	4.057874	1
ERVK13.1	-2.73459	-0.69068	0.043579	-1.21E-16	0.771645	2.357651	1
F11.AS1	-4.26321	-0.38649	0.139165	2.03E-16	0.666801	2.506424	1
FAM13A.AS1	-3.2773	-0.57888	0.069178	-4.04E-17	0.671151	2.236837	1
FAM157C	-3.43639	-0.50477	0.098145	-9.10E-17	0.62627	2.520796	1
FAM83A.AS1	-2.97688	-0.62188	0.103077	5.14E-17	0.653705	2.616549	1
FAM86B3P	-3.76104	-0.63278	0.03063	-1.84E-17	0.624143	2.7382	1
FAM86C2P	-3.73448	-0.62825	0.120577	-9.31E-19	0.657321	2.809567	1
FAM86DP	-4.0149	-0.59656	0.044516	-2.39E-17	0.567416	3.491895	1
FAM86EP	-3.42549	-0.59681	0.035871	-9.00E-18	0.684628	2.637861	1
FAM99A	-13.76	-0.23555	0.097658	-8.32E-17	0.488454	1.62487	1
FAM99B	-3.43328	-0.58907	0.183045	-3.77E-18	0.669958	2.120673	1
FGD5.AS1	-5.59445	-0.60795	-0.00513	-2.35E-16	0.667119	3.312547	1
FLNB.AS1	-3.30858	-0.6094	0.062447	-5.82E-17	0.64516	2.971136	1

APPENDIX B

Table B.1. LncRNA genes selected as predictors to explain the expression of warfarin pharmacogene *ABCB1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
AGAP2.AS1	AGAP2.AS1	ABCC6P1
ANO1.AS1	ANO1.AS1	ABCC6P2
APOA1.AS	APOA1.AS	ADORA2A.AS1
APOC1P1	APOC1P1	AGAP2.AS1
ARAP1.AS1	ARAP1.AS1	ALOX12P2
ASH1L.AS1	ASH1L.AS1	ANO1.AS1
ASMTL.AS1	ASMTL.AS1	AOC4P
AZGP1P1	AZGP1P1	APOA1.AS
BMS1P1	BMS1P1	APOC1P1
CA5BP1	CA5BP1	ARAP1.AS1
CACTIN.AS1	CACTIN.AS1	ARHGAP5.AS1
CASK.AS1	CASK.AS1	ARSD.AS1
CCT6P3	CCT6P3	ASB16.AS1
CERS6.AS1	CERS6.AS1	ASH1L.AS1
CES1P1	CES1P1	ASMTL.AS1
CKMT2.AS1	CKMT2.AS1	AZGP1P1
CROCCP2	CROCCP2	BCRP3
CYP2B7P	CYP2B7P	BDNF.AS
DENND5B.AS1	DENND5B.AS1	BMS1P1
DHRS4.AS1	DHRS4.AS1	BTN2A3P
DIO3OS	DIO3OS	C1RL.AS1
DNM3OS	DNM3OS	C3P1
DPY19L2P4	DPY19L2P4	CA5BP1
EPB41L4A.AS1	EPB41L4A.AS1	CACTIN.AS1
FAM157C	FAM157C	CASK.AS1
FAM86DP	FAM86DP	CATSPER2P1
FLNB.AS1	FLNB.AS1	CCT6P3
		CD27.AS1
		CD99P1
		CERS6.AS1
		CES1P1
		CKMT2.AS1
		CRNDE
		CROCCP2

CRYBB2P1
CYP2B7P
DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

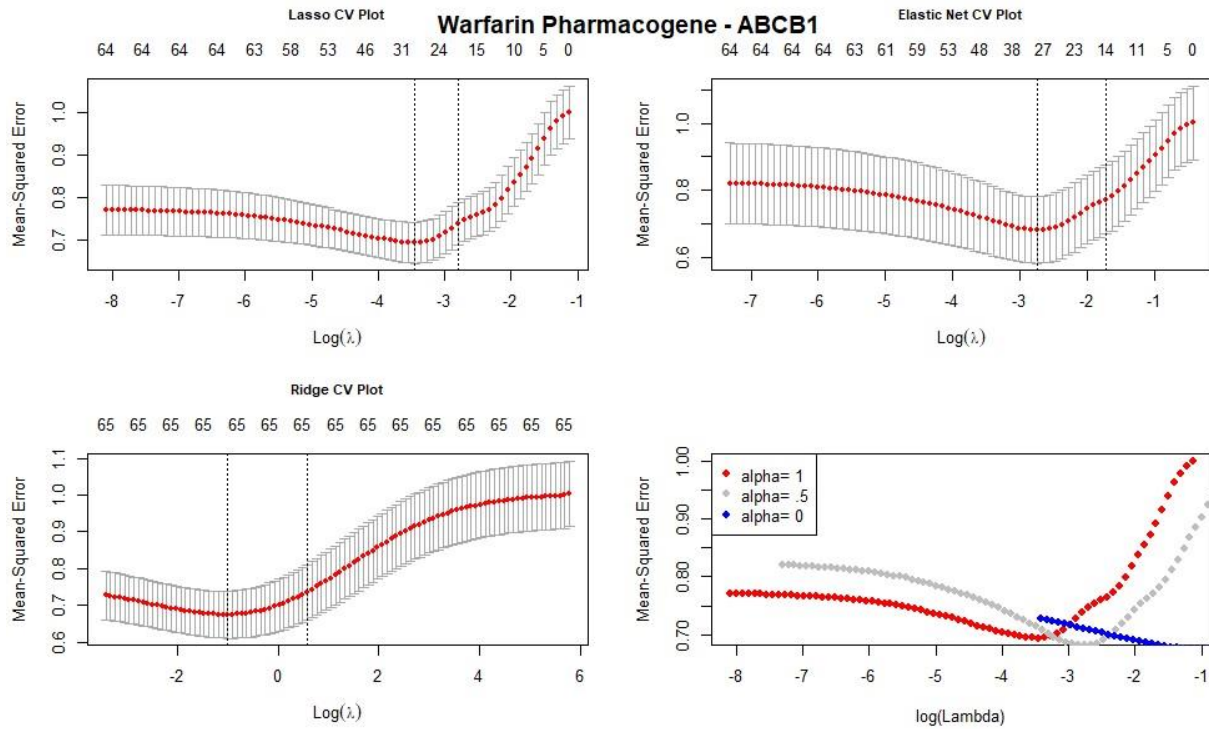


Figure B.1. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *ABCB1*.

Table B.2. LncRNA genes selected as predictors to explain the expression of warfarin pharmacogenes *APOB* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A1BG.AS1	A1BG.AS1
ABCC6P1	A2M.AS1	A2M.AS1
AOC4P	ABCC6P1	ABCC6P1
APOA1.AS	AGAP2.AS1	ABCC6P2
APOC1P1	ALOX12P2	ADORA2A.AS1
ARAP1.AS1	AOC4P	AGAP2.AS1
ARHGAP5.AS1	APOA1.AS	ALOX12P2
ARSD.AS1	APOC1P1	ANO1.AS1
ASB16.AS1	ARAP1.AS1	AOC4P
BMS1P1	ARHGAP5.AS1	APOA1.AS
CA5BP1	ARSD.AS1	APOC1P1
CATSPER2P1	ASB16.AS1	ARAP1.AS1
CCT6P3	ASMTL.AS1	ARHGAP5.AS1
CD27.AS1	BMS1P1	ARSD.AS1
CERS6.AS1	CA5BP1	ASB16.AS1
CRNDE	CATSPER2P1	ASH1L.AS1
CYP2B7P	CCT6P3	ASMTL.AS1
DENND5B.AS1	CD27.AS1	AZGP1P1
DICER1.AS1	CERS6.AS1	BCRP3
DLEU2	CKMT2.AS1	BDNF.AS
DNAJC9.AS1	CRNDE	BMS1P1
DPY19L2P4	CYP2B7P	BTN2A3P
EAF1.AS1	DANCR	C1RL.AS1
ENTPD1.AS1	DENND5B.AS1	C3P1
EPB41L4A.AS1	DICER1.AS1	CA5BP1
FAM157C	DLEU2	CACTIN.AS1
FAM83A.AS1	DNAJC9.AS1	CASK.AS1
FGD5.AS1	DPY19L2P4	CATSPER2P1
FLNB.AS1	EAF1.AS1	CCT6P3
	ENTPD1.AS1	CD27.AS1
	ENTPD3.AS1	CD99P1
	EPB41L4A.AS1	CERS6.AS1
	F11.AS1	CES1P1
	FAM13A.AS1	CKMT2.AS1
	FAM157C	CRNDE
	FAM83A.AS1	CROCCP2
	FAM86C2P	CRYBB2P1
	FAM86DP	CYP2B7P

FGD5.AS1
FLNB.AS1

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

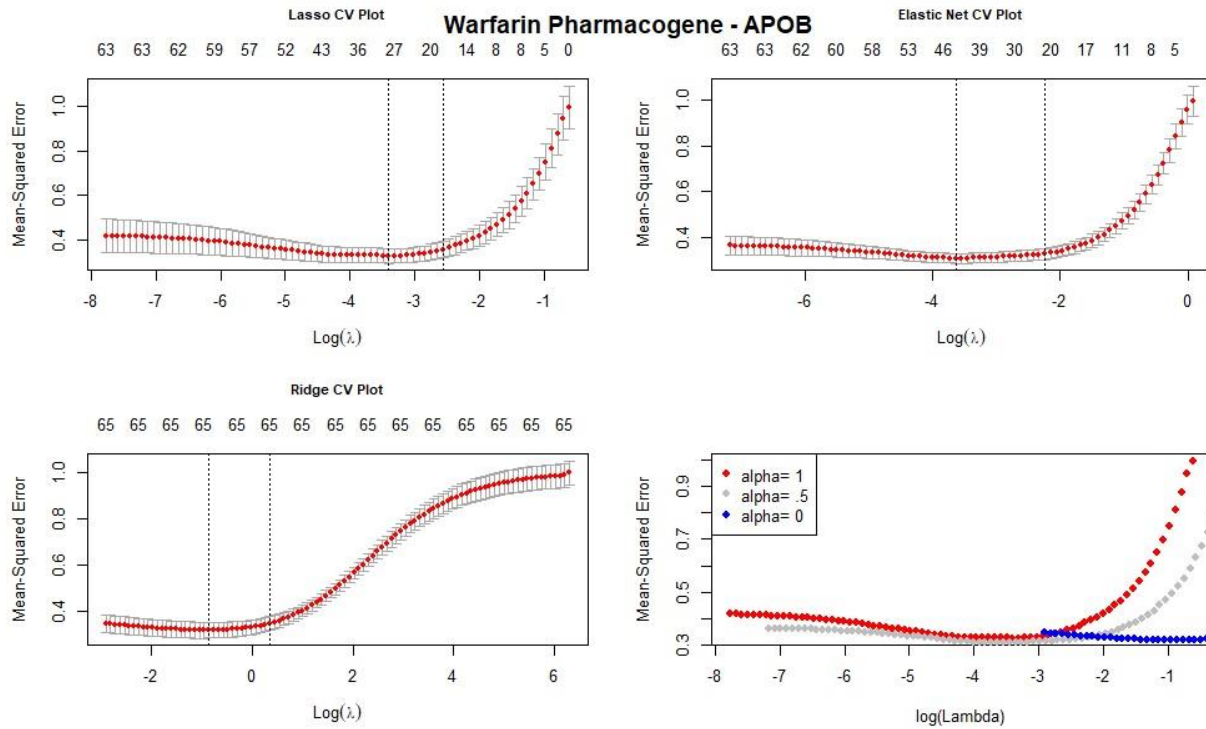


Figure B.2. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *APOB*.

Table B.3. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *APOC1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
ANO1.AS1	ADORA2A.AS1	A1BG.AS1
APOA1.AS	ANO1.AS1	A2M.AS1
APOC1P1	AOC4P	ABCC6P1
ARAP1.AS1	APOA1.AS	ABCC6P2
ARHGAP5.AS1	APOC1P1	ADORA2A.AS1
ASB16.AS1	ARAP1.AS1	AGAP2.AS1
AZGP1P1	ARHGAP5.AS1	ALOX12P2
BCRP3	ASB16.AS1	ANO1.AS1
BDNF.AS	ASH1L.AS1	AOC4P
BMS1P1	ASMTL.AS1	APOA1.AS
BTN2A3P	AZGP1P1	APOC1P1
C3P1	BCRP3	ARAP1.AS1
CACTIN.AS1	BDNF.AS	ARHGAP5.AS1
CASK.AS1	BMS1P1	ARSD.AS1
CD27.AS1	BTN2A3P	ASB16.AS1
CERS6.AS1	C3P1	ASH1L.AS1
CES1P1	CA5BP1	ASMTL.AS1
CRNDE	CACTIN.AS1	AZGP1P1
CRYBB2P1	CASK.AS1	BCRP3
DBH.AS1	CCT6P3	BDNF.AS
DIO3OS	CD27.AS1	BMS1P1
DNAJC9.AS1	CERS6.AS1	BTN2A3P
DNM3OS	CES1P1	C1RL.AS1
DPY19L2P4	CRNDE	C3P1
ENTPD1.AS1	CROCCP2	CA5BP1
ENTPD3.AS1	CRYBB2P1	CACTIN.AS1
EPB41L4A.AS1	DBH.AS1	CASK.AS1
ERVK13.1	DIO3OS	CATSPER2P1
F11.AS1	DNAJC9.AS1	CCT6P3
FAM83A.AS1	DNM3OS	CD27.AS1
FAM86B3P	DPY19L2P4	CD99P1
FAM86DP	ENTPD1.AS1	CERS6.AS1
FAM99A	ENTPD3.AS1	CES1P1
FGD5.AS1	EPB41L4A.AS1	CKMT2.AS1
	ERVK13.1	CRNDE
	F11.AS1	CROCCP2
	FAM157C	CRYBB2P1
	FAM83A.AS1	CYP2B7P

FAM86B3P	DANCR
FAM86DP	DBH.AS1
FAM86EP	DENND5B.AS1
FAM99A	DHRS4.AS1
FGD5.AS1	DICER1.AS1
	DIO3OS
	DLEU2
	DNAJC9.AS1
	DNM3OS
	DPY19L2P4
	EAF1.AS1
	ENTPD1.AS1
	ENTPD3.AS1
	EPB41L4A.AS1
	ERVK13.1
	F11.AS1
	FAM13A.AS1
	FAM157C
	FAM83A.AS1
	FAM86B3P
	FAM86C2P
	FAM86DP
	FAM86EP
	FAM99A
	FAM99B
	FGD5.AS1
	FLNB.AS1

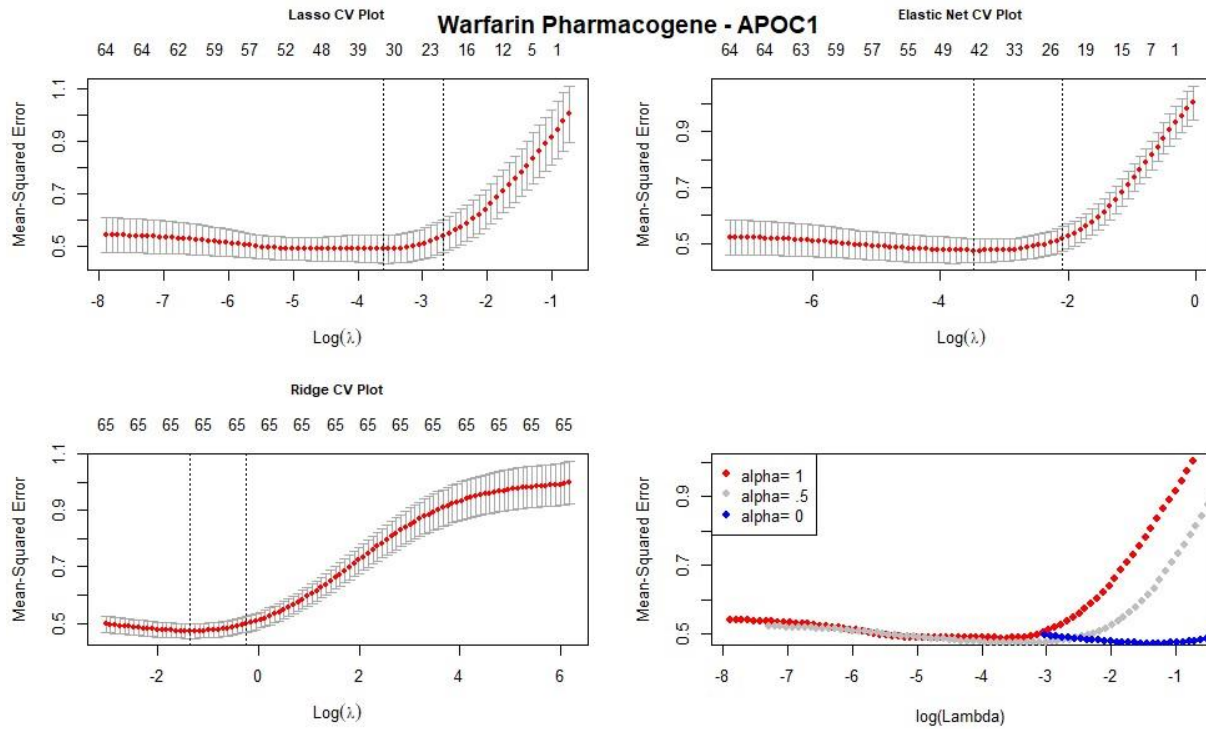


Figure B.3. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *APOC1*.

Table B.4. LncRNA genes selected as predictors to explain the expression of warfarin pharmacogene *APOE* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
ABCC6P2	ABCC6P2	A2M.AS1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P1
AGAP2.AS1	AGAP2.AS1	ABCC6P2
AOC4P	AOC4P	ADORA2A.AS1
APOA1.AS	APOA1.AS	AGAP2.AS1
APOC1P1	APOC1P1	ALOX12P2
ARAP1.AS1	ARAP1.AS1	ANO1.AS1
ARSD.AS1	ARSD.AS1	AOC4P
ASMTL.AS1	ASMTL.AS1	APOA1.AS
BCRP3	BCRP3	APOC1P1
BDNF.AS	BDNF.AS	ARAP1.AS1
BMS1P1	BMS1P1	ARHGAP5.AS1
C1RL.AS1	C1RL.AS1	ARSD.AS1
C3P1	C3P1	ASB16.AS1
CACTIN.AS1	CACTIN.AS1	ASH1L.AS1
CATSPER2P1	CATSPER2P1	ASMTL.AS1
CD99P1	CD99P1	AZGP1P1
CERS6.AS1	CERS6.AS1	BCRP3
CES1P1	CES1P1	BDNF.AS
CKMT2.AS1	CKMT2.AS1	BMS1P1
CRNDE	CRNDE	BTN2A3P
CROCCP2	CROCCP2	C1RL.AS1
CRYBB2P1	CRYBB2P1	C3P1
CYP2B7P	CYP2B7P	CA5BP1
DENND5B.AS1	DENND5B.AS1	CACTIN.AS1
DHRS4.AS1	DHRS4.AS1	CASK.AS1
DICER1.AS1	DICER1.AS1	CATSPER2P1
DIO3OS	DIO3OS	CCT6P3
DLEU2	DLEU2	CD27.AS1
DNAJC9.AS1	DNAJC9.AS1	CD99P1
DNM3OS	DNM3OS	CERS6.AS1
EAF1.AS1	EAF1.AS1	CES1P1
F11.AS1	F11.AS1	CKMT2.AS1
FAM13A.AS1	FAM13A.AS1	CRNDE
FAM157C	FAM157C	CROCCP2
FAM83A.AS1	FAM83A.AS1	CRYBB2P1
FAM86B3P	FAM86B3P	CYP2B7P

FAM86C2P	FAM86C2P	DANCR
FAM86DP	FAM86DP	DBH.AS1
FAM99A	FAM99A	DENND5B.AS1
FGD5.AS1	FGD5.AS1	DHRS4.AS1
		DICER1.AS1
		DIO3OS
		DLEU2
		DNAJC9.AS1
		DNM3OS
		DPY19L2P4
		EAF1.AS1
		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

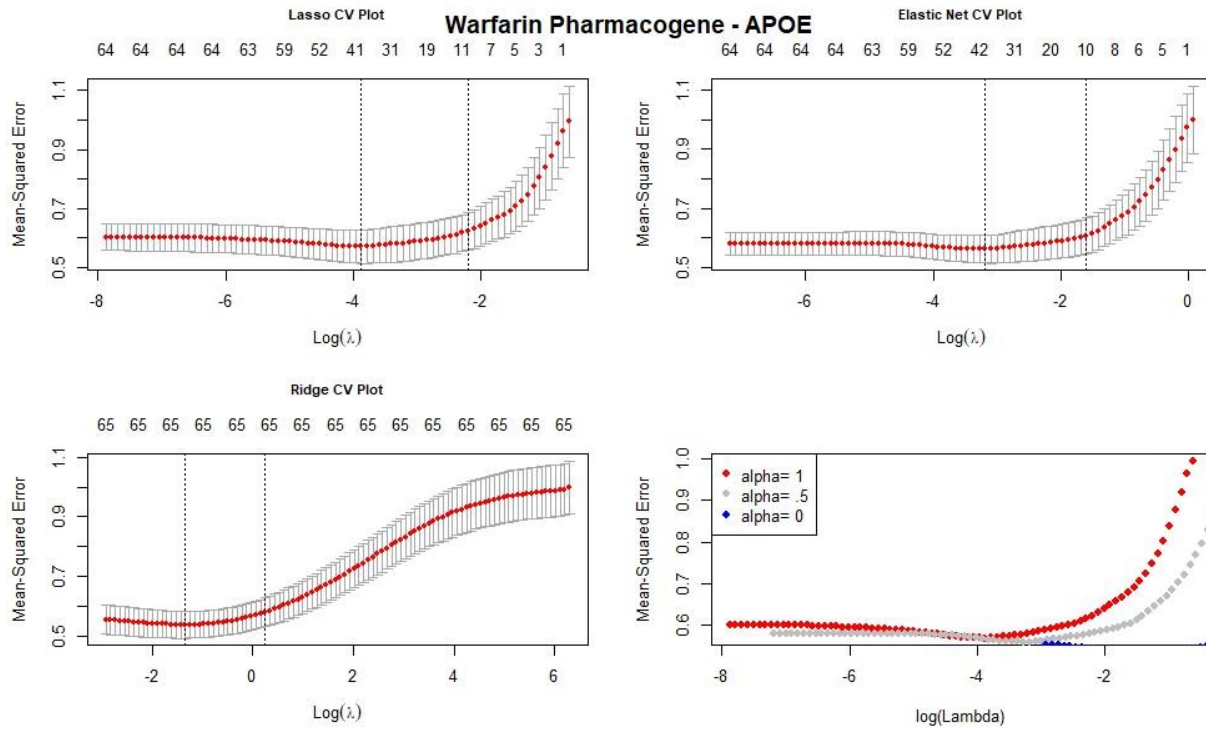


Figure B.4. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *APOE*.

Table B.5. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *ASPH* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ADORA2A.AS1	ANO1.AS1	ABCC6P2
ANO1.AS1	AOC4P	ADORA2A.AS1
AOC4P	APOA1.AS	AGAP2.AS1
APOA1.AS	ARAP1.AS1	ALOX12P2
ARAP1.AS1	ARHGAP5.AS1	ANO1.AS1
ARHGAP5.AS1	ARSD.AS1	AOC4P
ARSD.AS1	ASH1L.AS1	APOA1.AS
ASB16.AS1	AZGP1P1	APOC1P1
ASH1L.AS1	BCRP3	ARAP1.AS1
AZGP1P1	BDNF.AS	ARHGAP5.AS1
BCRP3	BMS1P1	ARSD.AS1
BDNF.AS	C1RL.AS1	ASB16.AS1
BMS1P1	C3P1	ASH1L.AS1
C1RL.AS1	CACTIN.AS1	ASMTL.AS1
C3P1	CATSPER2P1	AZGP1P1
CACTIN.AS1	CD99P1	BCRP3
CATSPER2P1	CERS6.AS1	BDNF.AS
CD99P1	CES1P1	BMS1P1
CERS6.AS1	CKMT2.AS1	BTN2A3P
CES1P1	CRNDE	C1RL.AS1
CKMT2.AS1	CROCCP2	C3P1
CRNDE	DENND5B.AS1	CA5BP1
CROCCP2	DICER1.AS1	CACTIN.AS1
DENND5B.AS1	DIO3OS	CASK.AS1
DICER1.AS1	DLEU2	CATSPER2P1
DIO3OS	DPY19L2P4	CCT6P3
DLEU2	EAF1.AS1	CD27.AS1
DPY19L2P4	EPB41L4A.AS1	CD99P1
EAF1.AS1	ERVK13.1	CERS6.AS1
EPB41L4A.AS1	F11.AS1	CES1P1
ERVK13.1	FAM157C	CKMT2.AS1
F11.AS1	FAM83A.AS1	CRNDE
FAM157C	FAM86C2P	CROCCP2
FAM83A.AS1	FAM86DP	CRYBB2P1
FAM86C2P	FAM99A	CYP2B7P

FAM86DP	FGD5.AS1	DANCR
FAM99A		DBH.AS1
FGD5.AS1		DENND5B.AS1
		DHRS4.AS1
		DICER1.AS1
		DIO3OS
		DLEU2
		DNAJC9.AS1
		DNM3OS
		DPY19L2P4
		EAF1.AS1
		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

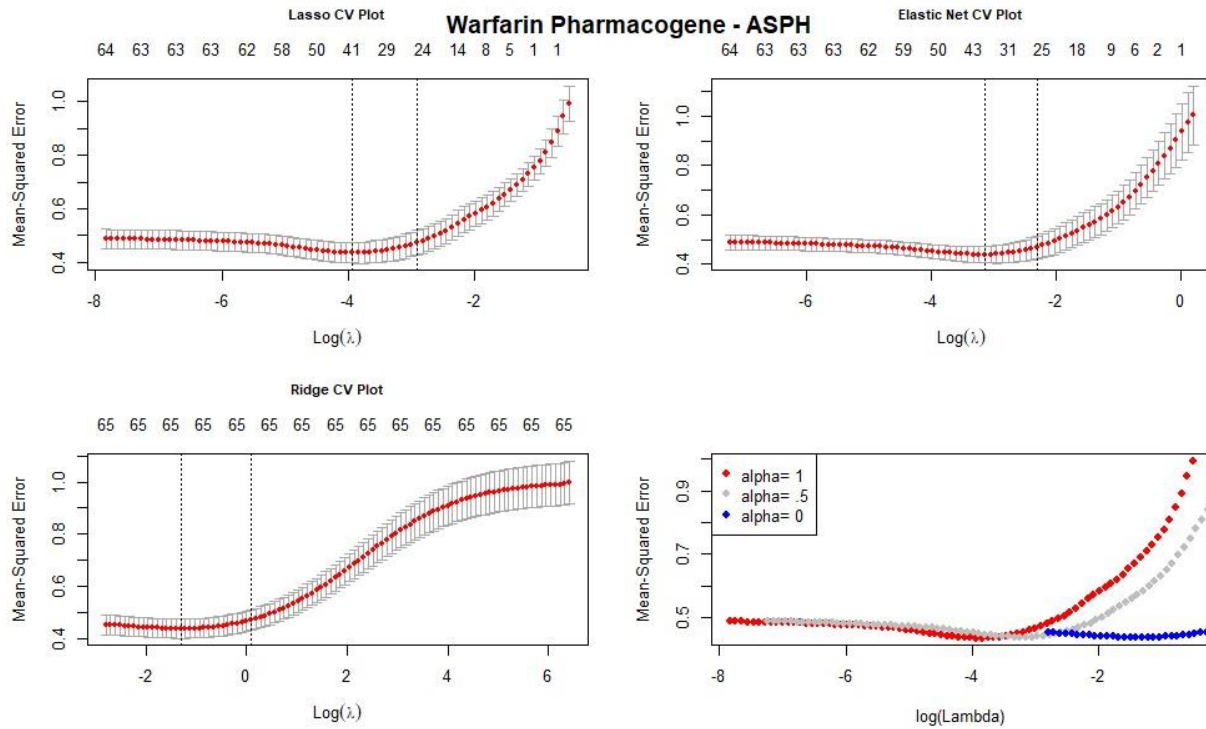


Figure B.5. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *ASPH*.

Table B.6. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CALU* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A2M.AS1	A1BG.AS1
ABCC6P2	ABCC6P2	A2M.AS1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P1
AGAP2.AS1	AGAP2.AS1	ABCC6P2
ANO1.AS1	ANO1.AS1	ADORA2A.AS1
AOC4P	AOC4P	AGAP2.AS1
APOA1.AS	APOA1.AS	ALOX12P2
APOC1P1	APOC1P1	ANO1.AS1
ARAP1.AS1	ARAP1.AS1	AOC4P
ARHGAP5.AS1	ARHGAP5.AS1	APOA1.AS
ARSD.AS1	ARSD.AS1	APOC1P1
ASB16.AS1	ASH1L.AS1	ARAP1.AS1
ASH1L.AS1	ASMTL.AS1	ARHGAP5.AS1
ASMTL.AS1	BCRP3	ARSD.AS1
BCRP3	BMS1P1	ASB16.AS1
BMS1P1	BTN2A3P	ASH1L.AS1
BTN2A3P	C3P1	ASMTL.AS1
C3P1	CA5BP1	AZGP1P1
CA5BP1	CACTIN.AS1	BCRP3
CACTIN.AS1	CATSPER2P1	BDNF.AS
CATSPER2P1	CD99P1	BMS1P1
CD99P1	CERS6.AS1	BTN2A3P
CERS6.AS1	CES1P1	C1RL.AS1
CES1P1	CKMT2.AS1	C3P1
CKMT2.AS1	CRNDE	CA5BP1
CRNDE	DBH.AS1	CACTIN.AS1
CROCCP2	DENND5B.AS1	CASK.AS1
DANCR	DHRS4.AS1	CATSPER2P1
DENND5B.AS1	DICER1.AS1	CCT6P3
DHRS4.AS1	DNAJC9.AS1	CD27.AS1
DICER1.AS1	DNM3OS	CD99P1
DLEU2	DPY19L2P4	CERS6.AS1
DNAJC9.AS1	EAF1.AS1	CES1P1
EAF1.AS1	ENTPD1.AS1	CKMT2.AS1
ENTPD1.AS1	ENTPD3.AS1	CRNDE
ENTPD3.AS1	EPB41L4A.AS1	CROCCP2
EPB41L4A.AS1	F11.AS1	CRYBB2P1
F11.AS1	FAM157C	CYP2B7P

FAM13A.AS1	FAM83A.AS1	DANCR
FAM157C	FAM86B3P	DBH.AS1
FAM83A.AS1	FAM86DP	DENND5B.AS1
FAM86B3P	FAM86EP	DHRS4.AS1
FAM86DP	FAM99B	DICER1.AS1
FAM86EP	FGD5.AS1	DIO3OS
FAM99B	FLNB.AS1	DLEU2
FGD5.AS1		DNAJC9.AS1
FLNB.AS1		DNM3OS
		DPY19L2P4
		EAF1.AS1
		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

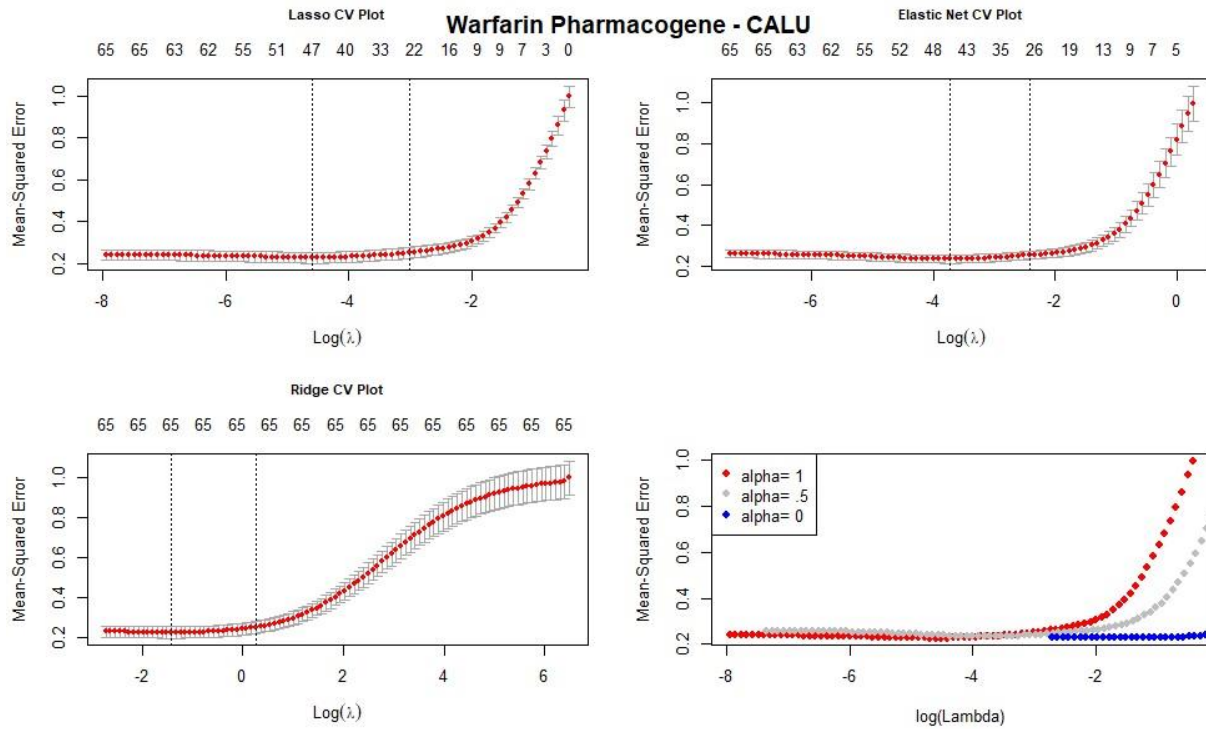


Figure B.6. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CALU*.

Table B.7. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CRP* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
ABCC6P2	ABCC6P2	ABCC6P1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P2
ALOX12P2	ALOX12P2	ADORA2A.AS1
ANO1.AS1	ANO1.AS1	AGAP2.AS1
AOC4P	AOC4P	ALOX12P2
APOA1.AS	APOA1.AS	ANO1.AS1
APOC1P1	APOC1P1	AOC4P
ARHGAP5.AS1	ARHGAP5.AS1	APOA1.AS
ASB16.AS1	ASB16.AS1	APOC1P1
ASH1L.AS1	ASH1L.AS1	ARAP1.AS1
ASMTL.AS1	ASMTL.AS1	ARHGAP5.AS1
C3P1	C3P1	ARSD.AS1
CACTIN.AS1	CACTIN.AS1	ASB16.AS1
CATSPER2P1	CATSPER2P1	ASH1L.AS1
CCT6P3	CCT6P3	ASMTL.AS1
CES1P1	CES1P1	AZGP1P1
CKMT2.AS1	CKMT2.AS1	BCRP3
CRNDE	CRNDE	BDNF.AS
CROCCP2	CROCCP2	BMS1P1
CYP2B7P	CYP2B7P	BTN2A3P
DANCR	DANCR	C1RL.AS1
DBH.AS1	DBH.AS1	C3P1
DENND5B.AS1	DENND5B.AS1	CA5BP1
DHRS4.AS1	DHRS4.AS1	CACTIN.AS1
DLEU2	DLEU2	CASK.AS1
DNAJC9.AS1	DNAJC9.AS1	CATSPER2P1
DNM3OS	DNM3OS	CCT6P3
EPB41L4A.AS1	EPB41L4A.AS1	CD27.AS1
FAM13A.AS1	FAM13A.AS1	CD99P1
FAM83A.AS1	FAM83A.AS1	CERS6.AS1
FAM86B3P	FAM86B3P	CES1P1
FAM86C2P	FAM86C2P	CKMT2.AS1
FAM86DP	FAM86DP	CRNDE
FAM86EP	FAM86EP	CROCCP2
FAM99A	FAM99A	CRYBB2P1
		CYP2B7P

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

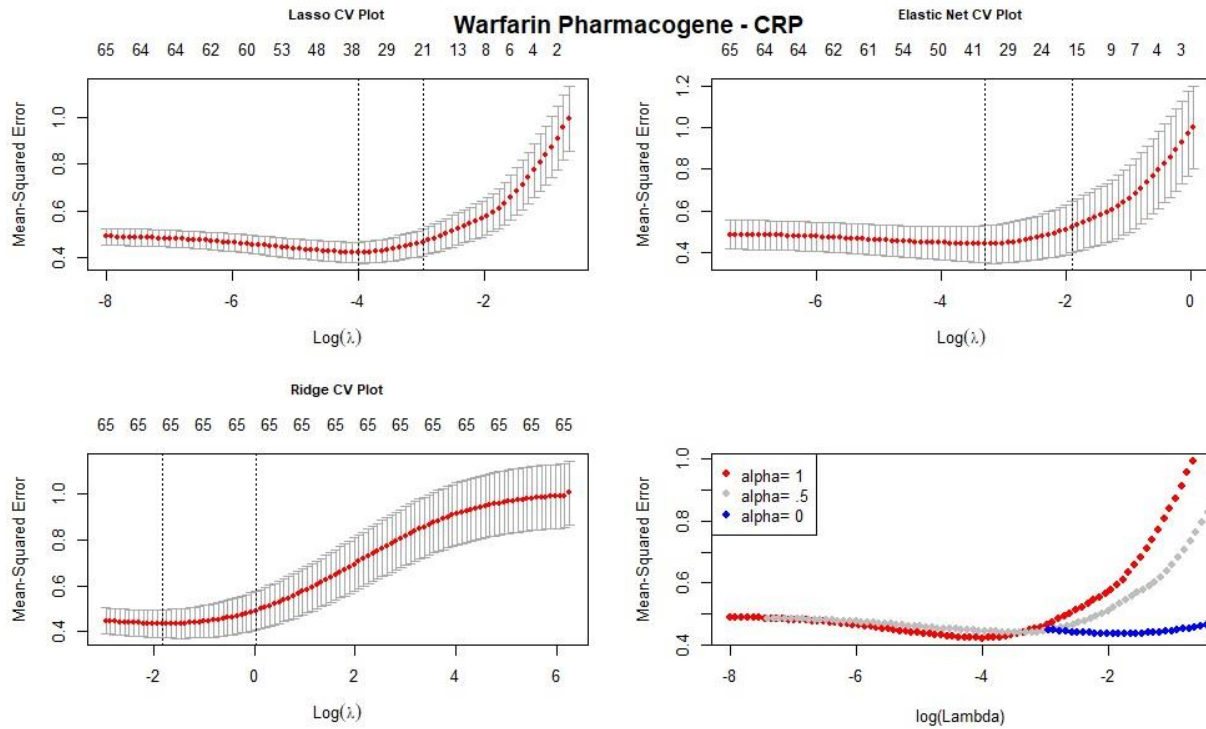


Figure B.7. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CRP*.

Table B.8. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP1A1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A2M.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
ABCC6P2	ABCC6P2	ABCC6P1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P2
AGAP2.AS1	AGAP2.AS1	ADORA2A.AS1
AOC4P	ANO1.AS1	AGAP2.AS1
APOA1.AS	AOC4P	ALOX12P2
APOC1P1	APOA1.AS	ANO1.AS1
ARSD.AS1	APOC1P1	AOC4P
ASB16.AS1	ARAP1.AS1	APOA1.AS
AZGP1P1	ARSD.AS1	APOC1P1
BDNF.AS	ASB16.AS1	ARAP1.AS1
CA5BP1	AZGP1P1	ARHGAP5.AS1
CACTIN.AS1	BDNF.AS	ARSD.AS1
CD27.AS1	C1RL.AS1	ASB16.AS1
CKMT2.AS1	CA5BP1	ASH1L.AS1
CROCCP2	CACTIN.AS1	ASMTL.AS1
CRYBB2P1	CATSPER2P1	AZGP1P1
DANCR	CCT6P3	BCRP3
DBH.AS1	CD27.AS1	BDNF.AS
DENND5B.AS1	CERS6.AS1	BMS1P1
DHRS4.AS1	CES1P1	BTN2A3P
DICER1.AS1	CKMT2.AS1	C1RL.AS1
DIO3OS	CRNDE	C3P1
DNM3OS	CROCCP2	CA5BP1
FAM13A.AS1	CRYBB2P1	CACTIN.AS1
FAM157C	DANCR	CASK.AS1
FAM83A.AS1	DBH.AS1	CATSPER2P1
FAM86B3P	DENND5B.AS1	CCT6P3
FLNB.AS1	DHRS4.AS1	CD27.AS1
	DICER1.AS1	CD99P1
	DIO3OS	CERS6.AS1
	DNAJC9.AS1	CES1P1
	DNM3OS	CKMT2.AS1
	EAF1.AS1	CRNDE
	ENTPD1.AS1	CROCCP2
	ERVK13.1	CRYBB2P1
	FAM13A.AS1	CYP2B7P

FAM157C	DANCR
FAM83A.AS1	DBH.AS1
FAM86B3P	DENND5B.AS1
FAM86C2P	DHRS4.AS1
FAM86EP	DICER1.AS1
FLNB.AS1	DIO3OS
	DLEU2
	DNAJC9.AS1
	DNM3OS
	DPY19L2P4
	EAF1.AS1
	ENTPD1.AS1
	ENTPD3.AS1
	EPB41L4A.AS1
	ERVK13.1
	F11.AS1
	FAM13A.AS1
	FAM157C
	FAM83A.AS1
	FAM86B3P
	FAM86C2P
	FAM86DP
	FAM86EP
	FAM99A
	FAM99B
	FGD5.AS1
	FLNB.AS1

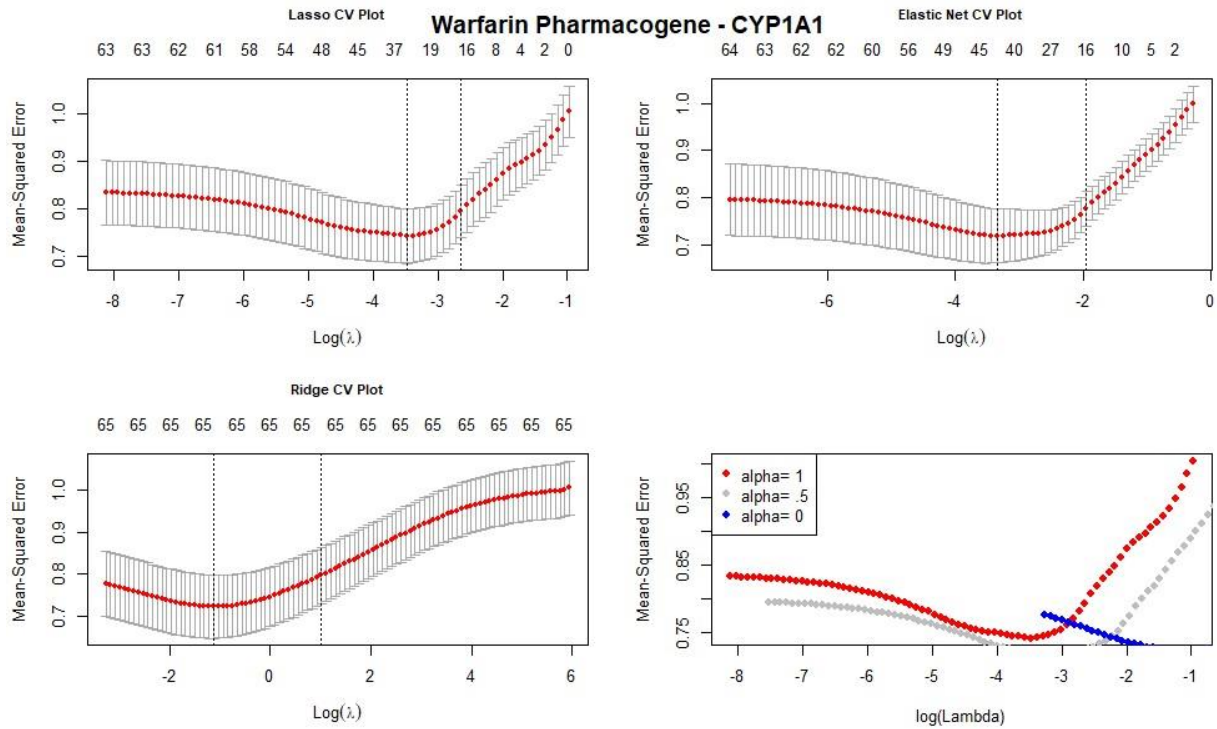


Figure B.8. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP1A1*.

Table B.9. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP1A2* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
ABCC6P1	ABCC6P1	A1BG.AS1
ABCC6P2	ABCC6P2	A2M.AS1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P1
AGAP2.AS1	AGAP2.AS1	ABCC6P2
ANO1.AS1	ANO1.AS1	ADORA2A.AS1
APOC1P1	APOC1P1	AGAP2.AS1
ARSD.AS1	ARSD.AS1	ALOX12P2
ASB16.AS1	ASB16.AS1	ANO1.AS1
ASH1L.AS1	ASH1L.AS1	AOC4P
AZGP1P1	AZGP1P1	APOA1.AS
BDNF.AS	BDNF.AS	APOC1P1
BMS1P1	BMS1P1	ARAP1.AS1
CA5BP1	CA5BP1	ARHGAP5.AS1
CATSPER2P1	CATSPER2P1	ARSD.AS1
CD27.AS1	CD27.AS1	ASB16.AS1
CERS6.AS1	CERS6.AS1	ASH1L.AS1
CKMT2.AS1	CKMT2.AS1	ASMTL.AS1
CRNDE	CRNDE	AZGP1P1
CRYBB2P1	CRYBB2P1	BCRP3
CYP2B7P	CYP2B7P	BDNF.AS
DBH.AS1	DBH.AS1	BMS1P1
DENND5B.AS1	DENND5B.AS1	BTN2A3P
DHRS4.AS1	DHRS4.AS1	C1RL.AS1
DICER1.AS1	DICER1.AS1	C3P1
DNAJC9.AS1	DNAJC9.AS1	CA5BP1
DNM3OS	DNM3OS	CACTIN.AS1
EAF1.AS1	EAF1.AS1	CASK.AS1
ENTPD1.AS1	ENTPD1.AS1	CATSPER2P1
ENTPD3.AS1	ENTPD3.AS1	CCT6P3
EPB41L4A.AS1	EPB41L4A.AS1	CD27.AS1
FAM157C	FAM157C	CD99P1
FAM83A.AS1	FAM83A.AS1	CERS6.AS1
FAM86B3P	FAM86B3P	CES1P1
FAM86DP	FAM86C2P	CKMT2.AS1
FAM99A	FAM86DP	CRNDE
FAM99B	FAM99A	CROCCP2
FLNB.AS1	FAM99B	CRYBB2P1
	FLNB.AS1	CYP2B7P

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

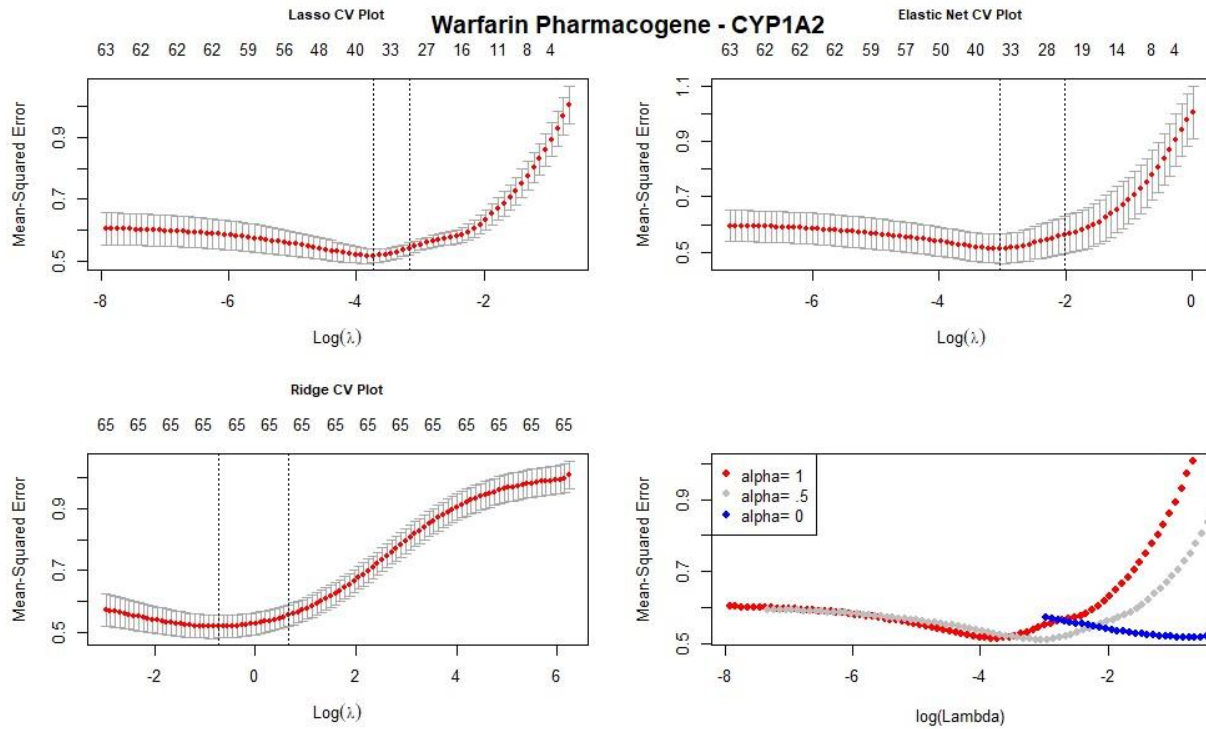


Figure B.9. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP1A2*.

Table B.10. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP2A6* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
AGAP2.AS1	ABCC6P2	ABCC6P1
ALOX12P2	AGAP2.AS1	ABCC6P2
ANO1.AS1	ALOX12P2	ADORA2A.AS1
AOC4P	ANO1.AS1	AGAP2.AS1
APOA1.AS	AOC4P	ALOX12P2
ASB16.AS1	APOA1.AS	ANO1.AS1
AZGP1P1	ARSD.AS1	AOC4P
BCRP3	ASB16.AS1	APOA1.AS
BMS1P1	AZGP1P1	APOC1P1
CASK.AS1	BCRP3	ARAP1.AS1
CATSPER2P1	BMS1P1	ARHGAP5.AS1
CCT6P3	CASK.AS1	ARSD.AS1
CD27.AS1	CATSPER2P1	ASB16.AS1
CERS6.AS1	CCT6P3	ASH1L.AS1
CES1P1	CD27.AS1	ASMTL.AS1
CKMT2.AS1	CERS6.AS1	AZGP1P1
CRNDE	CES1P1	BCRP3
CRYBB2P1	CKMT2.AS1	BDNF.AS
CYP2B7P	CRNDE	BMS1P1
DANCR	CRYBB2P1	BTN2A3P
DIO3OS	CYP2B7P	C1RL.AS1
DLEU2	DANCR	C3P1
DPY19L2P4	DICER1.AS1	CA5BP1
EAF1.AS1	DIO3OS	CACTIN.AS1
EPB41L4A.AS1	DLEU2	CASK.AS1
ERVK13.1	DPY19L2P4	CATSPER2P1
FAM86DP	EAF1.AS1	CCT6P3
FLNB.AS1	EPB41L4A.AS1	CD27.AS1
	ERVK13.1	CD99P1
	FAM86DP	CERS6.AS1
	FLNB.AS1	CES1P1
		CKMT2.AS1
		CRNDE
		CROCCP2
		CRYBB2P1
		CYP2B7P

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

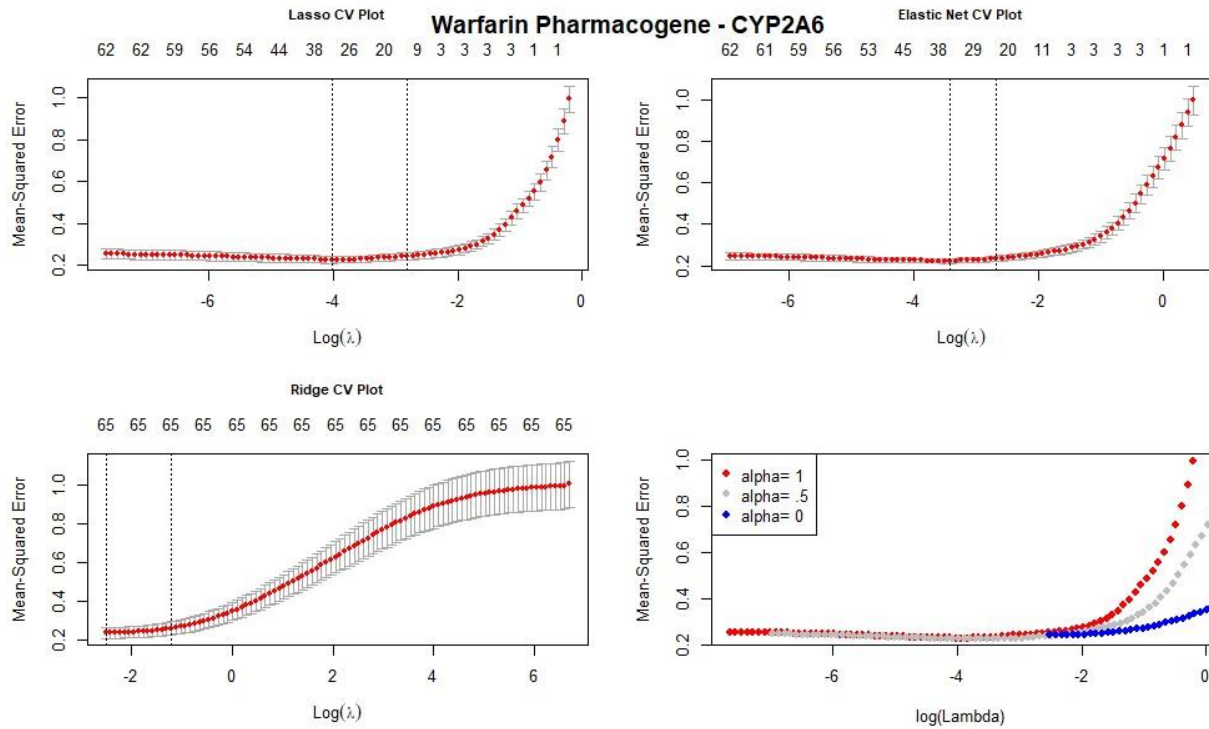


Figure B.10. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP2A6*.

Table B.11. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP2C18* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A2M.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
ABCC6P2	ABCC6P2	ABCC6P1
AGAP2.AS1	ALOX12P2	ABCC6P2
ALOX12P2	AOC4P	ADORA2A.AS1
ANO1.AS1	APOA1.AS	AGAP2.AS1
AOC4P	APOC1P1	ALOX12P2
APOA1.AS	ARAP1.AS1	ANO1.AS1
APOC1P1	ARHGAP5.AS1	AOC4P
ARAP1.AS1	ASH1L.AS1	APOA1.AS
ARHGAP5.AS1	AZGP1P1	APOC1P1
ASB16.AS1	BCRP3	ARAP1.AS1
ASH1L.AS1	BDNF.AS	ARHGAP5.AS1
AZGP1P1	C3P1	ARSD.AS1
BCRP3	CACTIN.AS1	ASB16.AS1
BDNF.AS	CATSPER2P1	ASH1L.AS1
C1RL.AS1	CCT6P3	ASMTL.AS1
C3P1	CERS6.AS1	AZGP1P1
CACTIN.AS1	CES1P1	BCRP3
CATSPER2P1	CRNDE	BDNF.AS
CCT6P3	CROCCP2	BMS1P1
CD27.AS1	CYP2B7P	BTN2A3P
CD99P1	DANCR	C1RL.AS1
CERS6.AS1	DBH.AS1	C3P1
CES1P1	DENND5B.AS1	CA5BP1
CRNDE	DHRS4.AS1	CACTIN.AS1
CROCCP2	DICER1.AS1	CASK.AS1
CYP2B7P	DIO3OS	CATSPER2P1
DANCR	DLEU2	CCT6P3
DBH.AS1	DNAJC9.AS1	CD27.AS1
DENND5B.AS1	DNM3OS	CD99P1
DHRS4.AS1	DPY19L2P4	CERS6.AS1
DICER1.AS1	ENTPD1.AS1	CES1P1
DIO3OS	ENTPD3.AS1	CKMT2.AS1
DLEU2	FAM13A.AS1	CRNDE
DNAJC9.AS1	FAM157C	CROCCP2
DNM3OS	FAM83A.AS1	CRYBB2P1
DPY19L2P4	FAM86B3P	CYP2B7P

ENTPD1.AS1	FAM86C2P	DANCR
ENTPD3.AS1	FAM86DP	DBH.AS1
FAM13A.AS1	FAM86EP	DENND5B.AS1
FAM157C	FAM99B	DHRS4.AS1
FAM83A.AS1	FGD5.AS1	DICER1.AS1
FAM86B3P	FLNB.AS1	DIO3OS
FAM86C2P		DLEU2
FAM86DP		DNAJC9.AS1
FAM86EP		DNM3OS
FAM99B		DPY19L2P4
FGD5.AS1		EAF1.AS1
FLNB.AS1		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

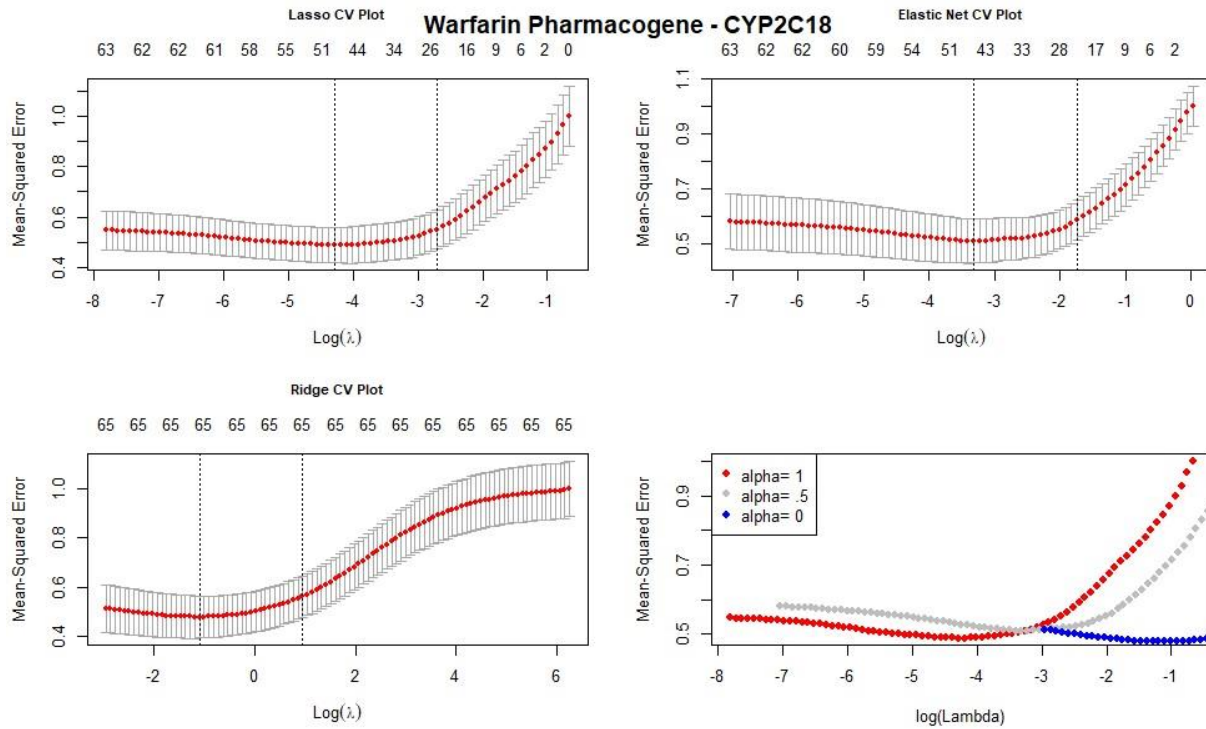


Figure B.11. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP2C18*.

Table B.12. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP2C19* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A2M.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
AGAP2.AS1	AGAP2.AS1	ABCC6P1
APOA1.AS	APOA1.AS	ABCC6P2
APOC1P1	APOC1P1	ADORA2A.AS1
ARSD.AS1	ARSD.AS1	AGAP2.AS1
ASB16.AS1	ASB16.AS1	ALOX12P2
ASH1L.AS1	ASH1L.AS1	ANO1.AS1
AZGP1P1	AZGP1P1	AOC4P
BDNF.AS	BDNF.AS	APOA1.AS
C1RL.AS1	BMS1P1	APOC1P1
C3P1	C1RL.AS1	ARAP1.AS1
CA5BP1	C3P1	ARHGAP5.AS1
CASK.AS1	CA5BP1	ARSD.AS1
CERS6.AS1	CASK.AS1	ASB16.AS1
CES1P1	CD27.AS1	ASH1L.AS1
CKMT2.AS1	CERS6.AS1	ASMTL.AS1
CRNDE	CES1P1	AZGP1P1
CYP2B7P	CKMT2.AS1	BCRP3
DBH.AS1	CRNDE	BDNF.AS
DENND5B.AS1	CRYBB2P1	BMS1P1
DICER1.AS1	CYP2B7P	BTN2A3P
DLEU2	DBH.AS1	C1RL.AS1
DNAJC9.AS1	DENND5B.AS1	C3P1
DNM3OS	DICER1.AS1	CA5BP1
ENTPD1.AS1	DLEU2	CACTIN.AS1
ENTPD3.AS1	DNAJC9.AS1	CASK.AS1
EPB41L4A.AS1	DNM3OS	CATSPER2P1
ERVK13.1	ENTPD1.AS1	CCT6P3
F11.AS1	ENTPD3.AS1	CD27.AS1
FAM13A.AS1	EPB41L4A.AS1	CD99P1
FAM157C	ERVK13.1	CERS6.AS1
FAM86B3P	F11.AS1	CES1P1
FAM86C2P	FAM13A.AS1	CKMT2.AS1
FAM86DP	FAM157C	CRNDE
FAM86EP	FAM86B3P	CROCCP2
	FAM86C2P	CRYBB2P1
	FAM86DP	CYP2B7P

FAM86EP

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

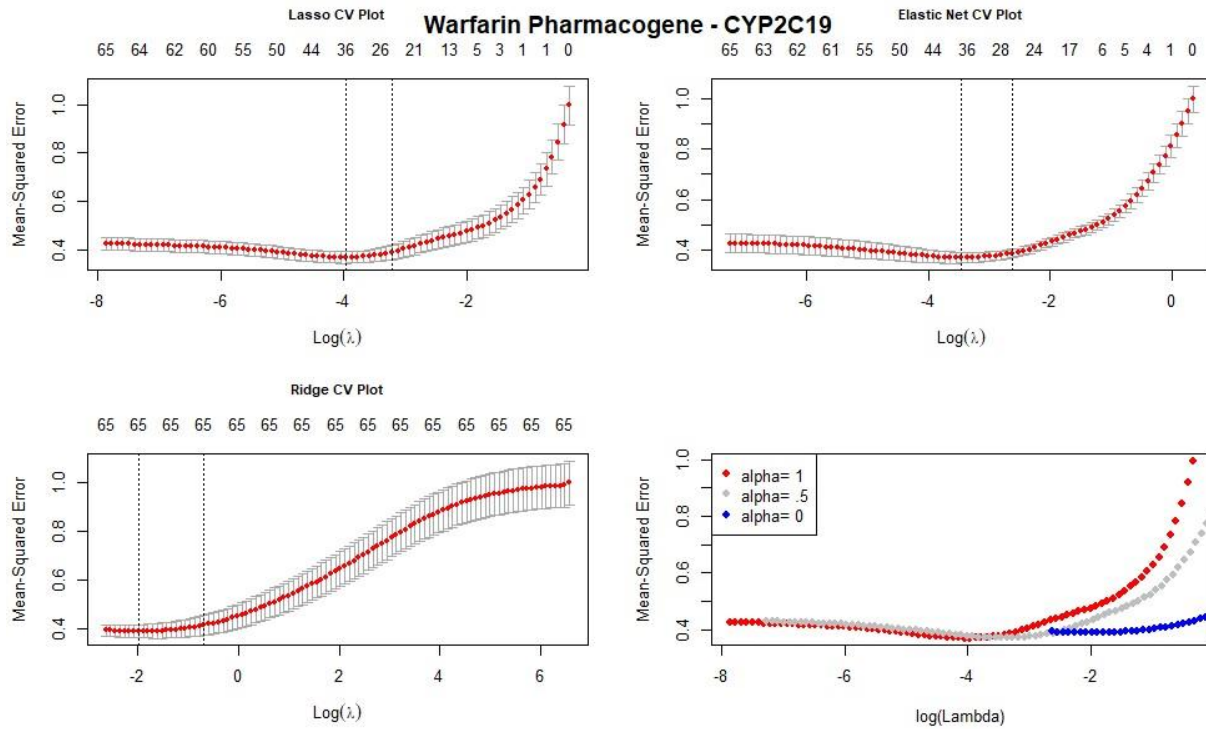


Figure B.12. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP2C19*.

Table B.13. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP2C9* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
ABCC6P1	ABCC6P1	A1BG.AS1
ABCC6P2	ABCC6P2	A2M.AS1
AGAP2.AS1	AGAP2.AS1	ABCC6P1
ALOX12P2	ALOX12P2	ABCC6P2
ANO1.AS1	ANO1.AS1	ADORA2A.AS1
APOA1.AS	APOA1.AS	AGAP2.AS1
ARHGAP5.AS1	ARHGAP5.AS1	ALOX12P2
ARSD.AS1	ARSD.AS1	ANO1.AS1
ASB16.AS1	ASB16.AS1	AOC4P
ASH1L.AS1	ASH1L.AS1	APOA1.AS
AZGP1P1	AZGP1P1	APOC1P1
BMS1P1	BMS1P1	ARAP1.AS1
CACTIN.AS1	CA5BP1	ARHGAP5.AS1
CASK.AS1	CACTIN.AS1	ARSD.AS1
CATSPER2P1	CASK.AS1	ASB16.AS1
CD27.AS1	CATSPER2P1	ASH1L.AS1
CD99P1	CD27.AS1	ASMTL.AS1
CERS6.AS1	CD99P1	AZGP1P1
CES1P1	CERS6.AS1	BCRP3
CRNDE	CES1P1	BDNF.AS
CROCCP2	CRNDE	BMS1P1
CRYBB2P1	CROCCP2	BTN2A3P
CYP2B7P	CRYBB2P1	C1RL.AS1
DHRS4.AS1	CYP2B7P	C3P1
DIO3OS	DHRS4.AS1	CA5BP1
DLEU2	DIO3OS	CACTIN.AS1
EAF1.AS1	DLEU2	CASK.AS1
ENTPD1.AS1	EAF1.AS1	CATSPER2P1
EPB41L4A.AS1	ENTPD1.AS1	CCT6P3
ERVK13.1	EPB41L4A.AS1	CD27.AS1
F11.AS1	ERVK13.1	CD99P1
FAM83A.AS1	F11.AS1	CERS6.AS1
FAM86B3P	FAM83A.AS1	CES1P1
FAM86DP	FAM86B3P	CKMT2.AS1
FAM99A	FAM86DP	CRNDE
FGD5.AS1	FAM99A	CROCCP2
FLNB.AS1	FGD5.AS1	CRYBB2P1
	FLNB.AS1	CYP2B7P

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

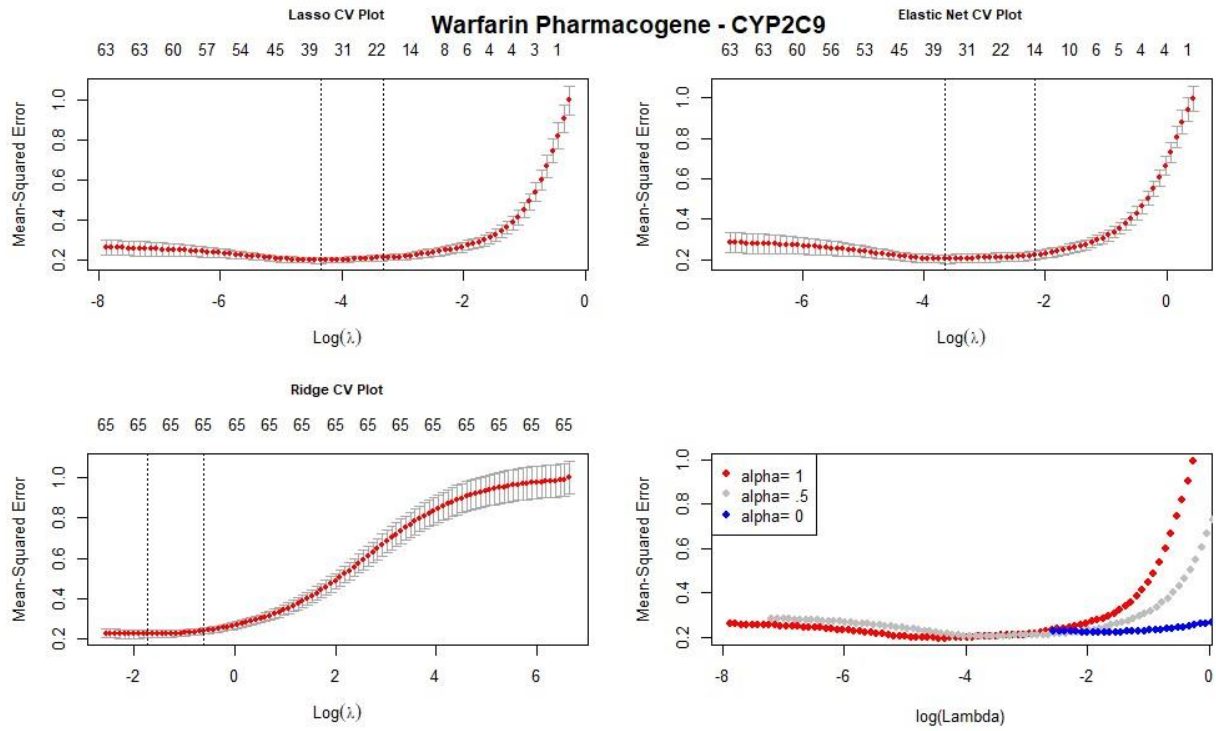


Figure B.13. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP2C9*.

Table B.14. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP3A4* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P1
ALOX12P2	ALOX12P2	ABCC6P2
ANO1.AS1	ANO1.AS1	ADORA2A.AS1
AOC4P	AOC4P	AGAP2.AS1
APOA1.AS	APOA1.AS	ALOX12P2
ARAP1.AS1	ARAP1.AS1	ANO1.AS1
ARHGAP5.AS1	ARHGAP5.AS1	AOC4P
ARSD.AS1	ARSD.AS1	APOA1.AS
ASH1L.AS1	ASH1L.AS1	APOC1P1
AZGP1P1	AZGP1P1	ARAP1.AS1
BDNF.AS	BDNF.AS	ARHGAP5.AS1
BMS1P1	BMS1P1	ARSD.AS1
C3P1	CASK.AS1	ASB16.AS1
CASK.AS1	CERS6.AS1	ASH1L.AS1
CD99P1	CES1P1	ASMTL.AS1
CERS6.AS1	CKMT2.AS1	AZGP1P1
CES1P1	CRNDE	BCRP3
CKMT2.AS1	CYP2B7P	BDNF.AS
CRNDE	DIO3OS	BMS1P1
CYP2B7P	DLEU2	BTN2A3P
DIO3OS	DNM3OS	C1RL.AS1
DLEU2	EPB41L4A.AS1	C3P1
DNM3OS	FAM83A.AS1	CA5BP1
ENTPD1.AS1	FAM86B3P	CACTIN.AS1
EPB41L4A.AS1	FAM86C2P	CASK.AS1
FAM83A.AS1	FAM86DP	CATSPER2P1
FAM86B3P	FAM99A	CCT6P3
FAM86C2P	FAM99B	CD27.AS1
FAM86DP	FGD5.AS1	CD99P1
FAM99A		CERS6.AS1
FAM99B		CES1P1
FGD5.AS1		CKMT2.AS1
		CRNDE
		CROCCP2
		CRYBB2P1
		CYP2B7P

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

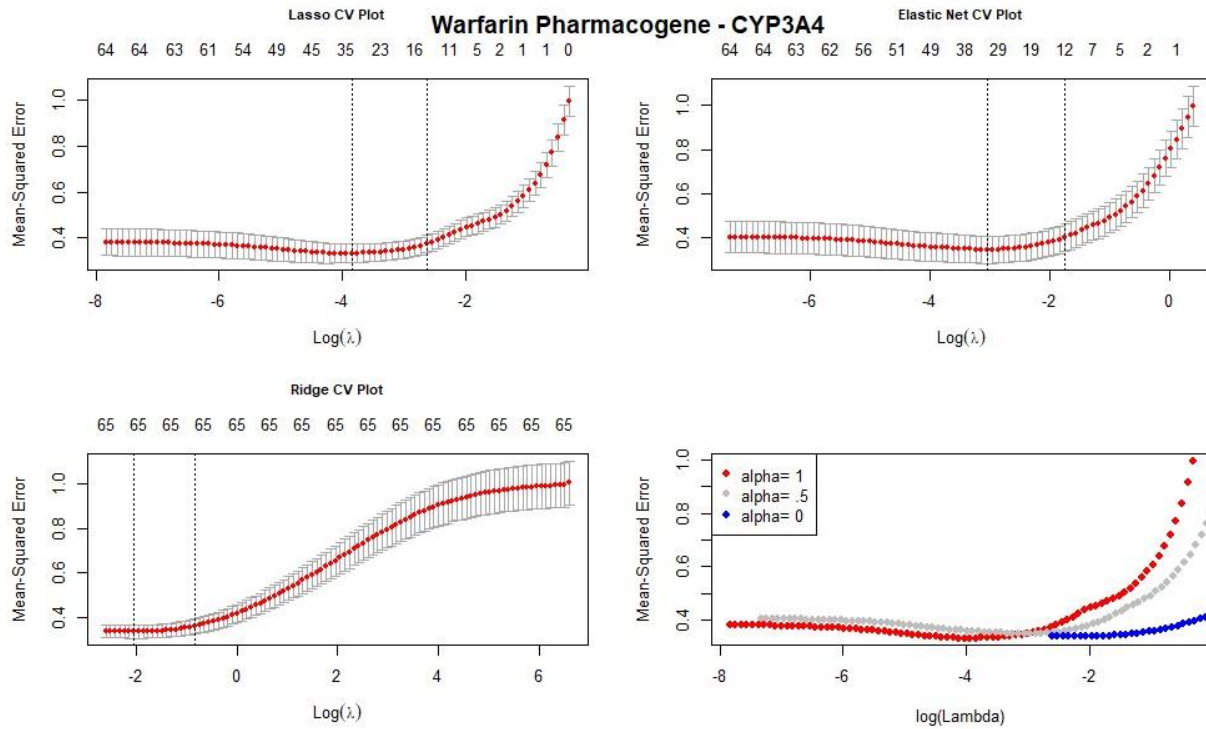


Figure B.14. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP3A4*.

Table B.15. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP4F11* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A1BG.AS1	A1BG.AS1
ABCC6P1	A2M.AS1	A2M.AS1
ADORA2A.AS1	ABCC6P1	ABCC6P1
AGAP2.AS1	ADORA2A.AS1	ABCC6P2
ALOX12P2	AGAP2.AS1	ADORA2A.AS1
ANO1.AS1	ALOX12P2	AGAP2.AS1
AOC4P	ANO1.AS1	ALOX12P2
APOA1.AS	AOC4P	ANO1.AS1
APOC1P1	APOA1.AS	AOC4P
ARAP1.AS1	APOC1P1	APOA1.AS
ARHGAP5.AS1	ARAP1.AS1	APOC1P1
ARSD.AS1	ARHGAP5.AS1	ARAP1.AS1
ASB16.AS1	ARSD.AS1	ARHGAP5.AS1
ASH1L.AS1	ASB16.AS1	ARSD.AS1
AZGP1P1	ASH1L.AS1	ASB16.AS1
BCRP3	AZGP1P1	ASH1L.AS1
BMS1P1	BCRP3	ASMTL.AS1
BTN2A3P	BMS1P1	AZGP1P1
C1RL.AS1	BTN2A3P	BCRP3
C3P1	C1RL.AS1	BDNF.AS
CACTIN.AS1	C3P1	BMS1P1
CASK.AS1	CACTIN.AS1	BTN2A3P
CATSPER2P1	CASK.AS1	C1RL.AS1
CCT6P3	CATSPER2P1	C3P1
CD27.AS1	CCT6P3	CA5BP1
CES1P1	CD27.AS1	CACTIN.AS1
CKMT2.AS1	CES1P1	CASK.AS1
CRYBB2P1	CKMT2.AS1	CATSPER2P1
DANCR	CRYBB2P1	CCT6P3
DBH.AS1	DANCR	CD27.AS1
DENND5B.AS1	DBH.AS1	CD99P1
DHRS4.AS1	DENND5B.AS1	CERS6.AS1
DICER1.AS1	DHRS4.AS1	CES1P1
DLEU2	DICER1.AS1	CKMT2.AS1
DPY19L2P4	DLEU2	CRNDE
EAF1.AS1	DPY19L2P4	CROCCP2
ENTPD1.AS1	EAF1.AS1	CRYBB2P1
ENTPD3.AS1	ENTPD1.AS1	CYP2B7P

EPB41L4A.AS1	ENTPD3.AS1	DANCR
ERVK13.1	EPB41L4A.AS1	DBH.AS1
F11.AS1	ERVK13.1	DENND5B.AS1
FAM13A.AS1	F11.AS1	DHRS4.AS1
FAM157C	FAM13A.AS1	DICER1.AS1
FAM83A.AS1	FAM157C	DIO3OS
FAM86B3P	FAM83A.AS1	DLEU2
FAM86C2P	FAM86B3P	DNAJC9.AS1
FAM86DP	FAM86C2P	DNM3OS
FAM99A	FAM86DP	DPY19L2P4
FAM99B	FAM99A	EAF1.AS1
FGD5.AS1	FAM99B	ENTPD1.AS1
FLNB.AS1	FGD5.AS1	ENTPD3.AS1
	FLNB.AS1	EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

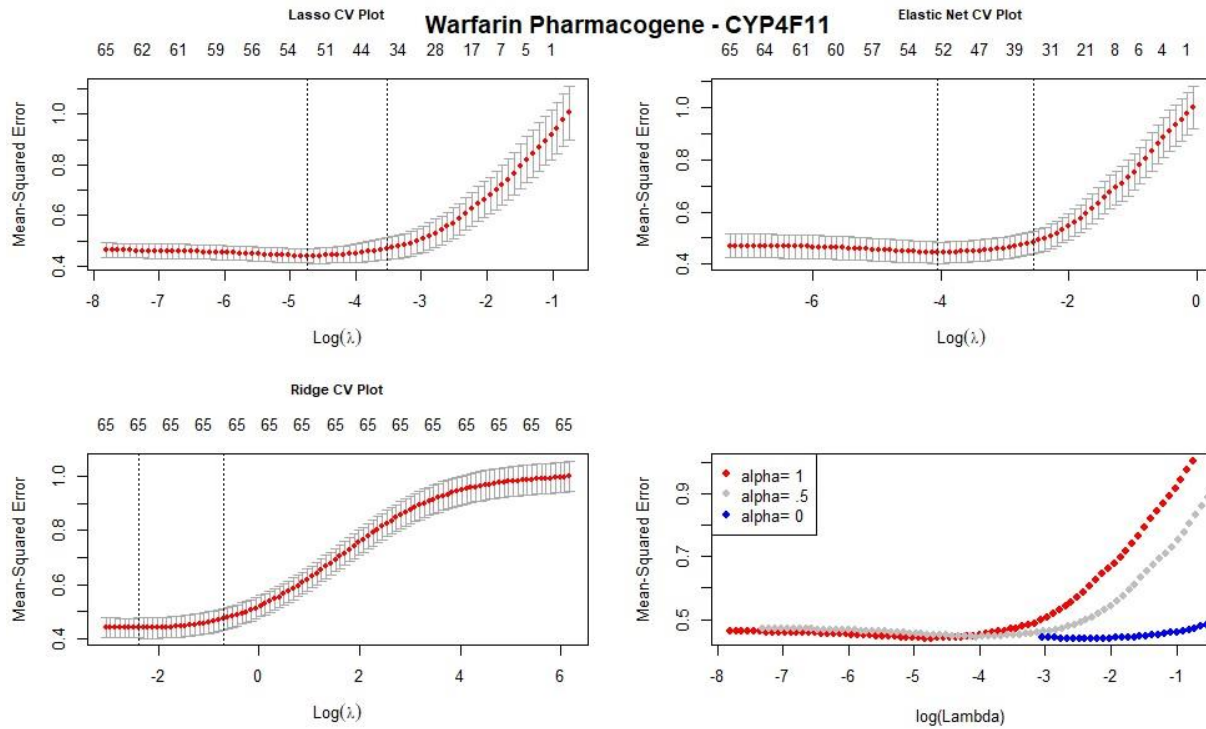


Figure B.15. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP4F11*.

Table B.16. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *CYP4F2* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ABCC6P2	ABCC6P2	ABCC6P2
ADORA2A.AS1	ADORA2A.AS1	ADORA2A.AS1
AGAP2.AS1	AGAP2.AS1	AGAP2.AS1
ALOX12P2	ALOX12P2	ALOX12P2
ANO1.AS1	ANO1.AS1	ANO1.AS1
AOC4P	AOC4P	AOC4P
APOA1.AS	APOA1.AS	APOA1.AS
APOC1P1	APOC1P1	APOC1P1
ARAP1.AS1	ARAP1.AS1	ARAP1.AS1
ARHGAP5.AS1	ARHGAP5.AS1	ARHGAP5.AS1
ARSD.AS1	ARSD.AS1	ARSD.AS1
ASH1L.AS1	ASH1L.AS1	ASB16.AS1
AZGP1P1	AZGP1P1	ASH1L.AS1
BCRP3	BCRP3	ASMTL.AS1
BDNF.AS	BDNF.AS	AZGP1P1
BMS1P1	BMS1P1	BCRP3
C3P1	BTN2A3P	BDNF.AS
CCT6P3	C1RL.AS1	BMS1P1
CES1P1	C3P1	BTN2A3P
CKMT2.AS1	CCT6P3	C1RL.AS1
CRNDE	CD99P1	C3P1
CRYBB2P1	CES1P1	CA5BP1
CYP2B7P	CKMT2.AS1	CACTIN.AS1
DANCR	CRNDE	CASK.AS1
DICER1.AS1	CRYBB2P1	CATSPER2P1
DNM3OS	CYP2B7P	CCT6P3
ENTPD1.AS1	DANCR	CD27.AS1
ENTPD3.AS1	DBH.AS1	CD99P1
ERVK13.1	DICER1.AS1	CERS6.AS1
F11.AS1	DIO3OS	CES1P1
FAM13A.AS1	DNM3OS	CKMT2.AS1
FAM83A.AS1	ENTPD1.AS1	CRNDE
FAM86C2P	ENTPD3.AS1	CROCCP2
FAM86DP	ERVK13.1	CRYBB2P1
FAM99A	F11.AS1	CYP2B7P

FAM13A.AS1	DANCR
FAM83A.AS1	DBH.AS1
FAM86C2P	DENND5B.AS1
FAM86DP	DHRS4.AS1
FAM99A	DICER1.AS1
	DIO3OS
	DLEU2
	DNAJC9.AS1
	DNM3OS
	DPY19L2P4
	EAF1.AS1
	ENTPD1.AS1
	ENTPD3.AS1
	EPB41L4A.AS1
	ERVK13.1
	F11.AS1
	FAM13A.AS1
	FAM157C
	FAM83A.AS1
	FAM86B3P
	FAM86C2P
	FAM86DP
	FAM86EP
	FAM99A
	FAM99B
	FGD5.AS1
	FLNB.AS1

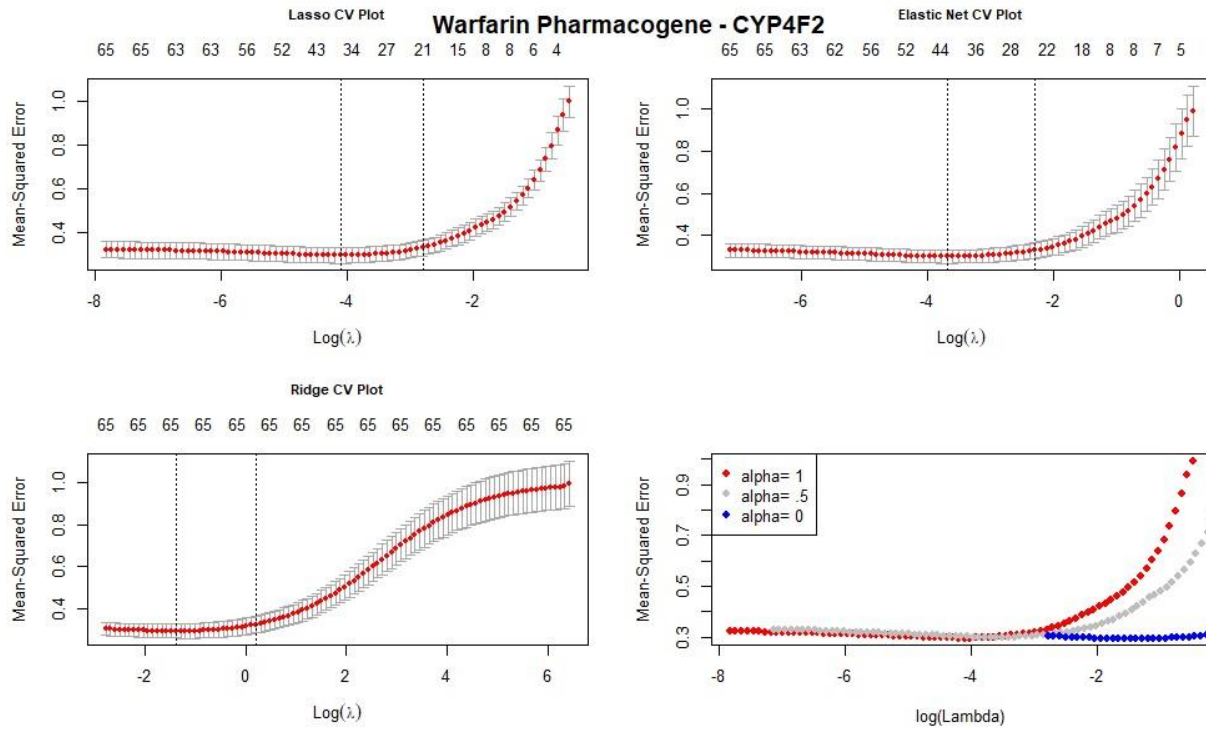


Figure B.16. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *CYP4F2*.

Table B.17. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *DDHD1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	ABCC6P1	A2M.AS1
ABCC6P1	ABCC6P2	ABCC6P1
ABCC6P2	ADORA2A.AS1	ABCC6P2
ADORA2A.AS1	AGAP2.AS1	ADORA2A.AS1
AGAP2.AS1	ALOX12P2	AGAP2.AS1
ALOX12P2	ANO1.AS1	ALOX12P2
ANO1.AS1	AOC4P	ANO1.AS1
AOC4P	APOA1.AS	AOC4P
APOA1.AS	APOC1P1	APOA1.AS
APOC1P1	ARHGAP5.AS1	APOC1P1
ARAP1.AS1	ARSD.AS1	ARAP1.AS1
ARHGAP5.AS1	ASB16.AS1	ARHGAP5.AS1
ARSD.AS1	ASH1L.AS1	ARSD.AS1
ASB16.AS1	AZGP1P1	ASB16.AS1
ASH1L.AS1	BCRP3	ASH1L.AS1
AZGP1P1	BDNF.AS	ASMTL.AS1
BCRP3	BMS1P1	AZGP1P1
BDNF.AS	BTN2A3P	BCRP3
BMS1P1	C1RL.AS1	BDNF.AS
BTN2A3P	CA5BP1	BMS1P1
C1RL.AS1	CACTIN.AS1	BTN2A3P
C3P1	CASK.AS1	C1RL.AS1
CA5BP1	CCT6P3	C3P1
CACTIN.AS1	CD27.AS1	CA5BP1
CASK.AS1	CD99P1	CACTIN.AS1
CATSPER2P1	CERS6.AS1	CASK.AS1
CCT6P3	CRNDE	CATSPER2P1
CD27.AS1	CROCCP2	CCT6P3
CD99P1	DANCR	CD27.AS1
CERS6.AS1	DBH.AS1	CD99P1
CES1P1	DHRS4.AS1	CERS6.AS1
CRNDE	DIO3OS	CES1P1
CROCCP2	DLEU2	CKMT2.AS1
DANCR	DNAJC9.AS1	CRNDE
DBH.AS1	DPY19L2P4	CROCCP2
DHRS4.AS1	ENTPD1.AS1	CRYBB2P1
DICER1.AS1	ENTPD3.AS1	CYP2B7P

DIO3OS	EPB41L4A.AS1	DANCR
DLEU2	F11.AS1	DBH.AS1
DNAJC9.AS1	FAM13A.AS1	DENND5B.AS1
DPY19L2P4	FAM157C	DHRS4.AS1
ENTPD1.AS1	FAM83A.AS1	DICER1.AS1
ENTPD3.AS1	FAM86B3P	DIO3OS
EPB41L4A.AS1	FAM86C2P	DLEU2
F11.AS1	FAM86EP	DNAJC9.AS1
FAM13A.AS1	FAM99B	DNM3OS
FAM157C	FGD5.AS1	DPY19L2P4
FAM83A.AS1	FLNB.AS1	EAF1.AS1
FAM86B3P		ENTPD1.AS1
FAM86C2P		ENTPD3.AS1
FAM86EP		EPB41L4A.AS1
FAM99B		ERVK13.1
FGD5.AS1		F11.AS1
FLNB.AS1		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

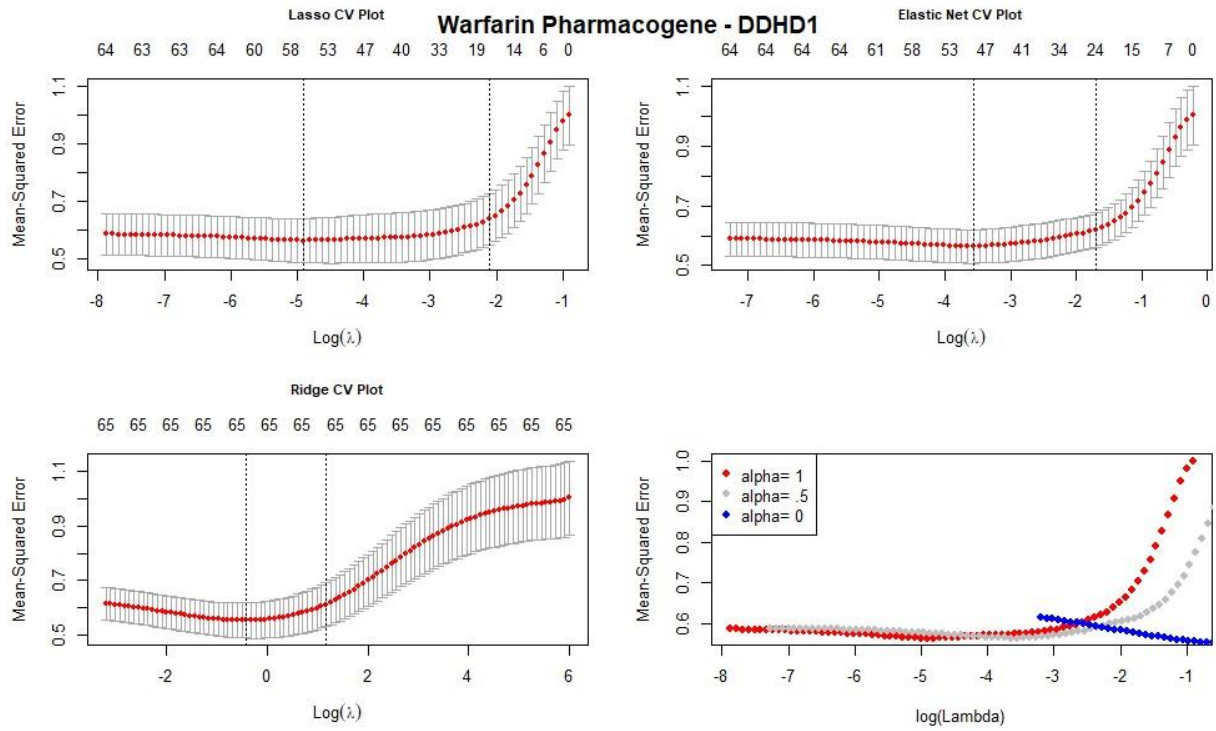


Figure B.17. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *DDHD1*.

Table B.18. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *DNMT3A* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A2M.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
ABCC6P2	ABCC6P2	ABCC6P1
AGAP2.AS1	AGAP2.AS1	ABCC6P2
ALOX12P2	ALOX12P2	ADORA2A.AS1
ANO1.AS1	ANO1.AS1	AGAP2.AS1
AOC4P	AOC4P	ALOX12P2
APOA1.AS	APOA1.AS	ANO1.AS1
APOC1P1	APOC1P1	AOC4P
ARAP1.AS1	ARAP1.AS1	APOA1.AS
ARHGAP5.AS1	ARHGAP5.AS1	APOC1P1
ARSD.AS1	ASB16.AS1	ARAP1.AS1
ASB16.AS1	ASH1L.AS1	ARHGAP5.AS1
ASH1L.AS1	ASMTL.AS1	ARSD.AS1
ASMTL.AS1	AZGP1P1	ASB16.AS1
AZGP1P1	BTN2A3P	ASH1L.AS1
BCRP3	C3P1	ASMTL.AS1
BTN2A3P	CACTIN.AS1	AZGP1P1
C3P1	CASK.AS1	BCRP3
CACTIN.AS1	CD99P1	BDNF.AS
CASK.AS1	CERS6.AS1	BMS1P1
CD99P1	CES1P1	BTN2A3P
CERS6.AS1	CKMT2.AS1	C1RL.AS1
CES1P1	CROCCP2	C3P1
CKMT2.AS1	CYP2B7P	CA5BP1
CROCCP2	DANCR	CACTIN.AS1
CYP2B7P	DBH.AS1	CASK.AS1
DANCR	DENND5B.AS1	CATSPER2P1
DBH.AS1	DHRS4.AS1	CCT6P3
DENND5B.AS1	DICER1.AS1	CD27.AS1
DHRS4.AS1	DLEU2	CD99P1
DICER1.AS1	DNAJC9.AS1	CERS6.AS1
DLEU2	DNM3OS	CES1P1
DNAJC9.AS1	DPY19L2P4	CKMT2.AS1
DNM3OS	EAF1.AS1	CRNDE
DPY19L2P4	ENTPD1.AS1	CROCCP2
EAF1.AS1	ENTPD3.AS1	CRYBB2P1
ENTPD1.AS1	F11.AS1	CYP2B7P

ENTPD3.AS1	FAM13A.AS1	DANCR
F11.AS1	FAM157C	DBH.AS1
FAM13A.AS1	FAM83A.AS1	DENND5B.AS1
FAM157C	FAM86B3P	DHRS4.AS1
FAM83A.AS1	FAM86C2P	DICER1.AS1
FAM86B3P	FAM86EP	DIO3OS
FAM86C2P	FLNB.AS1	DLEU2
FAM86EP		DNAJC9.AS1
FAM99A		DNM3OS
FGD5.AS1		DPY19L2P4
FLNB.AS1		EAF1.AS1
		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

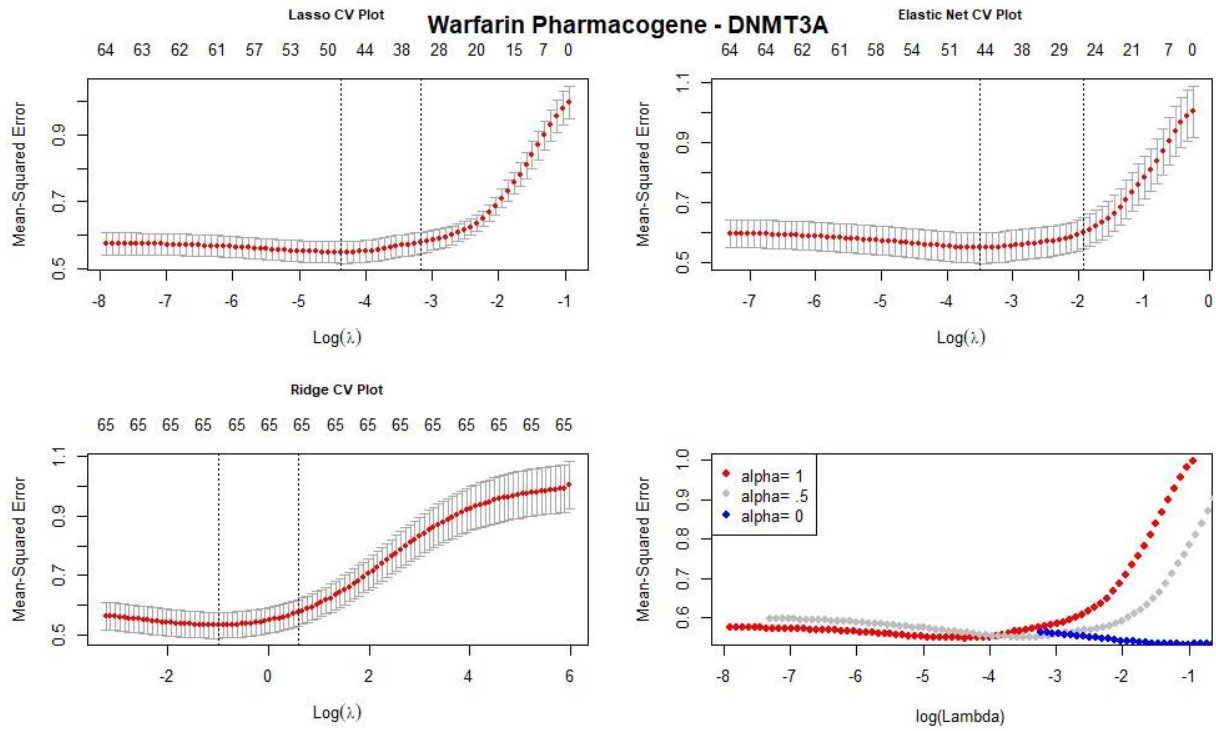


Figure B.18. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *DNMT3A*.

Table B.19. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *EPHX1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
ABCC6P1	A1BG.AS1	A1BG.AS1
ABCC6P2	ABCC6P1	A2M.AS1
AGAP2.AS1	ABCC6P2	ABCC6P1
ANO1.AS1	AGAP2.AS1	ABCC6P2
AOC4P	ANO1.AS1	ADORA2A.AS1
APOA1.AS	AOC4P	AGAP2.AS1
APOC1P1	APOA1.AS	ALOX12P2
ARAP1.AS1	APOC1P1	ANO1.AS1
ARHGAP5.AS1	ARAP1.AS1	AOC4P
ARSD.AS1	ARHGAP5.AS1	APOA1.AS
ASB16.AS1	ARSD.AS1	APOC1P1
ASH1L.AS1	ASB16.AS1	ARAP1.AS1
ASMTL.AS1	ASH1L.AS1	ARHGAP5.AS1
BCRP3	ASMTL.AS1	ARSD.AS1
BDNF.AS	AZGP1P1	ASB16.AS1
BMS1P1	BCRP3	ASH1L.AS1
C1RL.AS1	BDNF.AS	ASMTL.AS1
C3P1	BMS1P1	AZGP1P1
CA5BP1	C1RL.AS1	BCRP3
CACTIN.AS1	C3P1	BDNF.AS
CASK.AS1	CA5BP1	BMS1P1
CATSPER2P1	CACTIN.AS1	BTN2A3P
CCT6P3	CASK.AS1	C1RL.AS1
CD27.AS1	CATSPER2P1	C3P1
CD99P1	CCT6P3	CA5BP1
CES1P1	CD27.AS1	CACTIN.AS1
CRNDE	CD99P1	CASK.AS1
CROCCP2	CERS6.AS1	CATSPER2P1
CYP2B7P	CES1P1	CCT6P3
DANCR	CRNDE	CD27.AS1
DENND5B.AS1	CROCCP2	CD99P1
DICER1.AS1	CYP2B7P	CERS6.AS1
DIO3OS	DANCR	CES1P1
DLEU2	DENND5B.AS1	CKMT2.AS1
DNAJC9.AS1	DICER1.AS1	CRNDE
DPY19L2P4	DIO3OS	CROCCP2
EAF1.AS1	DLEU2	CRYBB2P1
ENTPD1.AS1	DNAJC9.AS1	CYP2B7P

ENTPD3.AS1	DPY19L2P4	DANCR
F11.AS1	EAF1.AS1	DBH.AS1
FAM83A.AS1	ENTPD1.AS1	DENND5B.AS1
FAM86B3P	ENTPD3.AS1	DHRS4.AS1
FAM86DP	F11.AS1	DICER1.AS1
FAM99A	FAM83A.AS1	DIO3OS
FAM99B	FAM86B3P	DLEU2
FLNB.AS1	FAM86DP	DNAJC9.AS1
	FAM86EP	DNM3OS
	FAM99A	DPY19L2P4
	FAM99B	EAF1.AS1
	FLNB.AS1	ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

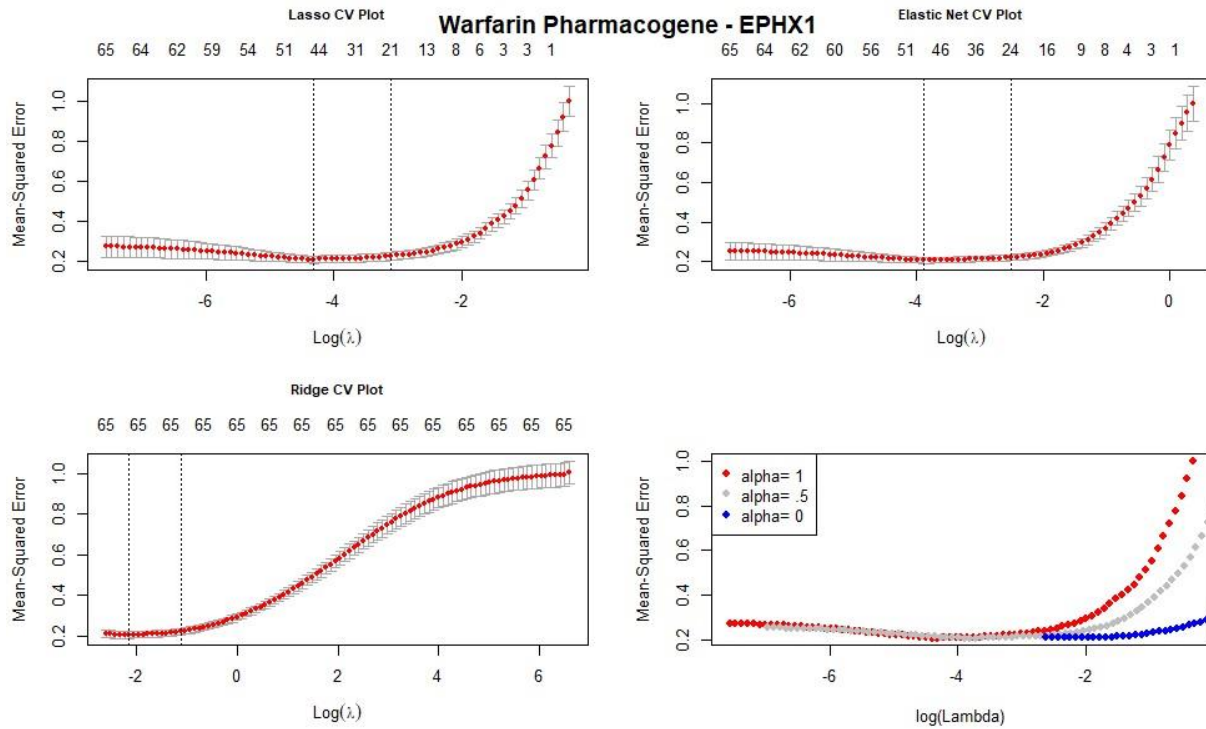


Figure B.19. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *EPHX1*.

Table B.20. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *VKORC1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P2	ABCC6P2	ABCC6P1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P2
AGAP2.AS1	AGAP2.AS1	ADORA2A.AS1
ANO1.AS1	ANO1.AS1	AGAP2.AS1
AOC4P	AOC4P	ALOX12P2
APOA1.AS	APOA1.AS	ANO1.AS1
APOC1P1	APOC1P1	AOC4P
ARAP1.AS1	ARAP1.AS1	APOA1.AS
ARHGAP5.AS1	ARHGAP5.AS1	APOC1P1
ARSD.AS1	ARSD.AS1	ARAP1.AS1
ASB16.AS1	ASB16.AS1	ARHGAP5.AS1
ASH1L.AS1	ASH1L.AS1	ARSD.AS1
ASMTL.AS1	ASMTL.AS1	ASB16.AS1
AZGP1P1	AZGP1P1	ASH1L.AS1
BCRP3	BCRP3	ASMTL.AS1
BTN2A3P	BTN2A3P	AZGP1P1
C1RL.AS1	C1RL.AS1	BCRP3
C3P1	C3P1	BDNF.AS
CA5BP1	CA5BP1	BMS1P1
CACTIN.AS1	CACTIN.AS1	BTN2A3P
CCT6P3	CCT6P3	C1RL.AS1
CD99P1	CD99P1	C3P1
CERS6.AS1	CERS6.AS1	CA5BP1
CES1P1	CES1P1	CACTIN.AS1
CKMT2.AS1	CKMT2.AS1	CASK.AS1
CRNDE	CRNDE	CATSPER2P1
CRYBB2P1	CRYBB2P1	CCT6P3
CYP2B7P	CYP2B7P	CD27.AS1
DANCR	DANCR	CD99P1
DBH.AS1	DBH.AS1	CERS6.AS1
DENND5B.AS1	DENND5B.AS1	CES1P1
DICER1.AS1	DICER1.AS1	CKMT2.AS1
DLEU2	DLEU2	CRNDE
DNAJC9.AS1	DNAJC9.AS1	CROCCP2
DNM3OS	DNM3OS	CRYBB2P1
DPY19L2P4	DPY19L2P4	CYP2B7P

ENTPD1.AS1	EAF1.AS1	DANCR
ENTPD3.AS1	ENTPD1.AS1	DBH.AS1
EPB41L4A.AS1	ENTPD3.AS1	DENND5B.AS1
ERVK13.1	EPB41L4A.AS1	DHRS4.AS1
F11.AS1	ERVK13.1	DICER1.AS1
FAM13A.AS1	F11.AS1	DIO3OS
FAM157C	FAM13A.AS1	DLEU2
FAM83A.AS1	FAM157C	DNAJC9.AS1
FAM86C2P	FAM83A.AS1	DNM3OS
FAM86DP	FAM86C2P	DPY19L2P4
FAM86EP	FAM86DP	EAF1.AS1
FAM99A	FAM86EP	ENTPD1.AS1
FAM99B	FAM99A	ENTPD3.AS1
FGD5.AS1	FAM99B	EPB41L4A.AS1
FLNB.AS1	FGD5.AS1	ERVK13.1
	FLNB.AS1	F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

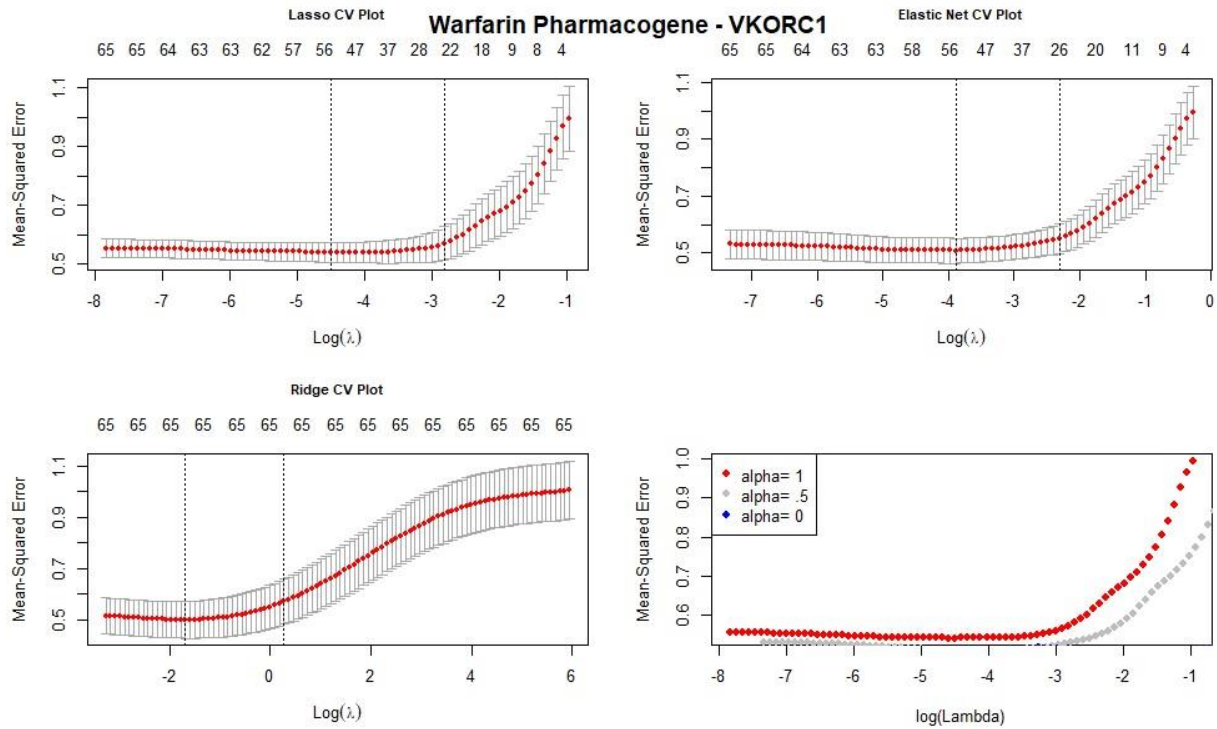


Figure B.20. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *VKORC1*.

Table B.21. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *FII* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
AGAP2.AS1	ABCC6P2	ABCC6P1
APOA1.AS	AGAP2.AS1	ABCC6P2
ARAP1.AS1	APOA1.AS	ADORA2A.AS1
ARHGAP5.AS1	ARAP1.AS1	AGAP2.AS1
ARSD.AS1	ARHGAP5.AS1	ALOX12P2
ASB16.AS1	ARSD.AS1	ANO1.AS1
ASH1L.AS1	ASB16.AS1	AOC4P
BDNF.AS	ASH1L.AS1	APOA1.AS
C3P1	BDNF.AS	APOC1P1
CERS6.AS1	C3P1	ARAP1.AS1
CES1P1	CD99P1	ARHGAP5.AS1
CRNDE	CERS6.AS1	ARSD.AS1
CROCCP2	CES1P1	ASB16.AS1
CRYBB2P1	CRNDE	ASH1L.AS1
CYP2B7P	CROCCP2	ASMTL.AS1
DBH.AS1	CRYBB2P1	AZGP1P1
DENND5B.AS1	CYP2B7P	BCRP3
DICER1.AS1	DBH.AS1	BDNF.AS
DNAJC9.AS1	DENND5B.AS1	BMS1P1
DNM3OS	DICER1.AS1	BTN2A3P
DPY19L2P4	DNAJC9.AS1	C1RL.AS1
EAF1.AS1	DNM3OS	C3P1
ENTPD1.AS1	DPY19L2P4	CA5BP1
EPB41L4A.AS1	EAF1.AS1	CACTIN.AS1
F11.AS1	ENTPD1.AS1	CASK.AS1
FAM13A.AS1	ENTPD3.AS1	CATSPER2P1
FAM157C	EPB41L4A.AS1	CCT6P3
FLNB.AS1	F11.AS1	CD27.AS1
	FAM13A.AS1	CD99P1
	FAM157C	CERS6.AS1
	FAM86EP	CES1P1
	FLNB.AS1	CKMT2.AS1
		CRNDE
		CROCCP2
		CRYBB2P1
		CYP2B7P

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

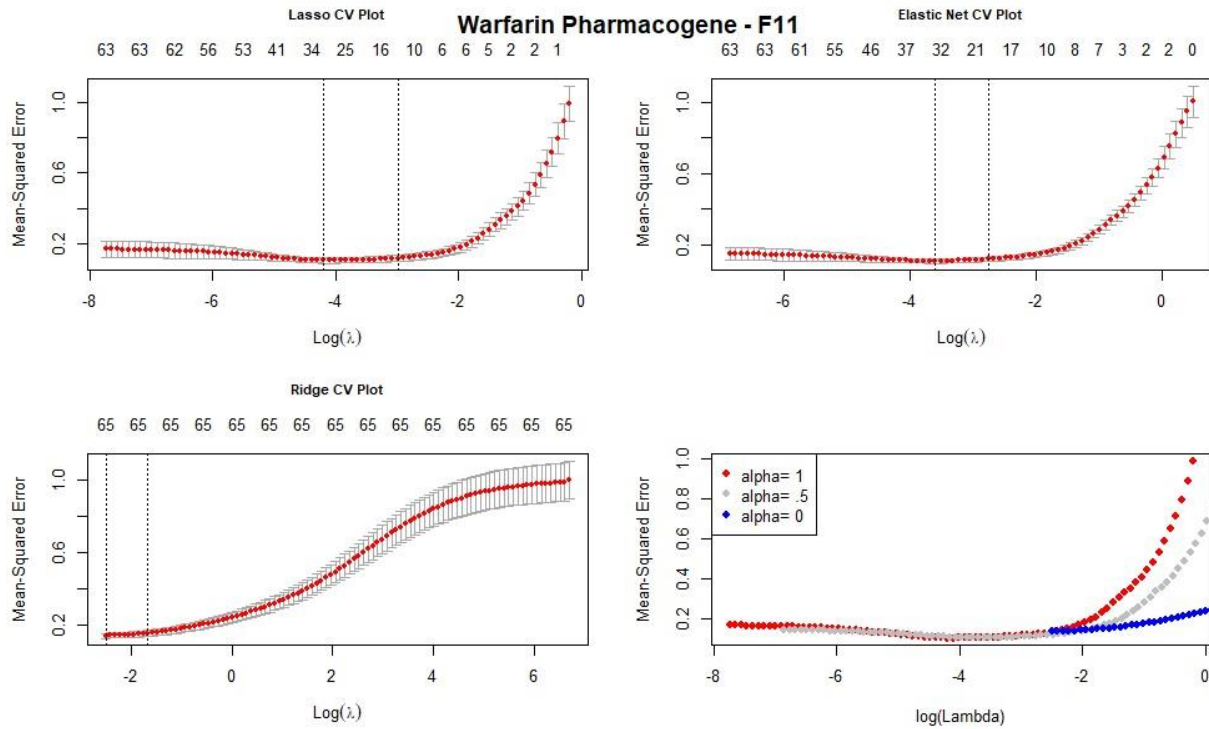


Figure B.21. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F11*.

Table B.22. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F13A1* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P2	ABCC6P2	ABCC6P1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P2
AGAP2.AS1	AGAP2.AS1	ADORA2A.AS1
ALOX12P2	ALOX12P2	AGAP2.AS1
APOC1P1	APOC1P1	ALOX12P2
ARAP1.AS1	ARAP1.AS1	ANO1.AS1
ARHGAP5.AS1	ARHGAP5.AS1	AOC4P
ASB16.AS1	ARSD.AS1	APOA1.AS
ASMTL.AS1	ASB16.AS1	APOC1P1
AZGP1P1	ASMTL.AS1	ARAP1.AS1
BDNF.AS	AZGP1P1	ARHGAP5.AS1
BMS1P1	BDNF.AS	ARSD.AS1
CASK.AS1	BMS1P1	ASB16.AS1
CATSPER2P1	CASK.AS1	ASH1L.AS1
CD27.AS1	CATSPER2P1	ASMTL.AS1
CERS6.AS1	CCT6P3	AZGP1P1
CES1P1	CD27.AS1	BCRP3
CKMT2.AS1	CERS6.AS1	BDNF.AS
CROCCP2	CES1P1	BMS1P1
CRYBB2P1	CKMT2.AS1	BTN2A3P
CYP2B7P	CROCCP2	C1RL.AS1
DENND5B.AS1	CRYBB2P1	C3P1
DHRS4.AS1	CYP2B7P	CA5BP1
DIO3OS	DENND5B.AS1	CACTIN.AS1
DNM3OS	DHRS4.AS1	CASK.AS1
EAF1.AS1	DIO3OS	CATSPER2P1
ENTPD1.AS1	DNM3OS	CCT6P3
ENTPD3.AS1	EAF1.AS1	CD27.AS1
EPB41L4A.AS1	ENTPD1.AS1	CD99P1
ERVK13.1	ENTPD3.AS1	CERS6.AS1
FAM13A.AS1	EPB41L4A.AS1	CES1P1
FAM157C	ERVK13.1	CKMT2.AS1
FAM83A.AS1	FAM13A.AS1	CRNDE
FAM86EP	FAM157C	CROCCP2
FAM99A	FAM83A.AS1	CRYBB2P1
FAM99B	FAM86EP	CYP2B7P

FGD5.AS1	FAM99A	DANCR
FLNB.AS1	FAM99B	DBH.AS1
	FGD5.AS1	DENND5B.AS1
	FLNB.AS1	DHRS4.AS1
		DICER1.AS1
		DIO3OS
		DLEU2
		DNAJC9.AS1
		DNM3OS
		DPY19L2P4
		EAF1.AS1
		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

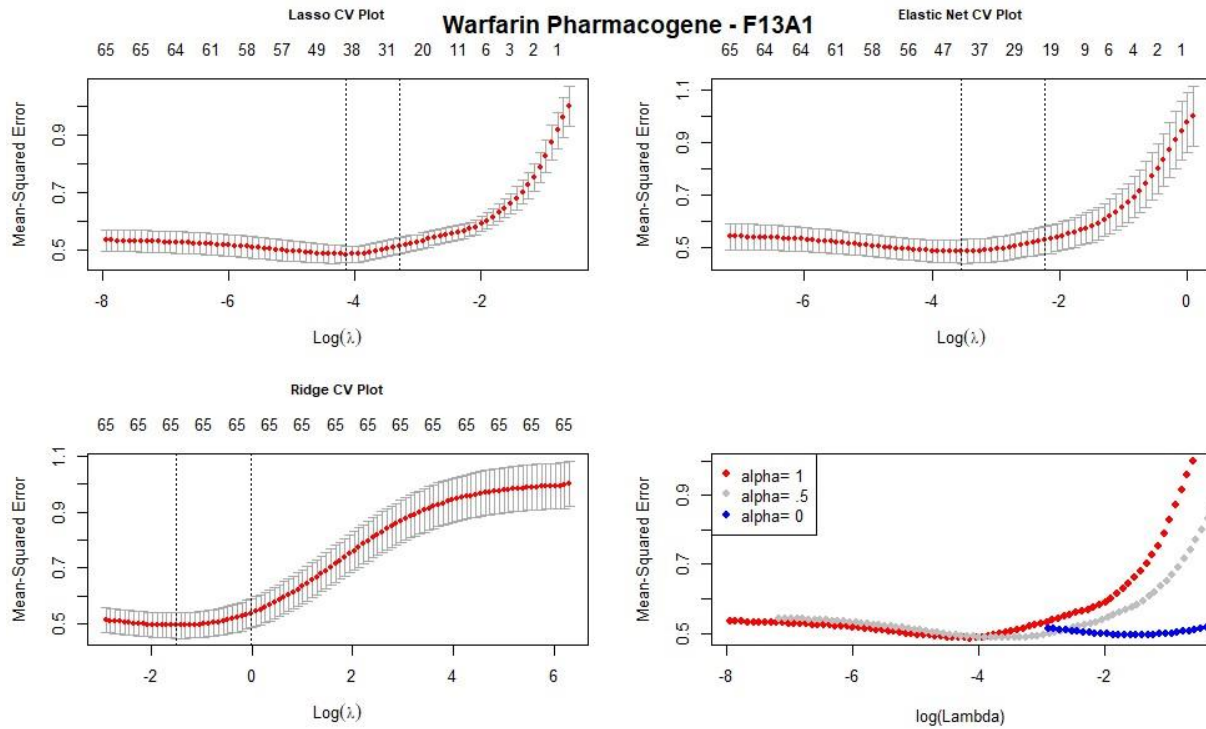


Figure B.22. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F13A1*.

Table B.23. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F2* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
ANO1.AS1	ANO1.AS1	ABCC6P1
AOC4P	AOC4P	ABCC6P2
APOC1P1	APOC1P1	ADORA2A.AS1
ARAP1.AS1	ARAP1.AS1	AGAP2.AS1
ARHGAP5.AS1	ARHGAP5.AS1	ALOX12P2
ARSD.AS1	ARSD.AS1	ANO1.AS1
ASH1L.AS1	ASH1L.AS1	AOC4P
AZGP1P1	AZGP1P1	APOA1.AS
BCRP3	BCRP3	APOC1P1
BDNF.AS	BDNF.AS	ARAP1.AS1
BMS1P1	BMS1P1	ARHGAP5.AS1
BTN2A3P	BTN2A3P	ARSD.AS1
C3P1	C3P1	ASB16.AS1
CA5BP1	CA5BP1	ASH1L.AS1
CACTIN.AS1	CACTIN.AS1	ASMTL.AS1
CASK.AS1	CASK.AS1	AZGP1P1
CATSPER2P1	CATSPER2P1	BCRP3
CCT6P3	CCT6P3	BDNF.AS
CD27.AS1	CD27.AS1	BMS1P1
CERS6.AS1	CERS6.AS1	BTN2A3P
CES1P1	CES1P1	C1RL.AS1
CRNDE	CRNDE	C3P1
DBH.AS1	DBH.AS1	CA5BP1
DICER1.AS1	DICER1.AS1	CACTIN.AS1
DLEU2	DLEU2	CASK.AS1
DNAJC9.AS1	DNAJC9.AS1	CATSPER2P1
DNM3OS	DNM3OS	CCT6P3
EAF1.AS1	EAF1.AS1	CD27.AS1
ENTPD1.AS1	ENTPD1.AS1	CD99P1
ENTPD3.AS1	ENTPD3.AS1	CERS6.AS1
ERVK13.1	FAM157C	CES1P1
FAM157C	FAM83A.AS1	CKMT2.AS1
FAM83A.AS1	FAM86B3P	CRNDE
FAM86B3P	FAM86EP	CROCCP2
FAM86DP	FAM99A	CRYBB2P1
FAM86EP	FAM99B	CYP2B7P

FAM99A	FLNB.AS1	DANCR
FAM99B		DBH.AS1
FLNB.AS1		DENND5B.AS1
		DHRS4.AS1
		DICER1.AS1
		DIO3OS
		DLEU2
		DNAJC9.AS1
		DNM3OS
		DPY19L2P4
		EAF1.AS1
		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

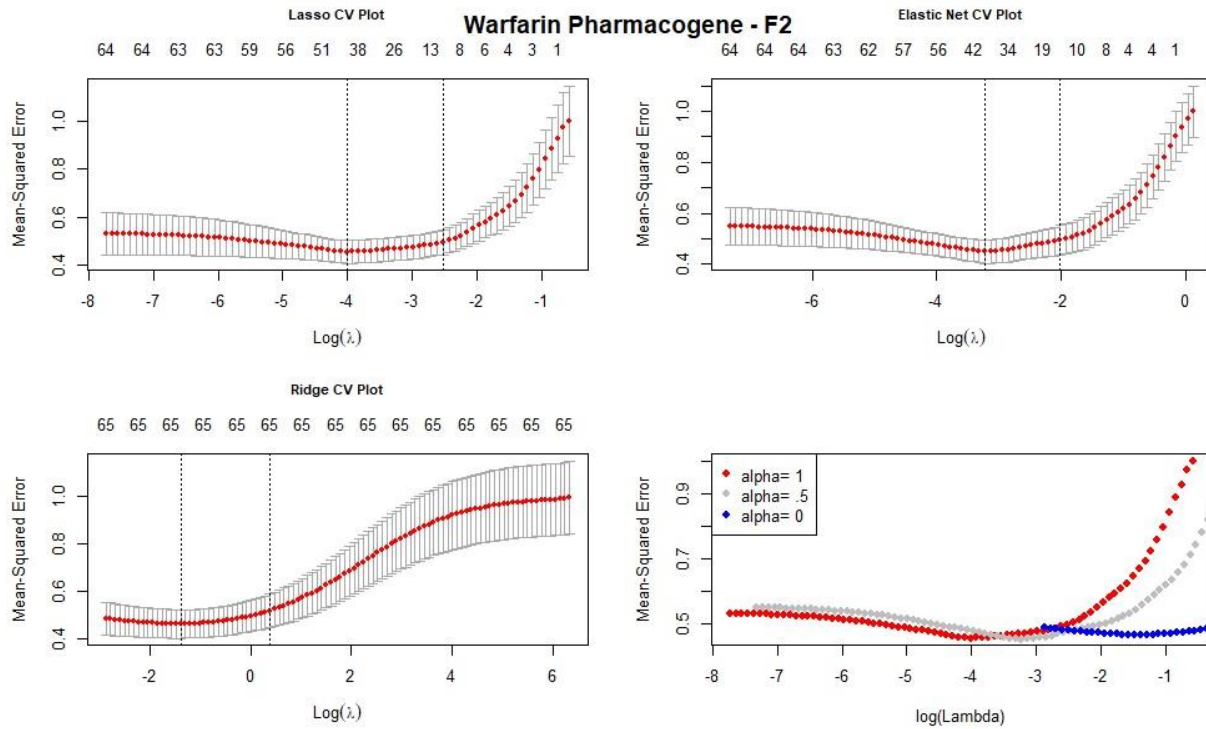


Figure B.23. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F2*.

Table B.24. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F5* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A2M.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P1
ALOX12P2	ALOX12P2	ABCC6P2
ANO1.AS1	ANO1.AS1	ADORA2A.AS1
APOA1.AS	APOA1.AS	AGAP2.AS1
APOC1P1	APOC1P1	ALOX12P2
ARSD.AS1	ARSD.AS1	ANO1.AS1
ASB16.AS1	ASB16.AS1	AOC4P
ASH1L.AS1	ASH1L.AS1	APOA1.AS
AZGP1P1	AZGP1P1	APOC1P1
BCRP3	BCRP3	ARAP1.AS1
BMS1P1	BTN2A3P	ARHGAP5.AS1
BTN2A3P	CA5BP1	ARSD.AS1
CA5BP1	CACTIN.AS1	ASB16.AS1
CACTIN.AS1	CATSPER2P1	ASH1L.AS1
CATSPER2P1	CCT6P3	ASMTL.AS1
CCT6P3	CD27.AS1	AZGP1P1
CD27.AS1	CERS6.AS1	BCRP3
CERS6.AS1	CES1P1	BDNF.AS
CES1P1	CKMT2.AS1	BMS1P1
CKMT2.AS1	CRNDE	BTN2A3P
CRNDE	CROCCP2	C1RL.AS1
CROCCP2	CYP2B7P	C3P1
CYP2B7P	DENND5B.AS1	CA5BP1
DENND5B.AS1	DICER1.AS1	CACTIN.AS1
DICER1.AS1	DNAJC9.AS1	CASK.AS1
DIO3OS	DPY19L2P4	CATSPER2P1
DNAJC9.AS1	EAF1.AS1	CCT6P3
DPY19L2P4	ENTPD1.AS1	CD27.AS1
EAF1.AS1	ENTPD3.AS1	CD99P1
ENTPD1.AS1	EPB41L4A.AS1	CERS6.AS1
ENTPD3.AS1	ERVK13.1	CES1P1
EPB41L4A.AS1	FAM13A.AS1	CKMT2.AS1
ERVK13.1	FAM157C	CRNDE
FAM13A.AS1	FAM83A.AS1	CROCCP2
FAM157C	FAM86B3P	CRYBB2P1
FAM83A.AS1	FAM86DP	CYP2B7P

FAM86B3P	FAM99A	DANCR
FAM86DP	FAM99B	DBH.AS1
FAM99A	FGD5.AS1	DENND5B.AS1
FAM99B	FLNB.AS1	DHRS4.AS1
FGD5.AS1		DICER1.AS1
FLNB.AS1		DIO3OS
		DLEU2
		DNAJC9.AS1
		DNM3OS
		DPY19L2P4
		EAF1.AS1
		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

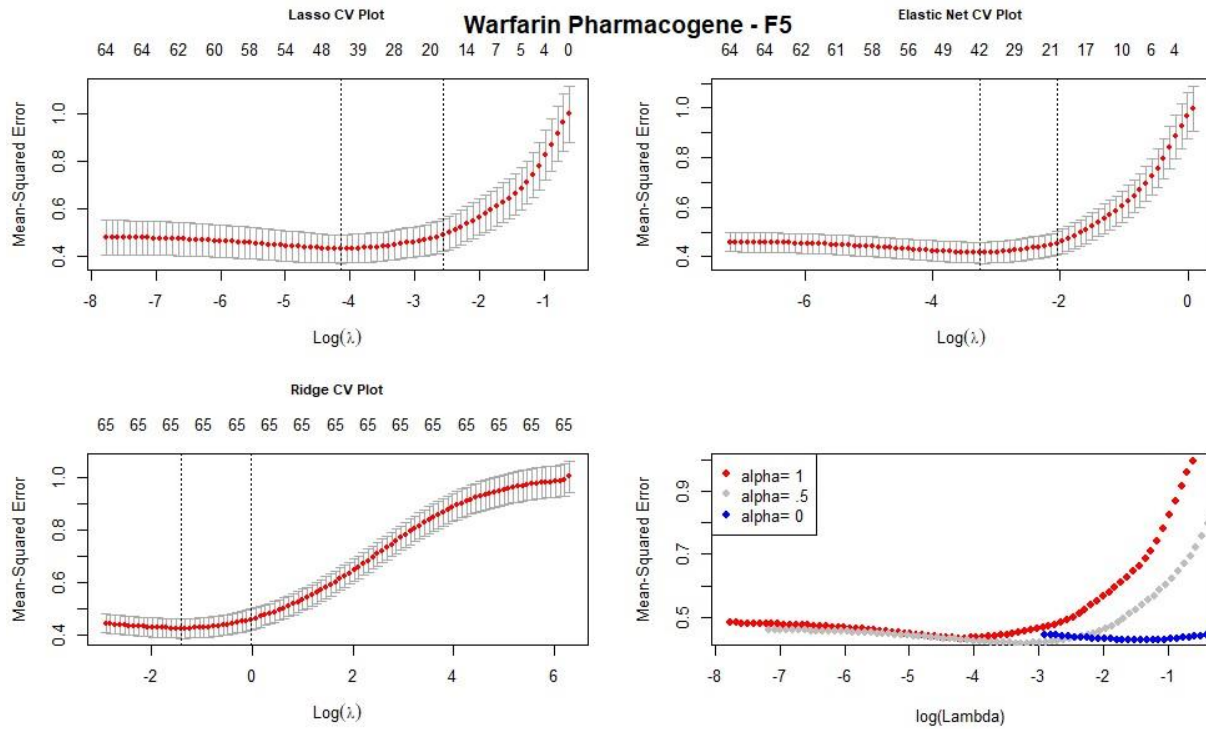


Figure B.24. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F5*.

Table B.25. LncRNA genes selected as predictors to explain the expression of the warfarin pharmacogene *F7* in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
ABCC6P1	ABCC6P1	A1BG.AS1
ABCC6P2	ABCC6P2	A2M.AS1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P1
AGAP2.AS1	AGAP2.AS1	ABCC6P2
ANO1.AS1	ANO1.AS1	ADORA2A.AS1
AOC4P	AOC4P	AGAP2.AS1
APOA1.AS	APOA1.AS	ALOX12P2
ARAP1.AS1	ARAP1.AS1	ANO1.AS1
ARSD.AS1	ARSD.AS1	AOC4P
ASB16.AS1	ASB16.AS1	APOA1.AS
ASH1L.AS1	ASH1L.AS1	APOC1P1
ASMTL.AS1	ASMTL.AS1	ARAP1.AS1
AZGP1P1	AZGP1P1	ARHGAP5.AS1
BCRP3	BCRP3	ARSD.AS1
CACTIN.AS1	CACTIN.AS1	ASB16.AS1
CATSPER2P1	CATSPER2P1	ASH1L.AS1
CCT6P3	CCT6P3	ASMTL.AS1
CD27.AS1	CD27.AS1	AZGP1P1
CD99P1	CD99P1	BCRP3
CES1P1	CES1P1	BDNF.AS
CKMT2.AS1	CKMT2.AS1	BMS1P1
CRNDE	CRNDE	BTN2A3P
CROCCP2	CROCCP2	C1RL.AS1
CRYBB2P1	CRYBB2P1	C3P1
CYP2B7P	CYP2B7P	CA5BP1
DANCR	DANCR	CACTIN.AS1
DIO3OS	DIO3OS	CASK.AS1
DNAJC9.AS1	DNAJC9.AS1	CATSPER2P1
DNM3OS	DNM3OS	CCT6P3
ENTPD3.AS1	ENTPD3.AS1	CD27.AS1
F11.AS1	F11.AS1	CD99P1
FAM13A.AS1	FAM13A.AS1	CERS6.AS1
FAM157C	FAM157C	CES1P1
FAM86C2P	FAM86C2P	CKMT2.AS1
FAM86DP	FAM86DP	CRNDE
FAM99B	FAM99B	CROCCP2
FGD5.AS1	FGD5.AS1	CRYBB2P1
		CYP2B7P

DANCR
DBH.AS1
DENND5B.AS1
DHRS4.AS1
DICER1.AS1
DIO3OS
DLEU2
DNAJC9.AS1
DNM3OS
DPY19L2P4
EAF1.AS1
ENTPD1.AS1
ENTPD3.AS1
EPB41L4A.AS1
ERVK13.1
F11.AS1
FAM13A.AS1
FAM157C
FAM83A.AS1
FAM86B3P
FAM86C2P
FAM86DP
FAM86EP
FAM99A
FAM99B
FGD5.AS1
FLNB.AS1

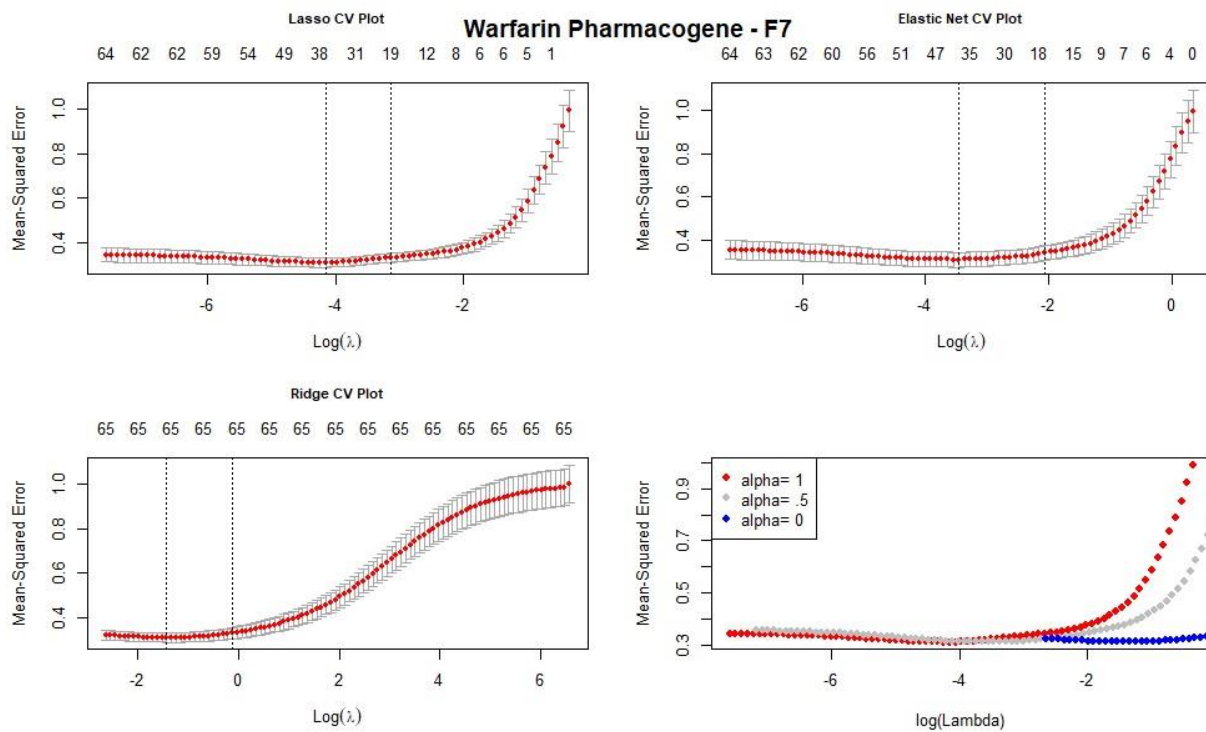


Figure B.25. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against warfarin pharmacogene *F7*.

APPENDIX C

Table C.1. LncRNA genes selected as predictors to explain the expression of all 25 warfarin pharmacogenes as one group in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ABCC6P2	ABCC6P2	ABCC6P2
ADORA2A.AS1	ADORA2A.AS1	ADORA2A.AS1
AGAP2.AS1	AGAP2.AS1	AGAP2.AS1
ALOX12P2	ALOX12P2	ALOX12P2
ANO1.AS1	ANO1.AS1	ANO1.AS1
AOC4P	AOC4P	AOC4P
APOA1.AS	APOA1.AS	APOA1.AS
APOC1P1	APOC1P1	APOC1P1
ARAP1.AS1	ARAP1.AS1	ARAP1.AS1
ARHGAP5.AS1	ARHGAP5.AS1	ARHGAP5.AS1
ARSD.AS1	ARSD.AS1	ARSD.AS1
ASB16.AS1	ASB16.AS1	ASB16.AS1
ASH1L.AS1	ASH1L.AS1	ASH1L.AS1
ASMTL.AS1	ASMTL.AS1	ASMTL.AS1
AZGP1P1	AZGP1P1	AZGP1P1
BCRP3	BCRP3	BCRP3
BDNF.AS	BDNF.AS	BDNF.AS
BMS1P1	BMS1P1	BMS1P1
BTN2A3P	BTN2A3P	BTN2A3P
C1RL.AS1	C1RL.AS1	C1RL.AS1
C3P1	C3P1	C3P1
CA5BP1	CA5BP1	CA5BP1
CACTIN.AS1	CACTIN.AS1	CACTIN.AS1
CASK.AS1	CASK.AS1	CASK.AS1
CATSPER2P1	CATSPER2P1	CATSPER2P1
CCT6P3	CCT6P3	CCT6P3
CD27.AS1	CD27.AS1	CD27.AS1
CD99P1	CD99P1	CD99P1
CERS6.AS1	CERS6.AS1	CERS6.AS1
CES1P1	CES1P1	CES1P1
CKMT2.AS1	CKMT2.AS1	CKMT2.AS1
CRNDE	CRNDE	CRNDE
CROCCP2	CROCCP2	CROCCP2

CRYBB2P1	CRYBB2P1	CRYBB2P1
CYP2B7P	CYP2B7P	CYP2B7P
DANCR	DANCR	DANCR
DBH.AS1	DBH.AS1	DBH.AS1
DENND5B.AS1	DENND5B.AS1	DENND5B.AS1
DHRS4.AS1	DHRS4.AS1	DHRS4.AS1
DICER1.AS1	DICER1.AS1	DICER1.AS1
DIO3OS	DIO3OS	DIO3OS
DLEU2	DLEU2	DLEU2
DNAJC9.AS1	DNAJC9.AS1	DNAJC9.AS1
DNM3OS	DNM3OS	DNM3OS
DPY19L2P4	DPY19L2P4	DPY19L2P4
EAF1.AS1	EAF1.AS1	EAF1.AS1
ENTPD1.AS1	ENTPD1.AS1	ENTPD1.AS1
ENTPD3.AS1	ENTPD3.AS1	ENTPD3.AS1
EPB41L4A.AS1	EPB41L4A.AS1	EPB41L4A.AS1
ERVK13.1	ERVK13.1	ERVK13.1
F11.AS1	F11.AS1	F11.AS1
FAM13A.AS1	FAM13A.AS1	FAM13A.AS1
FAM157C	FAM157C	FAM157C
FAM83A.AS1	FAM83A.AS1	FAM83A.AS1
FAM86B3P	FAM86B3P	FAM86B3P
FAM86C2P	FAM86C2P	FAM86C2P
FAM86DP	FAM86DP	FAM86DP
FAM86EP	FAM86EP	FAM86EP
FAM99A	FAM99A	FAM99A
FAM99B	FAM99B	FAM99B
FGD5.AS1	FGD5.AS1	FGD5.AS1
FLNB.AS1	FLNB.AS1	FLNB.AS1

Table C.2. LncRNA genes selected as predictors to explain the expression of 12 warfarin pharmacogenes (*APOC1*, *APOE*, *CYP1A2*, *CYP2A6*, *CYP2C19*, *CYP2C9*, *CYP3A4*, *CYP4F2*, *EPHX1*, *F5*, *F7*, and *F11*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ABCC6P2	ABCC6P2	ABCC6P2
ADORA2A.AS1	ADORA2A.AS1	ADORA2A.AS1
AGAP2.AS1	AGAP2.AS1	AGAP2.AS1
ALOX12P2	ALOX12P2	ALOX12P2
ANO1.AS1	ANO1.AS1	ANO1.AS1
AOC4P	AOC4P	AOC4P
APOA1.AS	APOA1.AS	APOA1.AS
APOC1P1	APOC1P1	APOC1P1
ARAP1.AS1	ARAP1.AS1	ARAP1.AS1
ARHGAP5.AS1	ARHGAP5.AS1	ARHGAP5.AS1
ARSD.AS1	ARSD.AS1	ARSD.AS1
ASB16.AS1	ASB16.AS1	ASB16.AS1
ASH1L.AS1	ASH1L.AS1	ASH1L.AS1
ASMTL.AS1	ASMTL.AS1	ASMTL.AS1
AZGP1P1	AZGP1P1	AZGP1P1
BCRP3	BCRP3	BCRP3
BDNF.AS	BDNF.AS	BDNF.AS
BMS1P1	BMS1P1	BMS1P1
BTN2A3P	BTN2A3P	BTN2A3P
C1RL.AS1	C1RL.AS1	C1RL.AS1
C3P1	C3P1	C3P1
CA5BP1	CA5BP1	CA5BP1
CACTIN.AS1	CACTIN.AS1	CACTIN.AS1
CASK.AS1	CASK.AS1	CASK.AS1
CATSPER2P1	CATSPER2P1	CATSPER2P1
CCT6P3	CCT6P3	CCT6P3
CD27.AS1	CD27.AS1	CD27.AS1
CD99P1	CD99P1	CD99P1
CERS6.AS1	CERS6.AS1	CERS6.AS1
CES1P1	CES1P1	CES1P1
CKMT2.AS1	CKMT2.AS1	CKMT2.AS1
CRNDE	CRNDE	CRNDE
CROCCP2	CROCCP2	CROCCP2

CRYBB2P1	CRYBB2P1	CRYBB2P1
CYP2B7P	CYP2B7P	CYP2B7P
DANCR	DANCR	DANCR
DBH.AS1	DBH.AS1	DBH.AS1
DENND5B.AS1	DENND5B.AS1	DENND5B.AS1
DICER1.AS1	DHRS4.AS1	DHRS4.AS1
DIO3OS	DICER1.AS1	DICER1.AS1
DLEU2	DIO3OS	DIO3OS
DNAJC9.AS1	DLEU2	DLEU2
DNM3OS	DNAJC9.AS1	DNAJC9.AS1
DPY19L2P4	DNM3OS	DNM3OS
EAF1.AS1	DPY19L2P4	DPY19L2P4
ENTPD1.AS1	EAF1.AS1	EAF1.AS1
ENTPD3.AS1	ENTPD1.AS1	ENTPD1.AS1
EPB41L4A.AS1	ENTPD3.AS1	ENTPD3.AS1
ERVK13.1	EPB41L4A.AS1	EPB41L4A.AS1
F11.AS1	ERVK13.1	ERVK13.1
FAM13A.AS1	F11.AS1	F11.AS1
FAM157C	FAM13A.AS1	FAM13A.AS1
FAM83A.AS1	FAM157C	FAM157C
FAM86B3P	FAM83A.AS1	FAM83A.AS1
FAM86C2P	FAM86B3P	FAM86B3P
FAM86DP	FAM86C2P	FAM86C2P
FAM86EP	FAM86DP	FAM86DP
FAM99A	FAM86EP	FAM86EP
FAM99B	FAM99A	FAM99A
FGD5.AS1	FAM99B	FAM99B
FLNB.AS1	FGD5.AS1	FGD5.AS1
	FLNB.AS1	FLNB.AS1

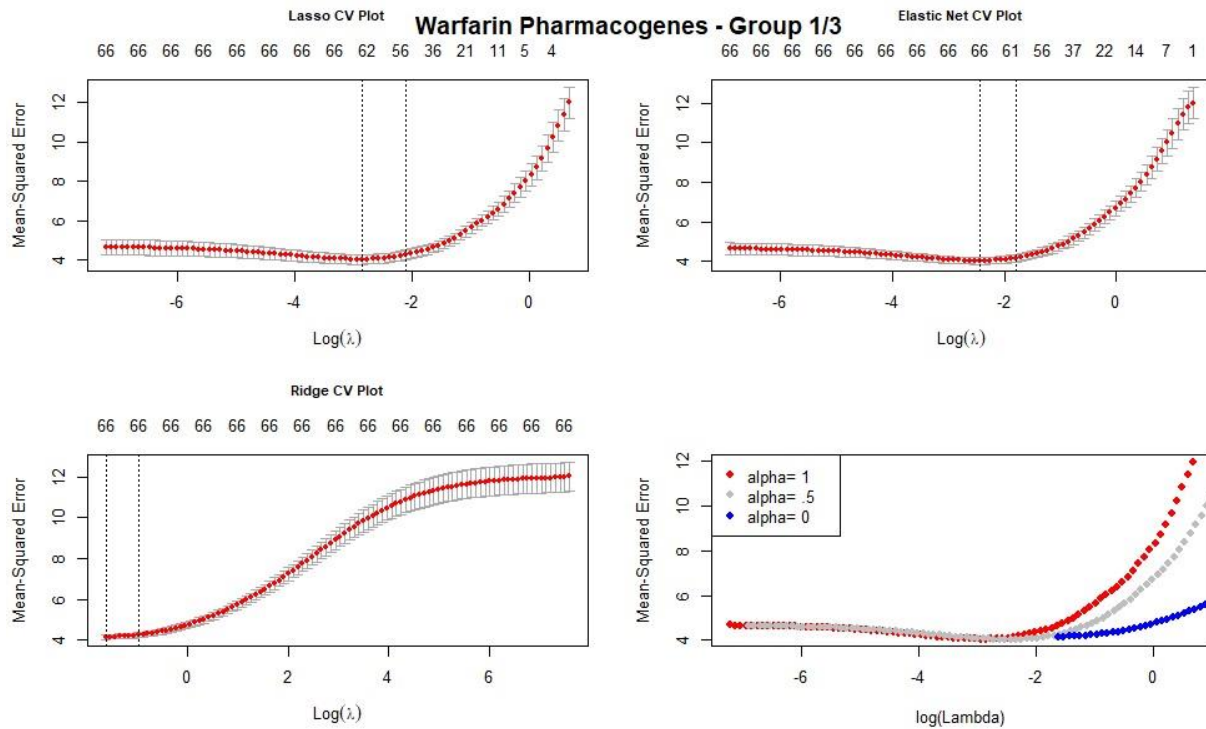


Figure C.2. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against 12 warfarin pharmacogenes (*APOC1*, *APOE*, *CYP1A2*, *CYP2A6*, *CYP2C19*, *CYP2C9*, *CYP3A4*, *CYP4F2*, *EPHX1*, *F5*, *F7*, and *F11*).

Table C.3. LncRNA genes selected as predictors to explain the expression of 11 warfarin pharmacogenes (*ABCB1*, *APOB*, *ASPH*, *CYP1A1*, *CYP2C18*, *CYP4F11*, *DDHD1*, *DNMT3A*, *F13A1*, *F2*, and *VKORC1*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ADORA2A.AS1	ABCC6P2	ABCC6P2
AGAP2.AS1	ADORA2A.AS1	ADORA2A.AS1
ALOX12P2	AGAP2.AS1	AGAP2.AS1
ANO1.AS1	ALOX12P2	ALOX12P2
AOC4P	ANO1.AS1	ANO1.AS1
APOA1.AS	AOC4P	AOC4P
APOC1P1	APOA1.AS	APOA1.AS
ARAP1.AS1	APOC1P1	APOC1P1
ARHGAP5.AS1	ARAP1.AS1	ARAP1.AS1
ARSD.AS1	ARHGAP5.AS1	ARHGAP5.AS1
ASB16.AS1	ARSD.AS1	ARSD.AS1
ASH1L.AS1	ASB16.AS1	ASB16.AS1
ASMTL.AS1	ASH1L.AS1	ASH1L.AS1
AZGP1P1	ASMTL.AS1	ASMTL.AS1
BCRP3	AZGP1P1	AZGP1P1
BDNF.AS	BCRP3	BCRP3
BMS1P1	BDNF.AS	BDNF.AS
BTN2A3P	BMS1P1	BMS1P1
C1RL.AS1	BTN2A3P	BTN2A3P
C3P1	C1RL.AS1	C1RL.AS1
CA5BP1	C3P1	C3P1
CACTIN.AS1	CA5BP1	CA5BP1
CASK.AS1	CACTIN.AS1	CACTIN.AS1
CATSPER2P1	CASK.AS1	CASK.AS1
CCT6P3	CATSPER2P1	CATSPER2P1
CD27.AS1	CCT6P3	CCT6P3
CD99P1	CD27.AS1	CD27.AS1
CERS6.AS1	CD99P1	CD99P1
CES1P1	CERS6.AS1	CERS6.AS1
CKMT2.AS1	CES1P1	CES1P1
CRNDE	CKMT2.AS1	CKMT2.AS1
CROCCP2	CRNDE	CRNDE
CRYBB2P1	CROCCP2	CROCCP2

CYP2B7P	CRYBB2P1	CRYBB2P1
DANCR	CYP2B7P	CYP2B7P
DBH.AS1	DANCR	DANCR
DENND5B.AS1	DBH.AS1	DBH.AS1
DHRS4.AS1	DENND5B.AS1	DENND5B.AS1
DICER1.AS1	DHRS4.AS1	DHRS4.AS1
DIO3OS	DICER1.AS1	DICER1.AS1
DLEU2	DIO3OS	DIO3OS
DNAJC9.AS1	DLEU2	DLEU2
DNM3OS	DNAJC9.AS1	DNAJC9.AS1
DPY19L2P4	DNM3OS	DNM3OS
EAF1.AS1	DPY19L2P4	DPY19L2P4
ENTPD1.AS1	EAF1.AS1	EAF1.AS1
ENTPD3.AS1	ENTPD1.AS1	ENTPD1.AS1
EPB41L4A.AS1	ENTPD3.AS1	ENTPD3.AS1
ERVK13.1	EPB41L4A.AS1	EPB41L4A.AS1
F11.AS1	ERVK13.1	ERVK13.1
FAM13A.AS1	F11.AS1	F11.AS1
FAM157C	FAM13A.AS1	FAM13A.AS1
FAM83A.AS1	FAM157C	FAM157C
FAM86B3P	FAM83A.AS1	FAM83A.AS1
FAM86C2P	FAM86B3P	FAM86B3P
FAM86DP	FAM86C2P	FAM86C2P
FAM86EP	FAM86DP	FAM86DP
FAM99B	FAM86EP	FAM86EP
FGD5.AS1	FAM99A	FAM99A
FLNB.AS1	FAM99B	FAM99B
	FGD5.AS1	FGD5.AS1
	FLNB.AS1	FLNB.AS1

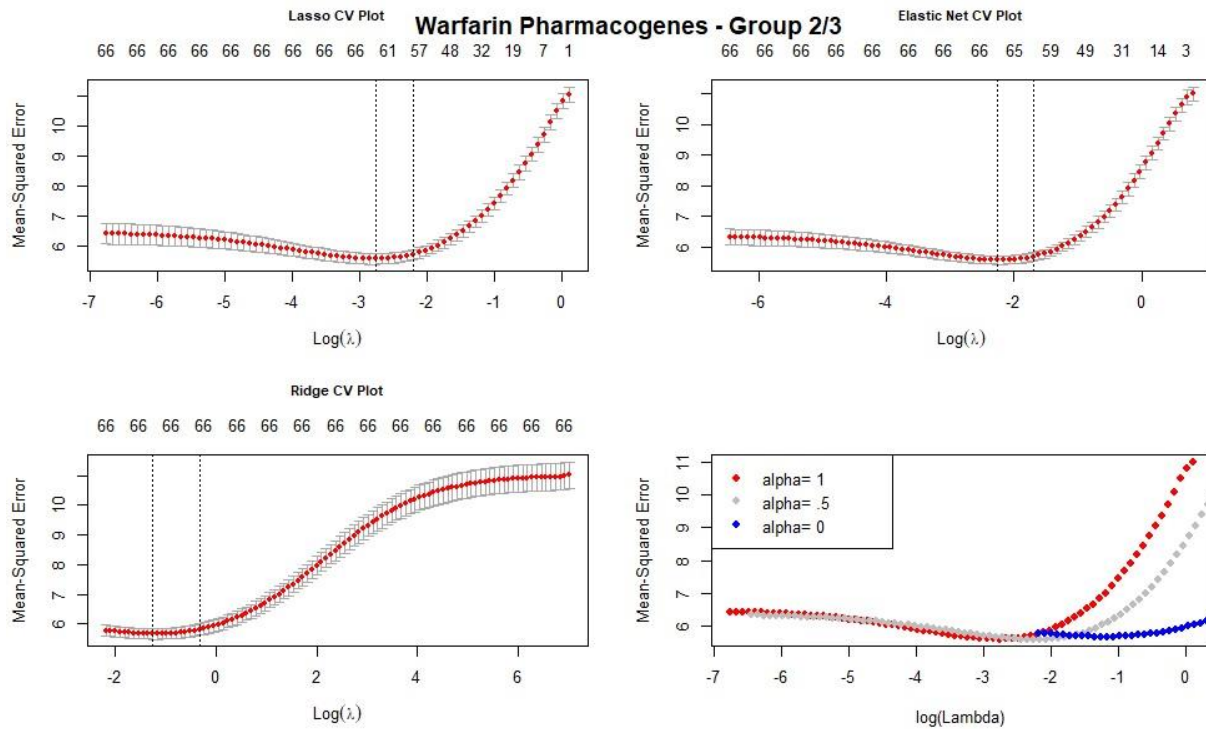


Figure C.3. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against 11 warfarin pharmacogenes (*ABCB1*, *APOB*, *ASPH*, *CYP1A1*, *CYP2C18*, *CYP4F11*, *DDHD1*, *DNMT3A*, *F13A1*, *F2*, and *VKORC1*).

Table C.4. LncRNA genes selected as predictors to explain the expression of two warfarin pharmacogenes (*CALU* and *CRP*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ABCC6P2	ABCC6P2	ABCC6P2
ADORA2A.AS1	ADORA2A.AS1	ADORA2A.AS1
AGAP2.AS1	AGAP2.AS1	AGAP2.AS1
ALOX12P2	ALOX12P2	ALOX12P2
ANO1.AS1	ANO1.AS1	ANO1.AS1
AOC4P	AOC4P	AOC4P
APOA1.AS	APOA1.AS	APOA1.AS
APOC1P1	APOC1P1	APOC1P1
ARHGAP5.AS1	ARAP1.AS1	ARAP1.AS1
ARSD.AS1	ARHGAP5.AS1	ARHGAP5.AS1
ASB16.AS1	ARSD.AS1	ARSD.AS1
ASH1L.AS1	ASB16.AS1	ASB16.AS1
ASMTL.AS1	ASH1L.AS1	ASH1L.AS1
BCRP3	ASMTL.AS1	ASMTL.AS1
BMS1P1	BCRP3	AZGP1P1
BTN2A3P	BMS1P1	BCRP3
C3P1	BTN2A3P	BDNF.AS
CA5BP1	C3P1	BMS1P1
CACTIN.AS1	CA5BP1	BTN2A3P
CATSPER2P1	CACTIN.AS1	C1RL.AS1
CCT6P3	CATSPER2P1	C3P1
CD99P1	CCT6P3	CA5BP1
CERS6.AS1	CD99P1	CACTIN.AS1
CES1P1	CERS6.AS1	CASK.AS1
CKMT2.AS1	CES1P1	CATSPER2P1
CRNDE	CKMT2.AS1	CCT6P3
CROCCP2	CRNDE	CD27.AS1
CYP2B7P	CROCCP2	CD99P1
DANCR	CYP2B7P	CERS6.AS1
DBH.AS1	DANCR	CES1P1
DENND5B.AS1	DBH.AS1	CKMT2.AS1
DHRS4.AS1	DENND5B.AS1	CRNDE
DICER1.AS1	DHRS4.AS1	CROCCP2
DLEU2	DICER1.AS1	CRYBB2P1
DNAJC9.AS1	DIO3OS	CYP2B7P

DNM3OS	DLEU2	DANCR
EAF1.AS1	DNAJC9.AS1	DBH.AS1
ENTPD1.AS1	DNM3OS	DENND5B.AS1
ENTPD3.AS1	EAF1.AS1	DHRS4.AS1
EPB41L4A.AS1	ENTPD1.AS1	DICER1.AS1
F11.AS1	ENTPD3.AS1	DIO3OS
FAM13A.AS1	EPB41L4A.AS1	DLEU2
FAM157C	F11.AS1	DNAJC9.AS1
FAM83A.AS1	FAM13A.AS1	DNM3OS
FAM86B3P	FAM157C	DPY19L2P4
FAM86C2P	FAM83A.AS1	EAF1.AS1
FAM86DP	FAM86B3P	ENTPD1.AS1
FAM86EP	FAM86C2P	ENTPD3.AS1
FGD5.AS1	FAM86DP	EPB41L4A.AS1
FLNB.AS1	FAM86EP	ERVK13.1
	FAM99A	F11.AS1
	FGD5.AS1	FAM13A.AS1
	FLNB.AS1	FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

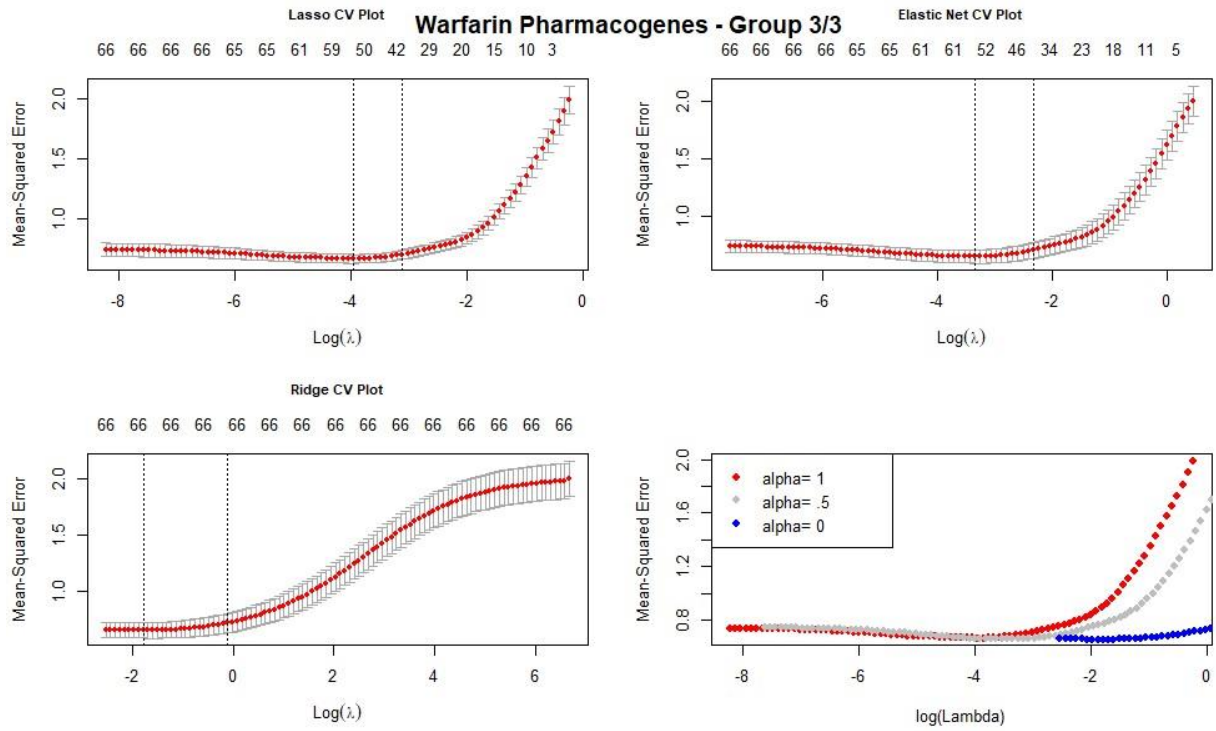


Figure C.4. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against two warfarin pharmacogenes (*CALU* and *CRP*).

Table C.5. LncRNA genes selected as predictors to explain the expression of nine warfarin pharmacogenes (*CYP1A2*, *CYP2A6*, *CYP2C9*, *CYP2C19*, *CYP3A4*, *CYP4F2*, *EPHX1*, *F7*, and *F11*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ABCC6P2	ABCC6P2	ABCC6P2
ADORA2A.AS1	ADORA2A.AS1	ADORA2A.AS1
AGAP2.AS1	AGAP2.AS1	AGAP2.AS1
ALOX12P2	ALOX12P2	ALOX12P2
ANO1.AS1	ANO1.AS1	ANO1.AS1
AOC4P	AOC4P	AOC4P
APOA1.AS	APOA1.AS	APOA1.AS
APOC1P1	APOC1P1	APOC1P1
ARAP1.AS1	ARAP1.AS1	ARAP1.AS1
ARHGAP5.AS1	ARHGAP5.AS1	ARHGAP5.AS1
ARSD.AS1	ARSD.AS1	ARSD.AS1
ASB16.AS1	ASB16.AS1	ASB16.AS1
ASH1L.AS1	ASH1L.AS1	ASH1L.AS1
ASMTL.AS1	ASMTL.AS1	ASMTL.AS1
AZGP1P1	AZGP1P1	AZGP1P1
BCRP3	BCRP3	BCRP3
BDNF.AS	BDNF.AS	BDNF.AS
BMS1P1	BMS1P1	BMS1P1
C1RL.AS1	C1RL.AS1	BTN2A3P
C3P1	C3P1	C1RL.AS1
CA5BP1	CA5BP1	C3P1
CACTIN.AS1	CACTIN.AS1	CA5BP1
CASK.AS1	CASK.AS1	CACTIN.AS1
CATSPER2P1	CATSPER2P1	CASK.AS1
CCT6P3	CCT6P3	CATSPER2P1
CD27.AS1	CD27.AS1	CCT6P3
CD99P1	CD99P1	CD27.AS1
CERS6.AS1	CERS6.AS1	CD99P1
CES1P1	CES1P1	CERS6.AS1
CKMT2.AS1	CKMT2.AS1	CES1P1
CRNDE	CRNDE	CKMT2.AS1
CROCCP2	CROCCP2	CRNDE
CRYBB2P1	CRYBB2P1	CROCCP2
CYP2B7P	CYP2B7P	CRYBB2P1
DANCR	DANCR	CYP2B7P

DBH.AS1	DBH.AS1	DANCR
DENND5B.AS1	DENND5B.AS1	DBH.AS1
DICER1.AS1	DHRS4.AS1	DENND5B.AS1
DIO3OS	DICER1.AS1	DHRS4.AS1
DLEU2	DIO3OS	DICER1.AS1
DNAJC9.AS1	DLEU2	DIO3OS
DNM3OS	DNAJC9.AS1	DLEU2
DPY19L2P4	DNM3OS	DNAJC9.AS1
EAF1.AS1	DPY19L2P4	DNM3OS
ENTPD1.AS1	EAF1.AS1	DPY19L2P4
ENTPD3.AS1	ENTPD1.AS1	EAF1.AS1
EPB41L4A.AS1	ENTPD3.AS1	ENTPD1.AS1
ERVK13.1	EPB41L4A.AS1	ENTPD3.AS1
F11.AS1	ERVK13.1	EPB41L4A.AS1
FAM13A.AS1	F11.AS1	ERVK13.1
FAM157C	FAM13A.AS1	F11.AS1
FAM83A.AS1	FAM157C	FAM13A.AS1
FAM86B3P	FAM83A.AS1	FAM157C
FAM86C2P	FAM86B3P	FAM83A.AS1
FAM86DP	FAM86C2P	FAM86B3P
FAM86EP	FAM86DP	FAM86C2P
FAM99A	FAM86EP	FAM86DP
FAM99B	FAM99A	FAM86EP
FGD5.AS1	FAM99B	FAM99A
FLNB.AS1	FGD5.AS1	FAM99B
	FLNB.AS1	FGD5.AS1
		FLNB.AS1

Table C.6. LncRNA genes selected as predictors to explain the expression of three warfarin pharmacogenes (*APOC1*, *APOE*, and *F5*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A2M.AS1	A2M.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
ABCC6P2	ABCC6P2	ABCC6P1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P2
AGAP2.AS1	AGAP2.AS1	ADORA2A.AS1
ALOX12P2	ALOX12P2	AGAP2.AS1
ANO1.AS1	ANO1.AS1	ALOX12P2
APOA1.AS	APOA1.AS	ANO1.AS1
APOC1P1	APOC1P1	AOC4P
ARAP1.AS1	ARAP1.AS1	APOA1.AS
ARHGAP5.AS1	ARHGAP5.AS1	APOC1P1
ARSD.AS1	ARSD.AS1	ARAP1.AS1
ASB16.AS1	ASB16.AS1	ARHGAP5.AS1
ASH1L.AS1	ASH1L.AS1	ARSD.AS1
AZGP1P1	AZGP1P1	ASB16.AS1
BCRP3	BCRP3	ASH1L.AS1
BDNF.AS	BDNF.AS	ASMTL.AS1
BMS1P1	BMS1P1	AZGP1P1
BTN2A3P	BTN2A3P	BCRP3
C3P1	C3P1	BDNF.AS
CA5BP1	CA5BP1	BMS1P1
CACTIN.AS1	CACTIN.AS1	BTN2A3P
CASK.AS1	CASK.AS1	C1RL.AS1
CATSPER2P1	CATSPER2P1	C3P1
CCT6P3	CCT6P3	CA5BP1
CD27.AS1	CD27.AS1	CACTIN.AS1
CD99P1	CD99P1	CASK.AS1
CERS6.AS1	CERS6.AS1	CATSPER2P1
CES1P1	CES1P1	CCT6P3
CKMT2.AS1	CKMT2.AS1	CD27.AS1
CRNDE	CRNDE	CD99P1
CROCCP2	CROCCP2	CERS6.AS1
CRYBB2P1	CRYBB2P1	CES1P1
CYP2B7P	CYP2B7P	CKMT2.AS1
DBH.AS1	DBH.AS1	CRNDE
DENND5B.AS1	DENND5B.AS1	CROCCP2
DHRS4.AS1	DHRS4.AS1	CRYBB2P1
DICER1.AS1	DICER1.AS1	CYP2B7P

DIO3OS	DIO3OS	DANCR
DLEU2	DLEU2	DBH.AS1
DNAJC9.AS1	DNAJC9.AS1	DENND5B.AS1
DNM3OS	DNM3OS	DHRS4.AS1
DPY19L2P4	DPY19L2P4	DICER1.AS1
EAF1.AS1	EAF1.AS1	DIO3OS
ENTPD1.AS1	ENTPD1.AS1	DLEU2
ENTPD3.AS1	ENTPD3.AS1	DNAJC9.AS1
EPB41L4A.AS1	EPB41L4A.AS1	DNM3OS
ERVK13.1	ERVK13.1	DPY19L2P4
F11.AS1	F11.AS1	EAF1.AS1
FAM13A.AS1	FAM13A.AS1	ENTPD1.AS1
FAM157C	FAM157C	ENTPD3.AS1
FAM83A.AS1	FAM83A.AS1	EPB41L4A.AS1
FAM86B3P	FAM86B3P	ERVK13.1
FAM86DP	FAM86DP	F11.AS1
FAM99A	FAM99A	FAM13A.AS1
FGD5.AS1	FGD5.AS1	FAM157C
FLNB.AS1	FLNB.AS1	FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

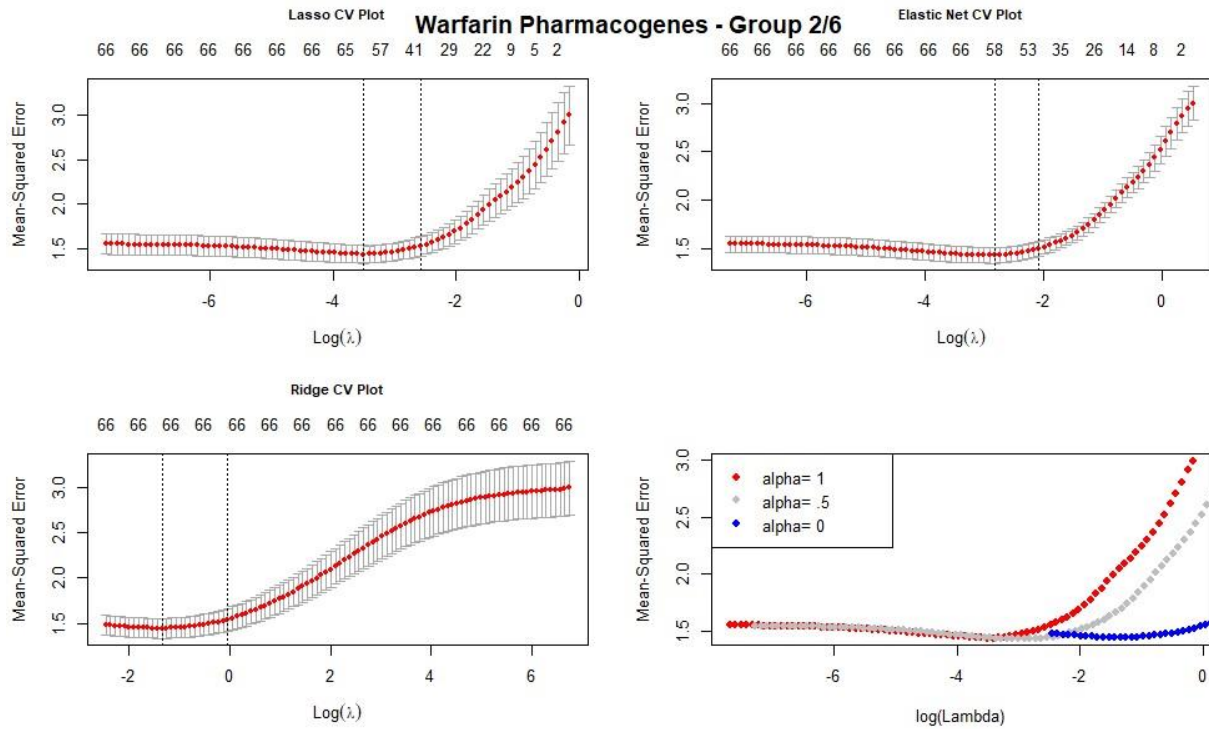


Figure C.6. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against three warfarin pharmacogenes (*APOC1*, *APOE*, and *F5*).

Table C.7. LncRNA genes selected as predictors to explain the expression of five warfarin pharmacogenes (*APOB*, *ASPH*, *CYP2C18*, *CYP4F11*, and *F2*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
AGAP2.AS1	AGAP2.AS1	ABCC6P2
ALOX12P2	ALOX12P2	ADORA2A.AS1
ANO1.AS1	ANO1.AS1	AGAP2.AS1
AOC4P	AOC4P	ALOX12P2
APOA1.AS	APOA1.AS	ANO1.AS1
APOC1P1	APOC1P1	AOC4P
ARAP1.AS1	ARAP1.AS1	APOA1.AS
ARHGAP5.AS1	ARHGAP5.AS1	APOC1P1
ARSD.AS1	ARSD.AS1	ARAP1.AS1
ASB16.AS1	ASB16.AS1	ARHGAP5.AS1
ASH1L.AS1	ASH1L.AS1	ARSD.AS1
AZGP1P1	AZGP1P1	ASB16.AS1
BCRP3	BCRP3	ASH1L.AS1
BDNF.AS	BDNF.AS	ASMTL.AS1
BMS1P1	BMS1P1	AZGP1P1
BTN2A3P	BTN2A3P	BCRP3
C1RL.AS1	C1RL.AS1	BDNF.AS
C3P1	C3P1	BMS1P1
CA5BP1	CA5BP1	BTN2A3P
CACTIN.AS1	CACTIN.AS1	C1RL.AS1
CASK.AS1	CASK.AS1	C3P1
CATSPER2P1	CATSPER2P1	CA5BP1
CCT6P3	CCT6P3	CACTIN.AS1
CD27.AS1	CD27.AS1	CASK.AS1
CERS6.AS1	CERS6.AS1	CATSPER2P1
CES1P1	CES1P1	CCT6P3
CKMT2.AS1	CKMT2.AS1	CD27.AS1
CRNDE	CRNDE	CD99P1
CROCCP2	CROCCP2	CERS6.AS1
CRYBB2P1	CRYBB2P1	CES1P1
CYP2B7P	CYP2B7P	CKMT2.AS1
DANCR	DANCR	CRNDE
DBH.AS1	DBH.AS1	CROCCP2
DENND5B.AS1	DENND5B.AS1	CRYBB2P1
DHRS4.AS1	DHRS4.AS1	CYP2B7P

DICER1.AS1	DICER1.AS1	DANCR
DIO3OS	DIO3OS	DBH.AS1
DLEU2	DLEU2	DENND5B.AS1
DNAJC9.AS1	DNAJC9.AS1	DHRS4.AS1
DNM3OS	DNM3OS	DICER1.AS1
DPY19L2P4	DPY19L2P4	DIO3OS
EAF1.AS1	EAF1.AS1	DLEU2
ENTPD1.AS1	ENTPD1.AS1	DNAJC9.AS1
ENTPD3.AS1	ENTPD3.AS1	DNM3OS
EPB41L4A.AS1	EPB41L4A.AS1	DPY19L2P4
ERVK13.1	ERVK13.1	EAF1.AS1
F11.AS1	F11.AS1	ENTPD1.AS1
FAM13A.AS1	FAM13A.AS1	ENTPD3.AS1
FAM157C	FAM157C	EPB41L4A.AS1
FAM83A.AS1	FAM83A.AS1	ERVK13.1
FAM86B3P	FAM86B3P	F11.AS1
FAM86C2P	FAM86C2P	FAM13A.AS1
FAM86DP	FAM86DP	FAM157C
FAM86EP	FAM86EP	FAM83A.AS1
FAM99A	FAM99A	FAM86B3P
FAM99B	FAM99B	FAM86C2P
FGD5.AS1	FGD5.AS1	FAM86DP
FLNB.AS1	FLNB.AS1	FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

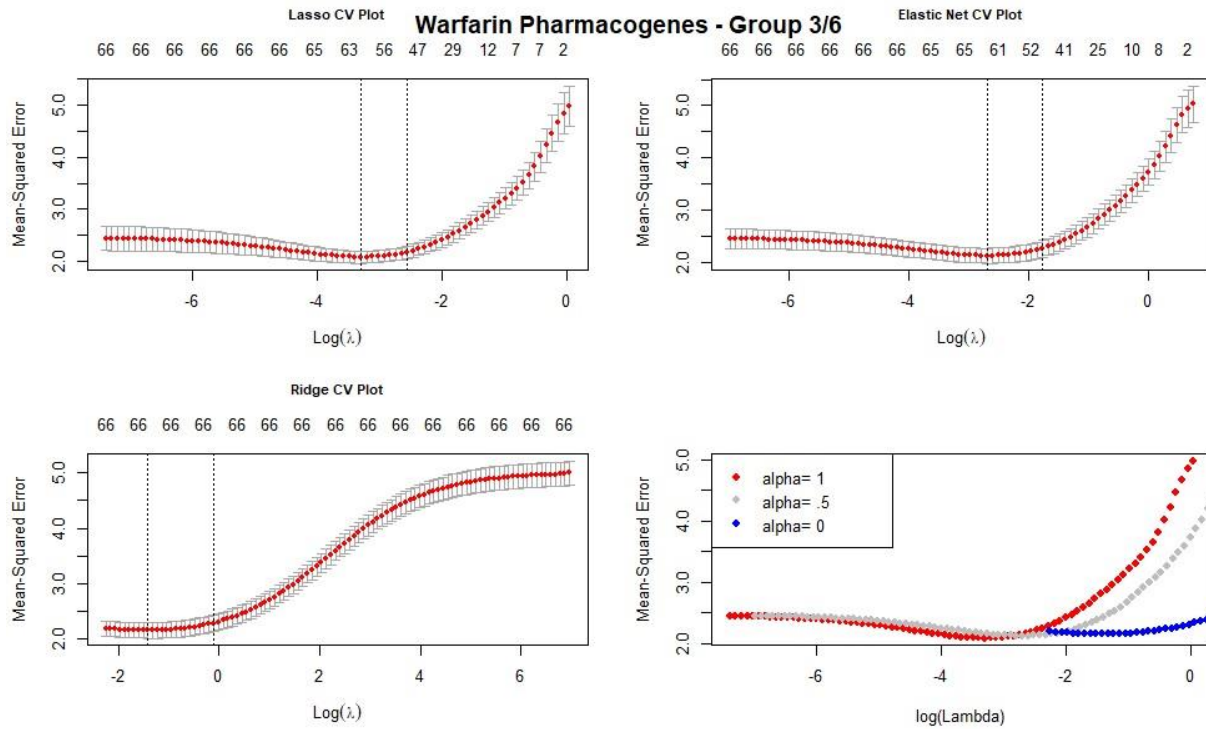


Figure C.7. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against five warfarin pharmacogenes (*APOB*, *ASPH*, *CYP2C18*, *CYP4F11*, and *F2*).

Table C.8. LncRNA genes selected as predictors to explain the expression of two warfarin pharmacogenes (*ABCB1* and *CYP1A1*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
ABCC6P1	ABCC6P1	A2M.AS1
ABCC6P2	ABCC6P2	ABCC6P1
ADORA2A.AS1	ADORA2A.AS1	ABCC6P2
AGAP2.AS1	AGAP2.AS1	ADORA2A.AS1
ANO1.AS1	ANO1.AS1	AGAP2.AS1
APOA1.AS	APOA1.AS	ALOX12P2
APOC1P1	APOC1P1	ANO1.AS1
ARAP1.AS1	ARAP1.AS1	AOC4P
ARSD.AS1	ARSD.AS1	APOA1.AS
ASB16.AS1	ASB16.AS1	APOC1P1
AZGP1P1	AZGP1P1	ARAP1.AS1
BDNF.AS	BDNF.AS	ARHGAP5.AS1
CA5BP1	BMS1P1	ARSD.AS1
CACTIN.AS1	CA5BP1	ASB16.AS1
CASK.AS1	CACTIN.AS1	ASH1L.AS1
CCT6P3	CASK.AS1	ASMTL.AS1
CD27.AS1	CCT6P3	AZGP1P1
CERS6.AS1	CD27.AS1	BCRP3
CES1P1	CERS6.AS1	BDNF.AS
CKMT2.AS1	CES1P1	BMS1P1
CRNDE	CKMT2.AS1	BTN2A3P
CROCCP2	CRNDE	C1RL.AS1
CRYBB2P1	CROCCP2	C3P1
CYP2B7P	CRYBB2P1	CA5BP1
DBH.AS1	CYP2B7P	CACTIN.AS1
DENND5B.AS1	DBH.AS1	CASK.AS1
DHRS4.AS1	DENND5B.AS1	CATSPER2P1
DICER1.AS1	DHRS4.AS1	CCT6P3
DIO3OS	DICER1.AS1	CD27.AS1
DLEU2	DIO3OS	CD99P1
DNM3OS	DLEU2	CERS6.AS1
DPY19L2P4	DNM3OS	CES1P1
EPB41L4A.AS1	DPY19L2P4	CKMT2.AS1
FAM13A.AS1	EPB41L4A.AS1	CRNDE
FAM157C	FAM13A.AS1	CROCCP2
FAM86B3P	FAM157C	CRYBB2P1
FAM86DP	FAM86B3P	CYP2B7P

FLNB.AS1	FAM86DP	DANCR
	FLNB.AS1	DBH.AS1
		DENND5B.AS1
		DHRS4.AS1
		DICER1.AS1
		DIO3OS
		DLEU2
		DNAJC9.AS1
		DNM3OS
		DPY19L2P4
		EAF1.AS1
		ENTPD1.AS1
		ENTPD3.AS1
		EPB41L4A.AS1
		ERVK13.1
		F11.AS1
		FAM13A.AS1
		FAM157C
		FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

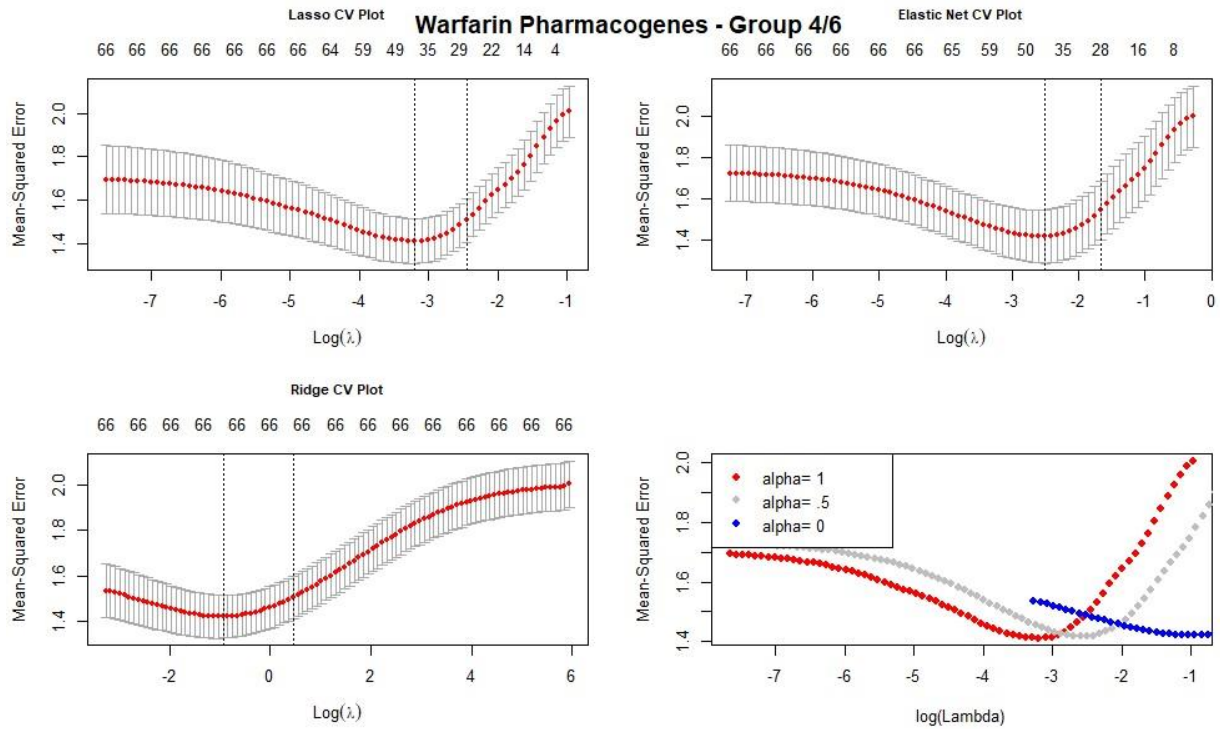


Figure C.8. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against two warfarin pharmacogenes (*ABCB1* and *CYP1A1*).

Table C.9. LncRNA genes selected as predictors to explain the expression of four warfarin pharmacogenes (*DDHD1*, *DNMT3A*, *F13A1*, and *VKORC1*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ABCC6P2	ABCC6P2	ABCC6P2
ADORA2A.AS1	ADORA2A.AS1	ADORA2A.AS1
AGAP2.AS1	AGAP2.AS1	AGAP2.AS1
ALOX12P2	ALOX12P2	ALOX12P2
ANO1.AS1	ANO1.AS1	ANO1.AS1
AOC4P	AOC4P	AOC4P
APOA1.AS	APOA1.AS	APOA1.AS
APOC1P1	APOC1P1	APOC1P1
ARAP1.AS1	ARAP1.AS1	ARAP1.AS1
ARHGAP5.AS1	ARHGAP5.AS1	ARHGAP5.AS1
ARSD.AS1	ARSD.AS1	ARSD.AS1
ASB16.AS1	ASB16.AS1	ASB16.AS1
ASH1L.AS1	ASH1L.AS1	ASH1L.AS1
ASMTL.AS1	ASMTL.AS1	ASMTL.AS1
AZGP1P1	AZGP1P1	AZGP1P1
BCRP3	BCRP3	BCRP3
BDNF.AS	BDNF.AS	BDNF.AS
BMS1P1	BMS1P1	BMS1P1
BTN2A3P	BTN2A3P	BTN2A3P
C1RL.AS1	C1RL.AS1	C1RL.AS1
C3P1	C3P1	C3P1
CA5BP1	CA5BP1	CA5BP1
CACTIN.AS1	CACTIN.AS1	CACTIN.AS1
CASK.AS1	CASK.AS1	CASK.AS1
CCT6P3	CCT6P3	CATSPER2P1
CD27.AS1	CD27.AS1	CCT6P3
CD99P1	CD99P1	CD27.AS1
CERS6.AS1	CERS6.AS1	CD99P1
CES1P1	CES1P1	CERS6.AS1
CKMT2.AS1	CKMT2.AS1	CES1P1
CRNDE	CRNDE	CKMT2.AS1
CROCCP2	CROCCP2	CRNDE
CRYBB2P1	CRYBB2P1	CROCCP2
CYP2B7P	CYP2B7P	CRYBB2P1
DANCR	DANCR	CYP2B7P

DBH.AS1	DBH.AS1	DANCR
DENND5B.AS1	DENND5B.AS1	DBH.AS1
DHRS4.AS1	DHRS4.AS1	DENND5B.AS1
DICER1.AS1	DICER1.AS1	DHRS4.AS1
DIO3OS	DIO3OS	DICER1.AS1
DLEU2	DLEU2	DIO3OS
DNAJC9.AS1	DNAJC9.AS1	DLEU2
DNM3OS	DNM3OS	DNAJC9.AS1
DPY19L2P4	DPY19L2P4	DNM3OS
EAF1.AS1	EAF1.AS1	DPY19L2P4
ENTPD1.AS1	ENTPD1.AS1	EAF1.AS1
ENTPD3.AS1	ENTPD3.AS1	ENTPD1.AS1
EPB41L4A.AS1	EPB41L4A.AS1	ENTPD3.AS1
ERVK13.1	ERVK13.1	EPB41L4A.AS1
F11.AS1	F11.AS1	ERVK13.1
FAM13A.AS1	FAM13A.AS1	F11.AS1
FAM157C	FAM157C	FAM13A.AS1
FAM83A.AS1	FAM83A.AS1	FAM157C
FAM86B3P	FAM86B3P	FAM83A.AS1
FAM86C2P	FAM86C2P	FAM86B3P
FAM86DP	FAM86DP	FAM86C2P
FAM86EP	FAM86EP	FAM86DP
FAM99A	FAM99A	FAM86EP
FAM99B	FAM99B	FAM99A
FGD5.AS1	FGD5.AS1	FAM99B
FLNB.AS1	FLNB.AS1	FGD5.AS1
		FLNB.AS1

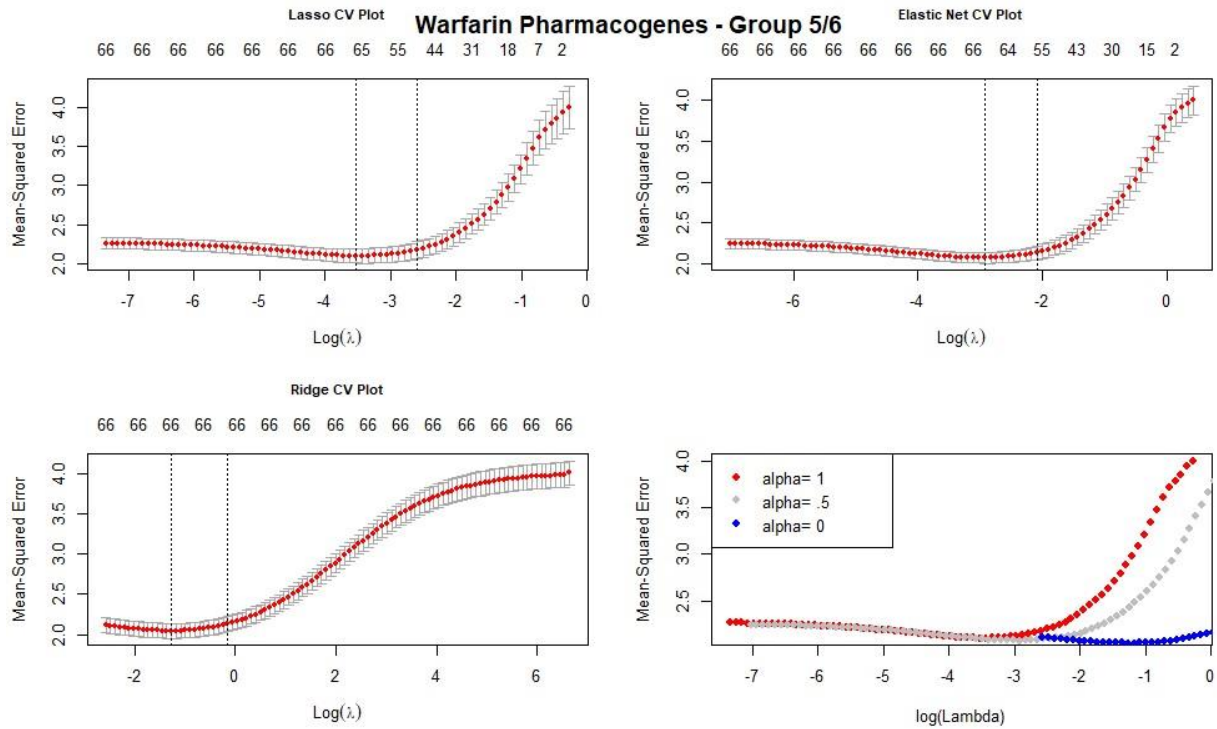


Figure C.9. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against four warfarin pharmacogenes (*DDHD1*, *DNMT3A*, *F13A1*, and *VKORC1*).

Table C.10. LncRNA genes selected as predictors to explain the expression of two warfarin pharmacogenes (*CALU* and *CRP*) in the context of the best fitting model under Lasso, Elastic Net, and Ridge penalized linear regression models.

Lasso	Elastic Net	Ridge
A1BG.AS1	A1BG.AS1	A1BG.AS1
A2M.AS1	A2M.AS1	A2M.AS1
ABCC6P1	ABCC6P1	ABCC6P1
ABCC6P2	ABCC6P2	ABCC6P2
ADORA2A.AS1	ADORA2A.AS1	ADORA2A.AS1
AGAP2.AS1	AGAP2.AS1	AGAP2.AS1
ALOX12P2	ALOX12P2	ALOX12P2
ANO1.AS1	ANO1.AS1	ANO1.AS1
AOC4P	AOC4P	AOC4P
APOA1.AS	APOA1.AS	APOA1.AS
APOC1P1	APOC1P1	APOC1P1
ARHGAP5.AS1	ARAP1.AS1	ARAP1.AS1
ARSD.AS1	ARHGAP5.AS1	ARHGAP5.AS1
ASB16.AS1	ARSD.AS1	ARSD.AS1
ASH1L.AS1	ASB16.AS1	ASB16.AS1
ASMTL.AS1	ASH1L.AS1	ASH1L.AS1
BCRP3	ASMTL.AS1	ASMTL.AS1
BMS1P1	BCRP3	AZGP1P1
BTN2A3P	BMS1P1	BCRP3
C3P1	BTN2A3P	BDNF.AS
CA5BP1	C1RL.AS1	BMS1P1
CACTIN.AS1	C3P1	BTN2A3P
CATSPER2P1	CA5BP1	C1RL.AS1
CCT6P3	CACTIN.AS1	C3P1
CD99P1	CATSPER2P1	CA5BP1
CERS6.AS1	CCT6P3	CACTIN.AS1
CES1P1	CD99P1	CASK.AS1
CKMT2.AS1	CERS6.AS1	CATSPER2P1
CRNDE	CES1P1	CCT6P3
CROCCP2	CKMT2.AS1	CD27.AS1
CYP2B7P	CRNDE	CD99P1
DANCR	CROCCP2	CERS6.AS1
DBH.AS1	CYP2B7P	CES1P1
DENND5B.AS1	DANCR	CKMT2.AS1
DHRS4.AS1	DBH.AS1	CRNDE
DICER1.AS1	DENND5B.AS1	CROCCP2
DLEU2	DHRS4.AS1	CRYBB2P1
DNAJC9.AS1	DICER1.AS1	CYP2B7P

DNM3OS	DIO3OS	DANCR
EAF1.AS1	DLEU2	DBH.AS1
ENTPD1.AS1	DNAJC9.AS1	DENND5B.AS1
ENTPD3.AS1	DNM3OS	DHRS4.AS1
EPB41L4A.AS1	EAF1.AS1	DICER1.AS1
F11.AS1	ENTPD1.AS1	DIO3OS
FAM13A.AS1	ENTPD3.AS1	DLEU2
FAM157C	EPB41L4A.AS1	DNAJC9.AS1
FAM83A.AS1	F11.AS1	DNM3OS
FAM86B3P	FAM13A.AS1	DPY19L2P4
FAM86C2P	FAM157C	EAF1.AS1
FAM86DP	FAM83A.AS1	ENTPD1.AS1
FAM86EP	FAM86B3P	ENTPD3.AS1
FAM99A	FAM86C2P	EPB41L4A.AS1
FGD5.AS1	FAM86DP	ERVK13.1
FLNB.AS1	FAM86EP	F11.AS1
	FAM99A	FAM13A.AS1
	FGD5.AS1	FAM157C
	FLNB.AS1	FAM83A.AS1
		FAM86B3P
		FAM86C2P
		FAM86DP
		FAM86EP
		FAM99A
		FAM99B
		FGD5.AS1
		FLNB.AS1

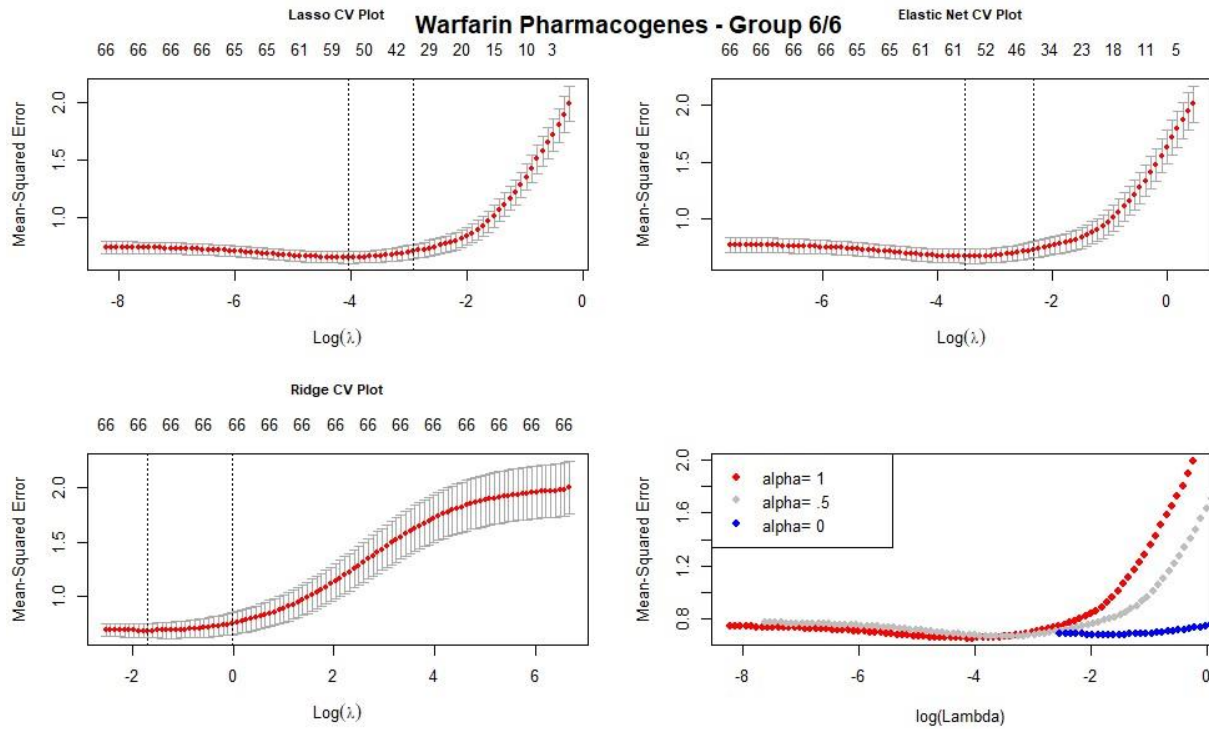


Figure C.10. Panel of Lasso ($\alpha = 1$), Elastic Net ($\alpha = 0.5$), and Ridge ($\alpha = 0$) penalized linear regression model cross-validation plots. All significant lncRNA genes were tested against two warfarin pharmacogenes (*CALU* and *CRP*).