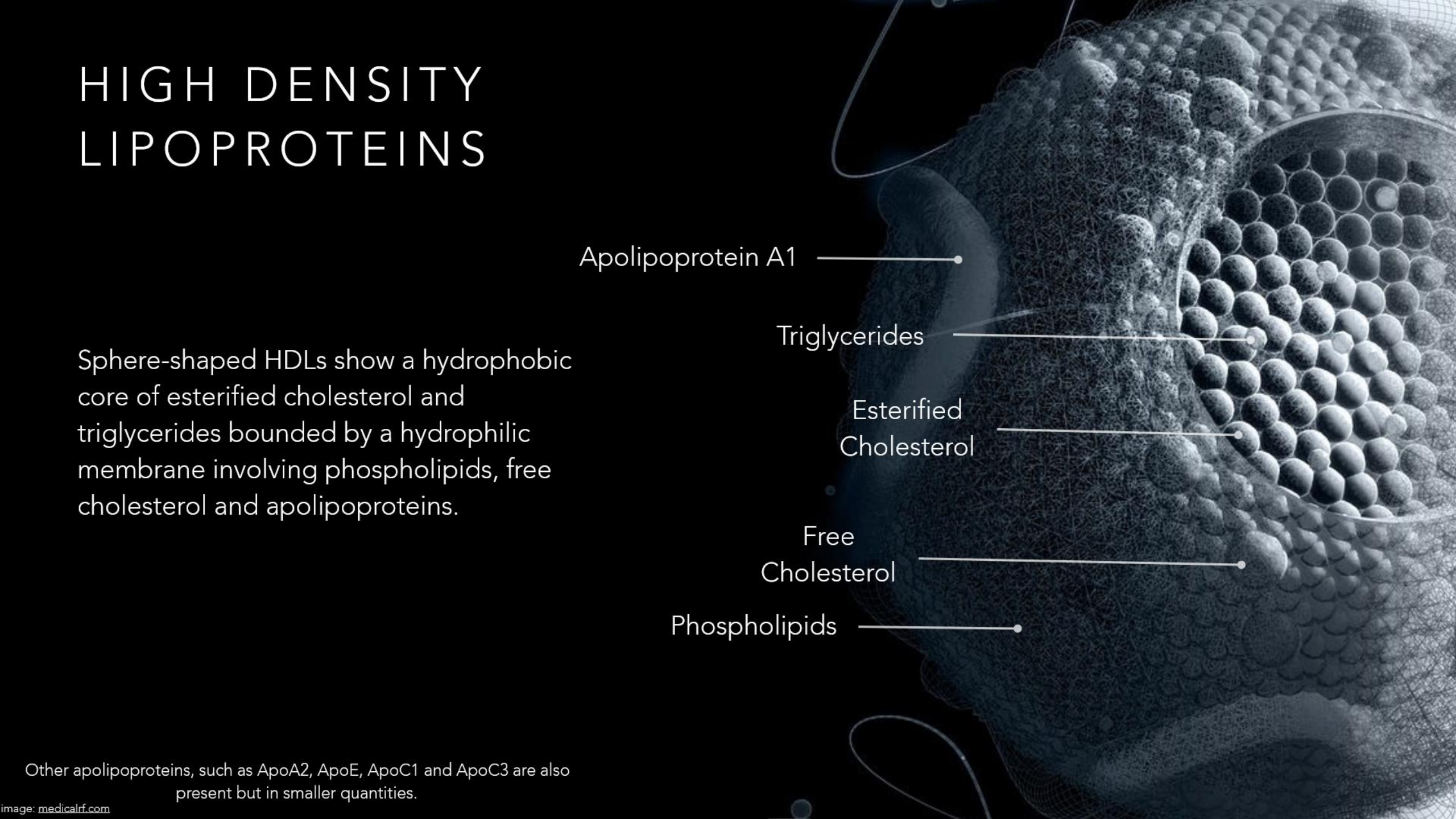
PANAGIOTIS TOGIAS

MULTIVARIATE DATA ANALYSIS LINKING STRUCTURAL AND FUNCTIONAL PROPERTIES OF HIGH DENSITY LIPOPROTEINS

UNIVERSITY OF PATRAS

MSC LIFE SCIENCES INFORMATICS

HIGH DENSITY LIPOPROTEINS



Sphere-shaped HDLs show a hydrophobic core of esterified cholesterol and triglycerides bounded by a hydrophilic membrane involving phospholipids, free cholesterol and apolipoproteins.

Other apolipoproteins, such as ApoA2, ApoE, ApoC1 and ApoC3 are also present but in smaller quantities.

image: medicalrf.com

HDL FUNCTIONS

Reverse Cholesterol Transfer

The removal of excessive cholesterol from peripheral tissues and its transport to the liver for catabolism

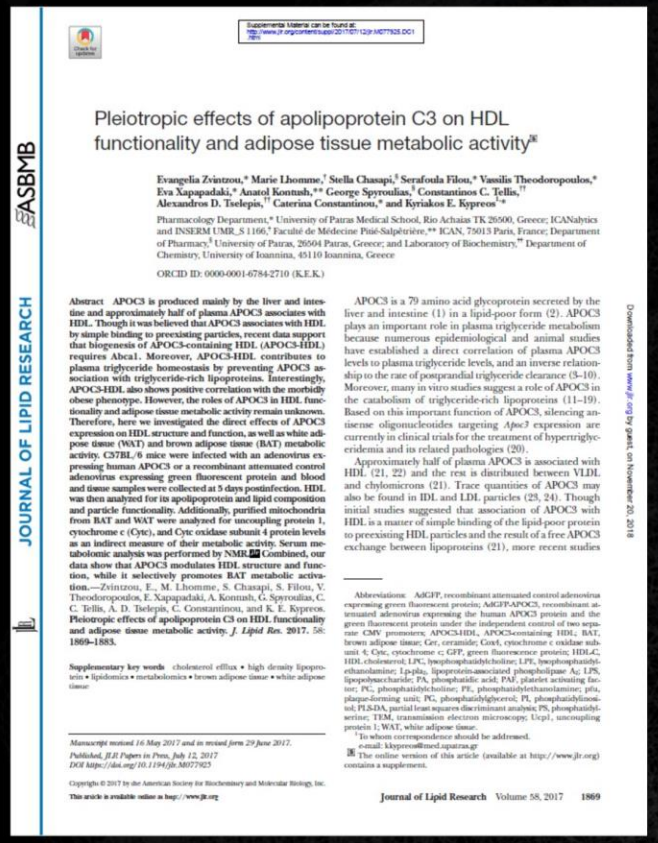
Antioxidant Properties

HDL weakens atherogenesis through ApoA1 which has the ability of eliminating oxidant molecules by binding them and carry them away

Anti-inflammatory Properties

HDL and its major apolipoproteins reduce free cholesterol from plasma membrane by suppressing the signaling pathways of inflammatory response

PREVIOUS WORK



Supplemental Material can be found at: http://www.jlr.org/content/suppl/2017/07/12/jlr.M077925.DC1

Pleiotropic effects of apolipoprotein C3 on HDL functionality and adipose tissue metabolic activity[S]

Evangelia Zvintzou,* Marie Lhomme,† Stella Chasapi,§ Serafoula Filou,* Vassilis Theodoropoulos,* Eva Xapapadaki,* Anatol Kontush,** George Spyroulias,§ Constantinos C. Tellis,†† Alexandros D. Tselepis,†† Caterina Constantinou,* and Kyriakos E. Kypreos*,1

Pharmacology Department,* University of Patras Medical School, Rio Achaias TK 26500, Greece; ICANalytics and INSERM UMR_S 1166,† Faculté de Médecine Pitié-Salpêtrière,** ICAN, 75013 Paris, France; Department of Pharmacy,§ University of Patras, 26504 Patras, Greece; and Laboratory of Biochemistry,†† Department of Chemistry, University of Ioannina, 45110 Ioannina, Greece

ORCID ID: 0000-0001-6784-2710 (K.E.K.)

Abstract **APOC3 is produced mainly by the liver and intestine and approximately half of plasma APOC3 associates with HDL. Though it was believed that APOC3 associates with HDL by simple binding to preexisting particles, recent data support that biogenesis of APOC3-containing HDL (APOC3-HDL) requires Abca1. Moreover, APOC3-HDL contributes to plasma triglyceride homeostasis by preventing APOC3 association with triglyceride-rich lipoproteins. Interestingly, APOC3-HDL also shows positive correlation with the morbidly obese phenotype. However, the roles of APOC3 in HDL functionality and adipose tissue metabolic activity remain unknown. Therefore, here we investigated the direct effects of APOC3 expression on HDL structure and function, as well as white adipose tissue (WAT) and brown adipose tissue (BAT) metabolic activity. C57BL/6 mice were infected with an adenovirus expressing human APOC3 or a recombinant attenuated control adenovirus expressing green fluorescent protein and blood and tissue samples were collected at 5 days postinfection. HDL was then analyzed for its apolipoprotein and lipid composition and particle functionality. Additionally, purified mitochondria from BAT and WAT were analyzed for uncoupling protein 1, cytochrome c (Cytc), and Cytc oxidase subunit 4 protein levels as an indirect measure of their metabolic activity. Serum metabolomic analysis was performed by NMR. Combined, our data show that APOC3 modulates HDL structure and function, while it selectively promotes BAT metabolic activation.**—Zvintzou, E., M. Lhomme, S. Chasapi, S. Filou, V. Theodoropoulos, E. Xapapadaki, A. Kontush, G. Spyroulias, C. C. Tellis, A. D. Tselepis, C. Constantinou, and K. E. Kypreos. **Pleiotropic effects of apolipoprotein C3 on HDL functionality and adipose tissue metabolic activity.** ***J. Lipid Res.*** **2017.** 58: **1869–1883.**

Supplementary key words cholesterol efflux • high density lipoprotein • lipidomics • metabolomics • brown adipose tissue • white adipose tissue

APOC3 is a 79 amino acid glycoprotein secreted by the liver and intestine (1) in a lipid-poor form (2). APOC3 plays an important role in plasma triglyceride metabolism because numerous epidemiological and animal studies have established a direct correlation of plasma APOC3 levels to plasma triglyceride levels, and an inverse relationship to the rate of postprandial triglyceride clearance (3–10). Moreover, many in vitro studies suggest a role of APOC3 in the catabolism of triglyceride-rich lipoproteins (11–19). Based on this important function of APOC3, silencing antisense oligonucleotides targeting *Apoc3* expression are currently in clinical trials for the treatment of hypertriglyceridemia and its related pathologies (20).

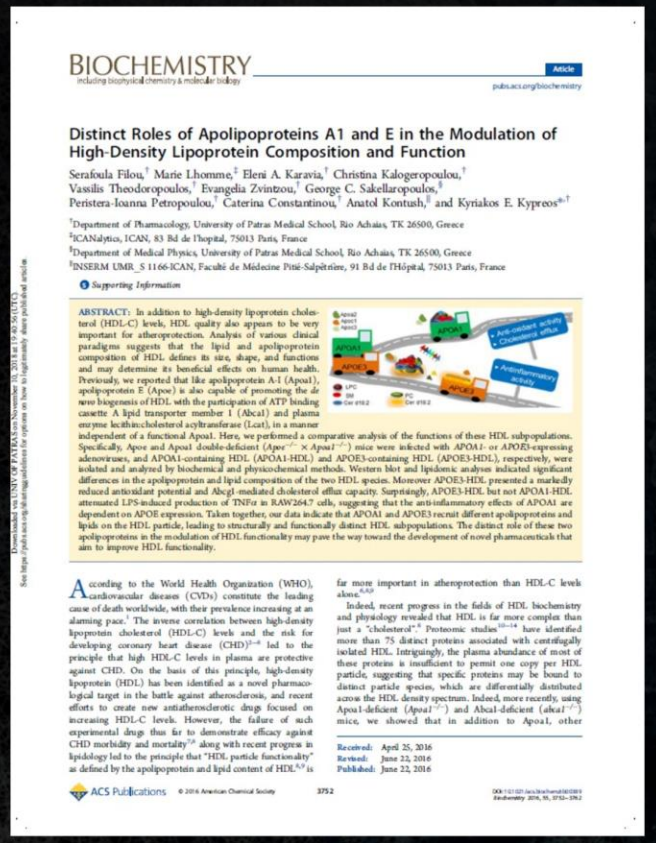
Approximately half of plasma APOC3 is associated with HDL (21, 22) and the rest is distributed between VLDL and chylomicrons (21). Trace quantities of APOC3 may also be found in IDL and LDL particles (23, 24). Though initial studies suggested that association of APOC3 with HDL is a matter of simple binding of the lipid-poor protein to preexisting HDL particles and the result of a free APOC3 exchange between lipoproteins (21), more recent studies

Abbreviations: AdGFP, recombinant attenuated control adenovirus expressing green fluorescent protein; AdGFP-APOC3, recombinant attenuated adenovirus expressing the human APOC3 protein and the green fluorescent protein under the independent control of two separate CMV promoters; APOC3-HDL, APOC3-containing HDL; BAT, brown adipose tissue; Cer, ceramide; Cox4, cytochrome c oxidase subunit 4; Cytc, cytochrome c; GFP, green fluorescence protein; HDL-C, HDL cholesterol; LPC, lysophosphatidylcholine; LPE, lysophosphatidylethanolamine; Lp-pla$_2$, lipoprotein-associated phospholipase A$_2$; LPS, lipopolysaccharide; PA, phosphatidic acid; PAF, platelet activating factor; PC, phosphatidylcholine; PE, phosphatidylethanolamine; pfu, plaque-forming unit; PG, phosphatidylglycerol; PI, phosphatidylinositol; PLS-DA, partial least squares discriminant analysis; PS, phosphatidylserine; TEM, transmission electron microscopy; Ucp1, uncoupling protein 1; WAT, white adipose tissue.

To whom correspondence should be addressed.
e-mail: kkypreos@med.upatras.gr

[S] The online version of this article (available at http://www.jlr.org) contains a supplement.

Manuscript received 16 May 2017 and in revised form 29 June 2017.
Published, JLR Papers in Press, July 12, 2017
DOI https://doi.org/10.1194/jlr.M077925



BIOCHEMISTRY including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Distinct Roles of Apolipoproteins A1 and E in the Modulation of High-Density Lipoprotein Composition and Function

Serafoula Filou,[†] Marie Lhomme,[‡] Eleni A. Karavia,[†] Christina Kalogeropoulou,[†] Vassilis Theodoropoulos,[†] Evangelia Zvintzou,[†] George C. Sakellaropoulos,[§] Peristera-Ioanna Petropoulou,[†] Caterina Constantinou,[†] Anatol Kontush,[∥] and Kyriakos E. Kypreos*[,†]

[†]Department of Pharmacology, University of Patras Medical School, Rio Achaias, TK 26500, Greece
[‡]ICANalytics, ICAN, 83 Bd de l'hopital, 75013 Paris, France
[§]Department of Medical Physics, University of Patras Medical School, Rio Achaias, TK 26500, Greece
[∥]INSERM UMR_S 1166-ICAN, Faculté de Médecine Pitié-Salpêtrière, 91 Bd de l'Hôpital, 75013 Paris, France

Supporting Information

ABSTRACT: In addition to high-density lipoprotein cholesterol (HDL-C) levels, HDL quality also appears to be very important for atheroprotection. Analysis of various clinical paradigms suggests that the lipid and apolipoprotein composition of HDL defines its size, shape, and functions and may determine its beneficial effects on human health. Previously, we reported that like apolipoprotein A-I (Apoa1), apolipoprotein E (Apoe) is also capable of promoting the de novo biogenesis of HDL with the participation of ATP binding cassette A1 lipid transporter member 1 (Abca1) and plasma enzyme lecithin:cholesterol acyltransferase (Lcat), in a manner independent of a functional Apoa1. Here, we performed a comparative analysis of the functions of these HDL subpopulations. Specifically, Apoe and Apoa1 double-deficient (*Apoe*$^{-/-}$ × *Apoa1*$^{-/-}$) mice were infected with *APOA1*- or *APOE3*-expressing adenoviruses, and APOA1-containing HDL (APOA1-HDL) and APOE3-containing HDL (APOE3-HDL), respectively, were isolated and analyzed by biochemical and physicochemical methods. Western blot and lipidomic analyses indicated significant differences in the apolipoprotein and lipid composition of the two HDL species. Moreover APOE3-HDL presented a markedly reduced antioxidant potential and Abcg1-mediated cholesterol efflux capacity. Surprisingly, APOE3-HDL but not APOA1-HDL attenuated LPS-induced production of TNFα in RAW264.7 cells, suggesting that the anti-inflammatory effects of APOA1 are dependent on APOE expression. Taken together, our data indicate that APOA1 and APOE3 recruit different apolipoproteins and lipids on the HDL particle, leading to structurally and functionally distinct HDL subpopulations. The distinct role of these two apolipoproteins in the modulation of HDL functionality may pave the way toward the development of novel pharmaceuticals that aim to improve HDL functionality.

According to the World Health Organization (WHO), cardiovascular diseases (CVDs) constitute the leading cause of death worldwide, with their prevalence increasing at an alarming pace.[1] The inverse correlation between high-density lipoprotein cholesterol (HDL-C) levels and the risk for developing coronary heart disease (CHD)[2−6] led to the principle that high HDL-C levels in plasma are protective against CHD. On the basis of this principle, high-density lipoprotein (HDL) has been identified as a novel pharmacological target in the battle against atherosclerosis, and recent efforts to create new antiatherosclerotic drugs focused on increasing HDL-C levels. However, the failure of such experimental drugs thus far to demonstrate efficacy against CHD morbidity and mortality[7,8] along with recent progress in lipidology led to the principle that "HDL particle functionality" as defined by the apolipoprotein and lipid content of HDL[9,5] is far more important in atheroprotection than HDL-C levels alone.[5,6,9]

Indeed, recent progress in the fields of HDL biochemistry and physiology revealed that HDL is far more complex than just a "cholesterol".[6] Proteomic studies[10−14] have identified more than 75 distinct proteins associated with centrifugally isolated HDL. Intriguingly, the plasma abundance of most of these proteins is insufficient to permit one copy per HDL particle, suggesting that specific proteins may be bound to distinct particle species, which are differentially distributed across the HDL density spectrum. Indeed, more recently, using Apoa1-deficient (*Apoa1*$^{-/-}$) and Abca1-deficient (*abca1*$^{-/-}$) mice, we showed that in addition to Apoa1, other

Received: April 25, 2016
Revised: June 22, 2016
Published: June 22, 2016

E. Zvintzou, M. Lhomme, S. Chasapi, S. Filou, V. Theodoropoulos, E. Xapapadaki, A. Kontush, G. Spyroulias, C. C. Tellis, A. D. Tselepis, C. Constantinou and K. E. Kypreos, "**Pleiotropic effects of apolipoprotein C3 on HDL functionality and adipose tissue metabolic activity,**" Journal of Lipid Research, 2017.

S. Filou, M. Lhomme, E. A. Karavia, C. Kalogeropoulou, V. Theodoropoulos, E. Zvintzou, G. C. Sakellaropoulos, P. I. Petropoulou, C. Constantinou, A. Kontush and K. E. Kypreos, "**Distinct Roles of Apolipoproteins A1 and E in the Modulation of High-Density Lipoprotein Composition and Function**," Biochemistry, 2016

DATA

Categories

- Lipid Classes + 3 Functions
- Lipid Species + 3 Functions
- Polyunsaturated Fatty Acids + 3 Functions
- Saturated and Monounsaturated Fatty Acids + 3 Functions
- Polyunsaturated Fatty Acids and Phospholipids + Antioxidant Function

Groups

- AdGFP
- AdGFP-APOC3
- AdGFP-APOE3
- AdGFP-APOA1

Lipid Subclasses

Lipids

L(yso)PC	Lysophosphatidylcholines
L(yso)PE	Lysophosphatidylethanolamines
PA	Phosphatidic Acids
PC	Phosphatidylcholines
PE	Phosphatidylethanolamines
PG	Phosphatidylglycerols
PI	Phosphatidylinositols
PS	Phosphatidylserines
Cer	Ceramides

Lipids

PL	Phospholipids
SM	Sphingomyelins

miscellaneous

CE	Cholesteryl Ester
FC	Free Cholesterol
TG	Triglycerides

FRAMEWORK

4 Stages

Data Cleaning | Profiling & Exploratory Analysis | Random Forests & Variable Importance | Partial Dependence Plots

DATA CLEANING

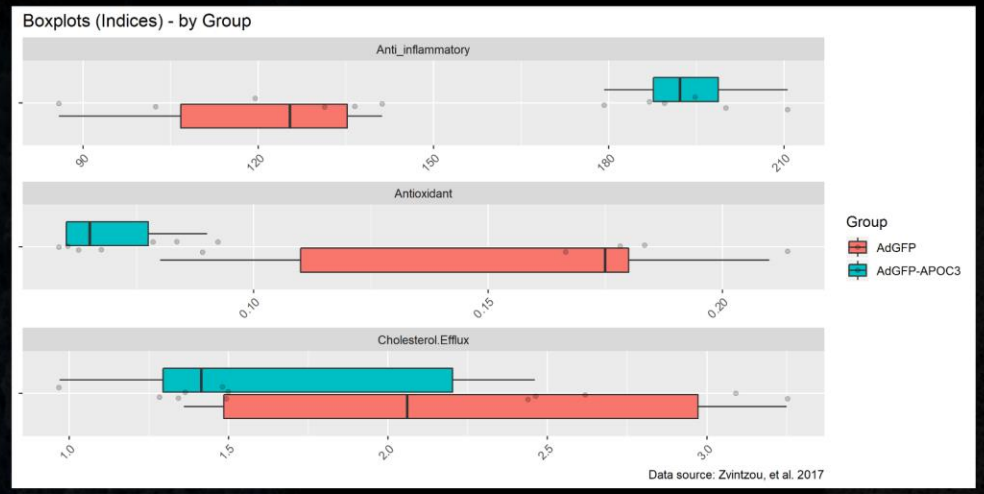
To avoid problems of low or zero variance, the following tests were performed removing each variable met one of the following conditions:

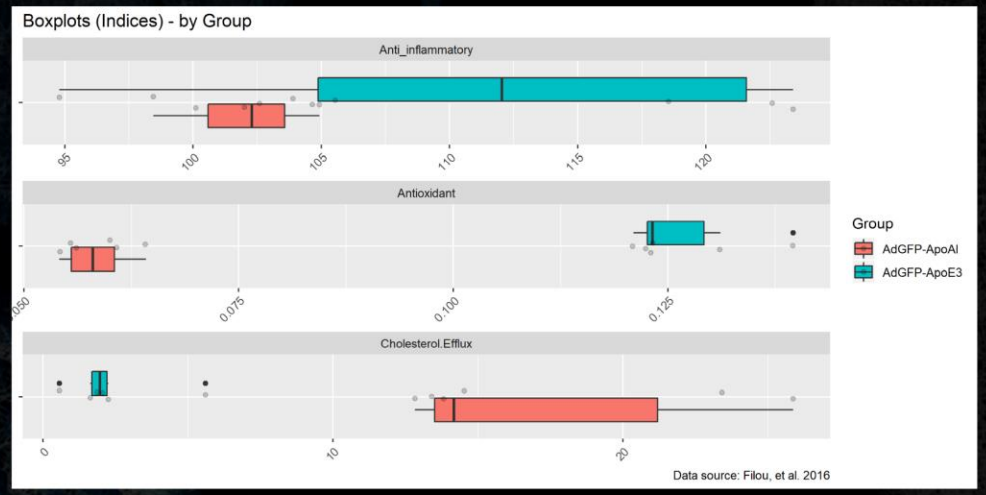
- Variables with a unique value across each sample (zero variance)
- Variables that belong to one of the following two cases
  - have very few unique values relative to the number of samples (at least 3 out of 6)
  - the ratio of the frequency of the most frequently occurring value to the frequency of the second most frequently occurring value is large (>= 0.8)

PROFILING & EXPLORATORY ANALYSIS

Treatment of variables that contained more than 6 observations

z-score scaling was applied in order to include additional information in the analysis but no significant effect was observed





Basic descriptives
grouped by AdGFP

var	*n*	*mean*	*sd*	*range*	*skew*
Cholesterol.Efflux	6	2.22	0.87	1.89 (1.36-3.25)	0.2
Antioxidant	6	0.15	0.05	0.13 (0.08-0.21)	-0.69
Anti_inflammatory	6	119.5	21.55	55.33 (85.9-141.23)	-0.78

Basic descriptives
grouped by AdGFP-ApoAI

var	*n*	*mean*	*sd*	*range*	*skew*
Cholesterol.Efflux	6	17.31	5.76	13.04 (12.83-25.87)	1.01
Antioxidant	6	0.06	0	0.01 (0.05-0.06)	0.47
Anti_inflammatory	6	102	2.39	6.48 (98.45-104.93)	-0.41

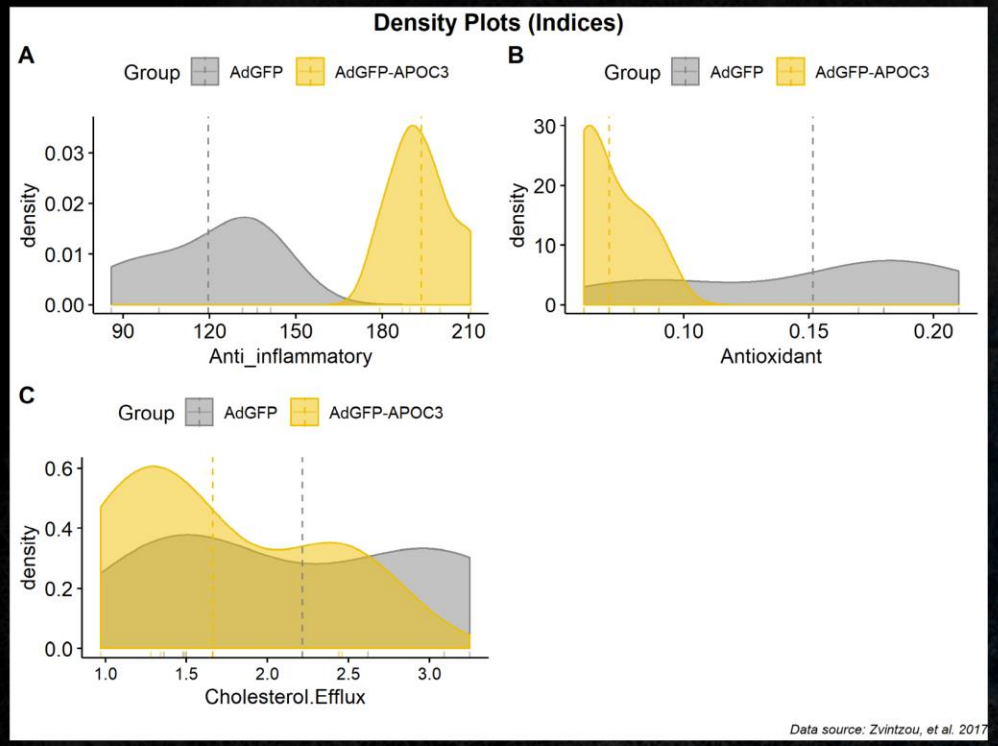
Basic descriptives
grouped by AdGFP-APOC3

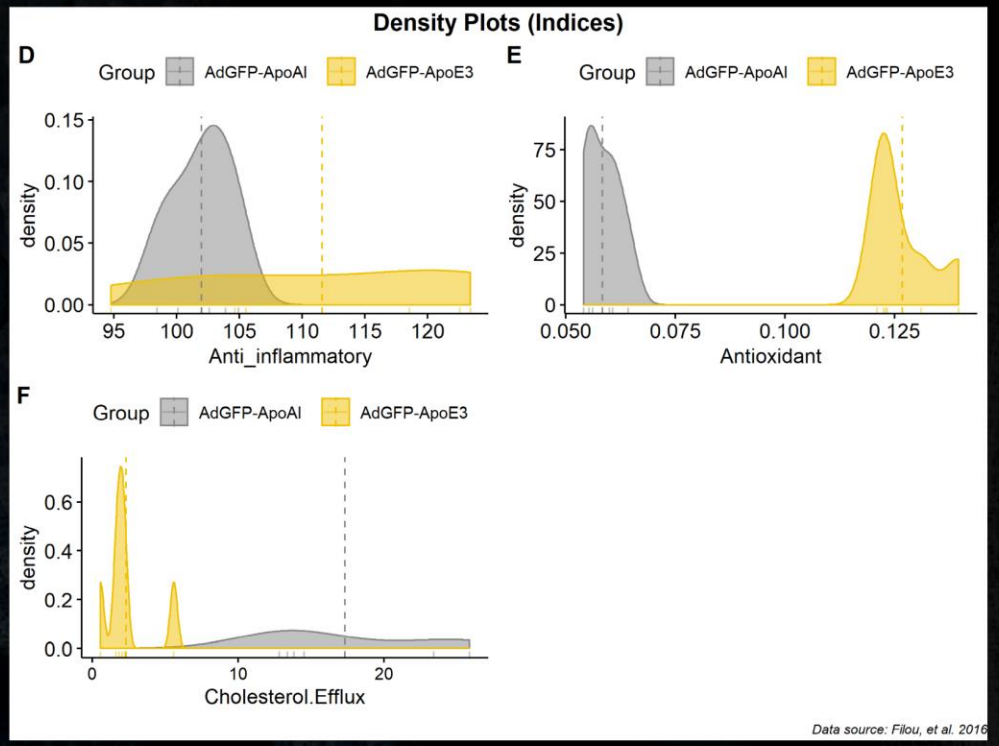
var	*n*	*mean*	*sd*	*range*	*skew*
Cholesterol.Efflux	6	1.66	0.63	1.49 (0.97-2.46)	0.65
Antioxidant	6	0.07	0.01	0.03 (0.06-0.09)	0.89
Anti_inflammatory	6	193.52	10.96	31.43 (179.19-210.62)	0.45

Basic descriptives
grouped by AdGFP-ApoE3

var	*n*	*mean*	*sd*	*range*	*skew*
Cholesterol.Efflux	6	2.32	1.71	5.04 (0.56-5.6)	1.75
Antioxidant	6	0.13	0.01	0.02 (0.12-0.14)	1.46
Anti_inflammatory	6	111.59	11.63	28.62 (94.78-123.41)	-0.38

PROFILING & EXPLORATORY ANALYSIS

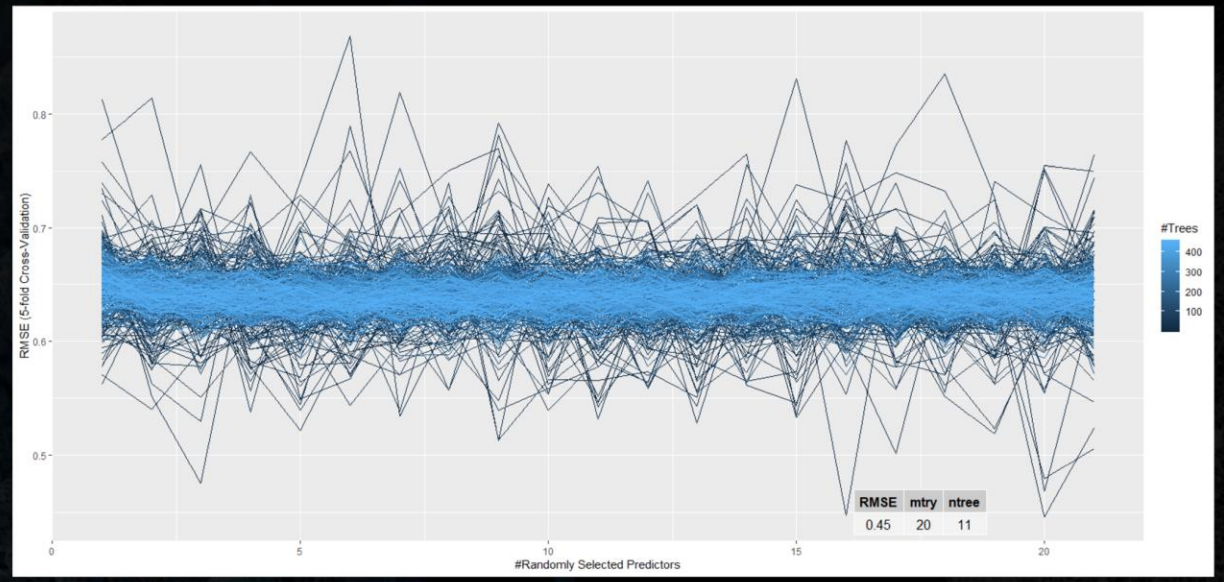


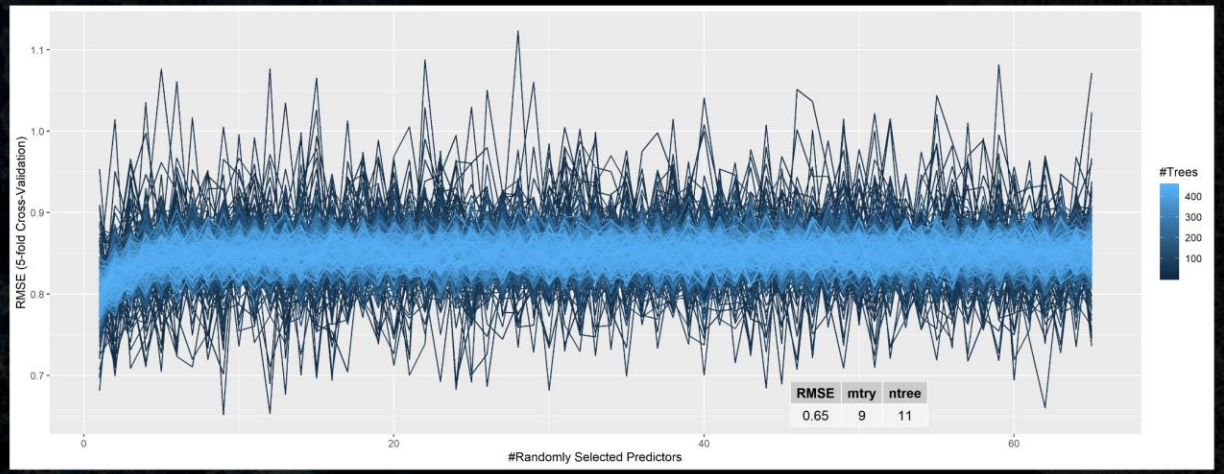


RANDOM FORESTS & VARIABLE IMPORTANCE

Due to the small sample size, choosing an appropriate number of folds for cross-validation turned out to be a challenge

Models created using 5 or 4-fold cross-validation, were ovefitted (r-squared=1) while ones created using other values were severely underperforming with high RMSEs

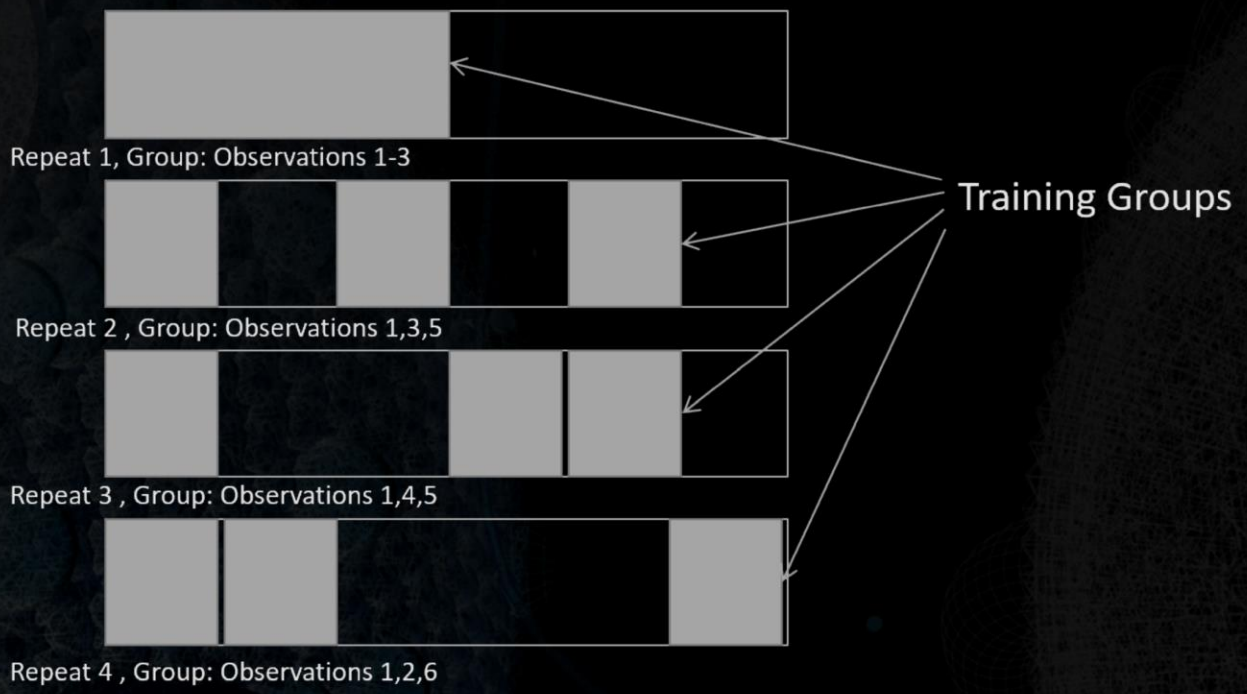




RANDOM FORESTS & VARIABLE IMPORTANCE

Leave-Group-Out cross-validation was used with 2-folds (groups) repeated 10 times.

Randomly selected (without replacement) fractions of data are chosen to form the training set while the remaining ones are used for testing. This process is then repeated multiple times, generating (at random) new training and test partitions each time. Since the partitions are done independently for each run, the same point can appear in the test set multiple times.



RANDOM FORESTS & VARIABLE IMPORTANCE

One of the most robust metrics for measuring importance for numerical variables is the *Percent Increase in Mean Square Error* (%IncMSE) which is computed by permuting each variable present in the model:

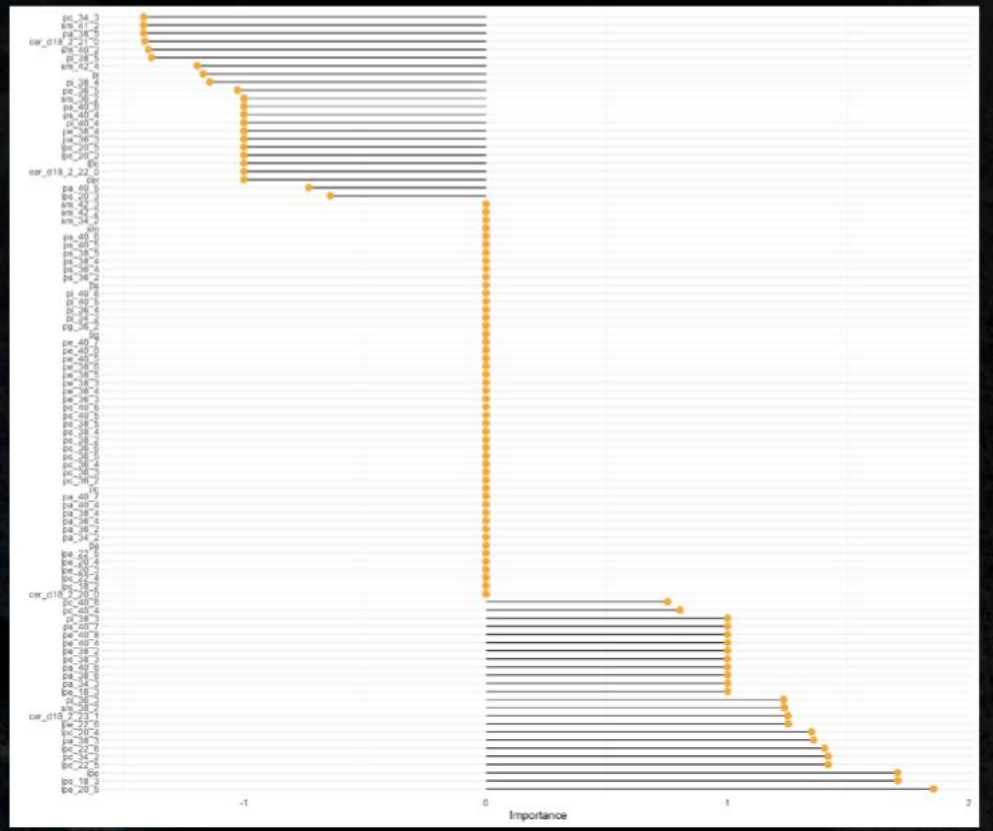
a. the model's MSE is calculated
b. for each model-specific variable, the following take place:
    i. Permutation
    ii. Computation of new model MSE giving variable permutation
    iii. The difference between initial MSE and MSE based on permutation is recorded after being scaled and multiplied by 100
c. List every difference from the previous step for each variable
d. Rank variables by importance based on Percent Increase in Mean Squared Error (higher is better)

RANDOM FORESTS & VARIABLE IMPORTANCE

Percent Increase in Mean Square Error (%IncMSE)

AdGFP, PUFA_PL, Antioxidant Function, Zvintzou et al

An additional step for variable selection in order to make it more stringent and unbiased entails the filtering of variables that do not pass a certain threshold. Values indicating this threshold were calculated by taking the absolute value of the highest negative ranked variable (per %IncMSE) and then select only those positive ranked variables that pass the respective threshold.

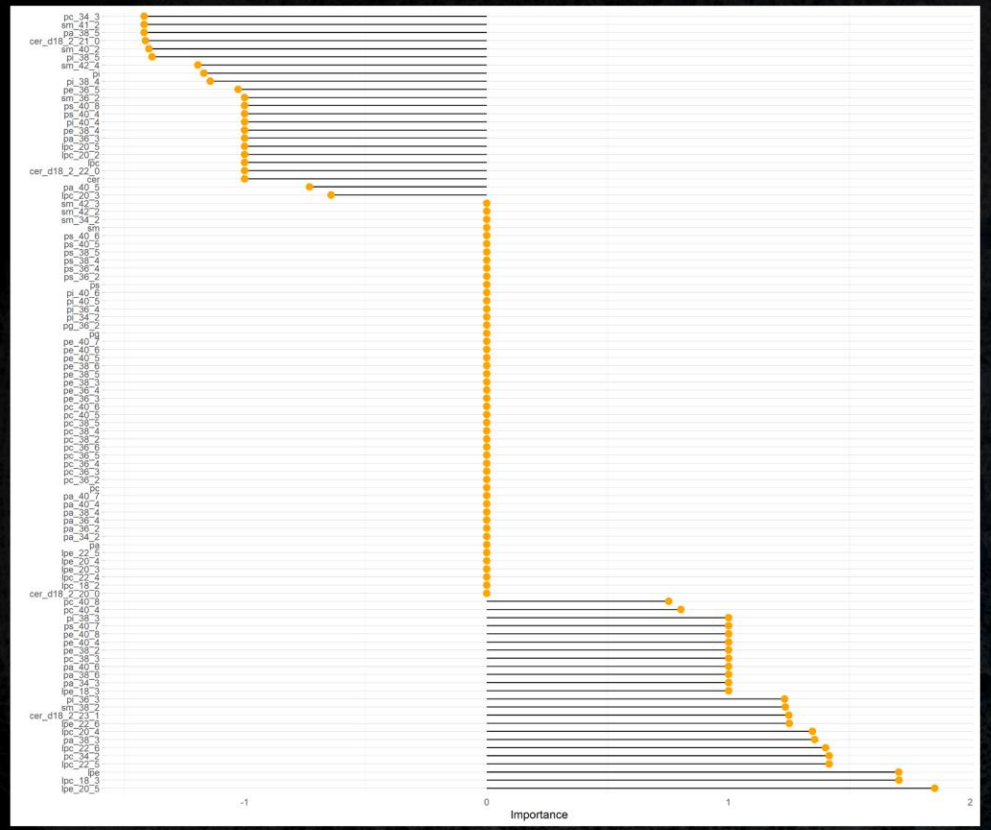


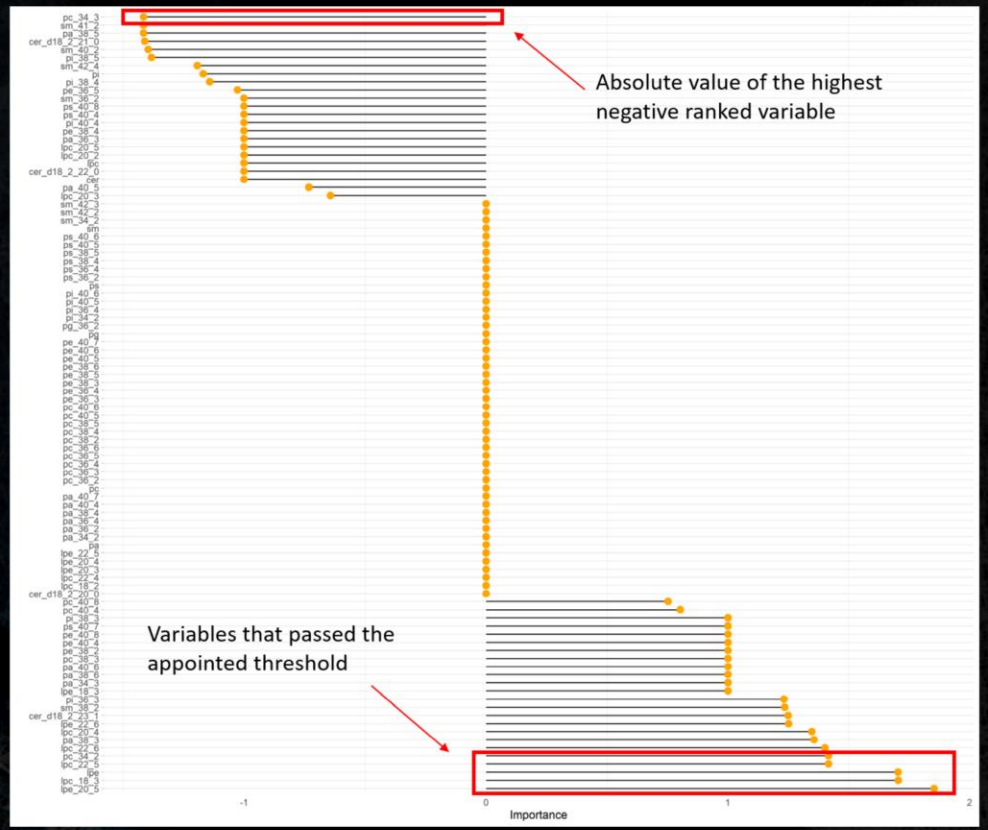
C. Strobl, J. Malley and G. Tutz, "An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests," Psychological Methods, vol. 4, no. 14, pp. 323-348, 2009.

RANDOM FORESTS & VARIABLE IMPORTANCE

Percent Increase in Mean Square Error (%IncMSE)

AdGFP, PUFA_PL, Antioxidant Function, Zvintzou et al





RANDOM FORESTS & VARIABLE IMPORTANCE

Variable	Subclass	Group	Target	RMSE	Rsquared	MAE	Overall	Source	mtry	Category
PI(34:3)	Phosphatidylinositol	AdGFP-ApoE3	Anti_inflammatory	11.3457	0.51085	10.2137	1	filou	196	LipidSpecies
PI(36:2)	Phosphatidylinositol	AdGFP-ApoE3	Anti_inflammatory	11.3457	0.51085	10.2137	2	filou	196	LipidSpecies
PI(34:2)	Phosphatidylinositol	AdGFP-ApoE3	Anti_inflammatory	11.3457	0.51085	10.2137	3	filou	196	LipidSpecies
PG(34:1)	Phosphatidylglycerol	AdGFP-ApoE3	Anti_inflammatory	11.3457	0.51085	10.2137	4	filou	196	LipidSpecies
PE(34:1)	Phosphatidylethanolamine	AdGFP-ApoAI	Anti_inflammatory	2.0579	0.69792	1.77628	1	filou	182	LipidSpecies
Cer(d18:2-16:0)	Ceramides d18:2	AdGFP-ApoAI	Anti_inflammatory	2.0579	0.69792	1.77628	2	filou	182	LipidSpecies
SM(36:2)	Sphingomyelin	AdGFP-ApoAI	Anti_inflammatory	2.0579	0.69792	1.77628	3	filou	182	LipidSpecies
PC(40:8)	Phosphatidylcholine	AdGFP-ApoAI	Anti_inflammatory	2.0579	0.69792	1.77628	4	filou	182	LipidSpecies
Cer(d18:0-16:0)	Ceramides d18:0	AdGFP-ApoAI	Anti_inflammatory	2.0579	0.69792	1.77628	5	filou	182	LipidSpecies
Cer(d18:2-24:1)	Ceramides d18:2	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	1	filou	94	LipidSpecies
PI(40:7)	Phosphatidylinositol	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	2	filou	94	LipidSpecies
PI(40:5)	Phosphatidylinositol	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	3	filou	94	LipidSpecies
PC(28:0)	Phosphatidylcholine	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	4	filou	94	LipidSpecies
PG(30:0)	Phosphatidylglycerol	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	5	filou	94	LipidSpecies
PI(38:4)	Phosphatidylinositol	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	6	filou	94	LipidSpecies
PI(36:3)	Phosphatidylinositol	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	7	filou	94	LipidSpecies
Cer(d18:2-15:0)	Ceramides d18:2	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	8	filou	94	LipidSpecies
PI(38:5)	Phosphatidylinositol	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	9	filou	94	LipidSpecies
PG(32:1)	Phosphatidylglycerol	AdGFP-ApoE3	Antioxidant	0.00643	0.45782	0.00542	10	filou	94	LipidSpecies
SM(41:2)	Sphingomyelin	AdGFP-ApoAI	Antioxidant	0.0037	0.47456	0.00315	1	filou	16	LipidSpecies
PI(36:3)	Phosphatidylinositol	AdGFP-ApoE3	Anti_inflammatory	11.2913	0.49495	10.1549	1	filou	18	PUFA
PI(40:4)	Phosphatidylinositol	AdGFP-ApoE3	Anti_inflammatory	11.2913	0.49495	10.1549	2	filou	18	PUFA
PC(40:5)	Phosphatidylcholine	AdGFP-ApoE3	Anti_inflammatory	11.2913	0.49495	10.1549	3	filou	18	PUFA
Cer(d18:2-24:1)	Ceramides d18:2	AdGFP-ApoE3	Anti_inflammatory	11.2913	0.49495	10.1549	4	filou	18	PUFA
PI(40:7)	Phosphatidylinositol	AdGFP-ApoE3	Anti_inflammatory	11.2913	0.49495	10.1549	5	filou	18	PUFA
PC(38:2)	Phosphatidylcholine	AdGFP-ApoE3	Anti_inflammatory	11.2913	0.49495	10.1549	6	filou	18	PUFA
PI(40:7)	Phosphatidylinositol	AdGFP-ApoAI	Anti_inflammatory	2.07209	0.64958	1.78168	1	filou	50	PUFA
PI(38:4)	Phosphatidylinositol	AdGFP-ApoAI	Anti_inflammatory	2.07209	0.64958	1.78168	2	filou	50	PUFA
PS(40:4)	Phosphatidylserine	AdGFP-ApoAI	Anti_inflammatory	2.07209	0.64958	1.78168	3	filou	50	PUFA
PA(38:2)	Phosphatidic Acid	AdGFP-ApoAI	Anti_inflammatory	2.07209	0.64958	1.78168	4	filou	50	PUFA
PC(36:2)	Phosphatidylcholine	AdGFP-ApoE3	Antioxidant	0.00636	0.40719	0.00533	1	filou	93	PUFA
PA(36:4)	Phosphatidic Acid	AdGFP-ApoE3	Antioxidant	0.00636	0.40719	0.00533	2	filou	93	PUFA
PI(38:4)	Phosphatidylinositol	AdGFP-ApoE3	Antioxidant	0.00636	0.40719	0.00533	3	filou	93	PUFA
PS(40:7)	Phosphatidylserine	AdGFP-ApoE3	Antioxidant	0.00636	0.40719	0.00533	4	filou	93	PUFA
PE(34:2)	Phosphatidylethanolamine	AdGFP-ApoAI	Antioxidant	0.00366	0.43685	0.00309	1	filou	2	PUFA
SM(34:2)	Sphingomyelin	AdGFP-ApoAI	Antioxidant	0.00366	0.43685	0.00309	2	filou	2	PUFA
PE(36:4)	Phosphatidylethanolamine	AdGFP-ApoAI	Antioxidant	0.00366	0.43685	0.00309	3	filou	2	PUFA
PA(40:5)	Phosphatidic Acid	AdGFP-ApoE3	Cholesterol.Efflux	1.55787	0.45596	1.21006	1	filou	82	PUFA
Cer(d18:2-16:0)	Ceramides d18:2	AdGFP-ApoE3	Cholesterol.Efflux	1.55787	0.45596	1.21006	2	filou	82	PUFA
PI(38:3)	Phosphatidylinositol	AdGFP-ApoE3	Cholesterol.Efflux	1.55787	0.45596	1.21006	3	filou	82	PUFA
PE(40:5)	Phosphatidylethanolamine	AdGFP-ApoE3	Cholesterol.Efflux	1.55787	0.45596	1.21006	4	filou	82	PUFA
PE(38:4)	Phosphatidylethanolamine	AdGFP-ApoE3	Cholesterol.Efflux	1.55787	0.45596	1.21006	5	filou	82	PUFA

Antioxidant Function

	RMSE	R-SQUARED	MAE
ADGFP	0.0522	0.450	0.0447
ADGFP-APOA1	0.00366	0.440	0.00309
ADGFP-APOC3	0.0140	0.604	0.0123
ADGFP-APOE3	0.00621	0.451	0.00524

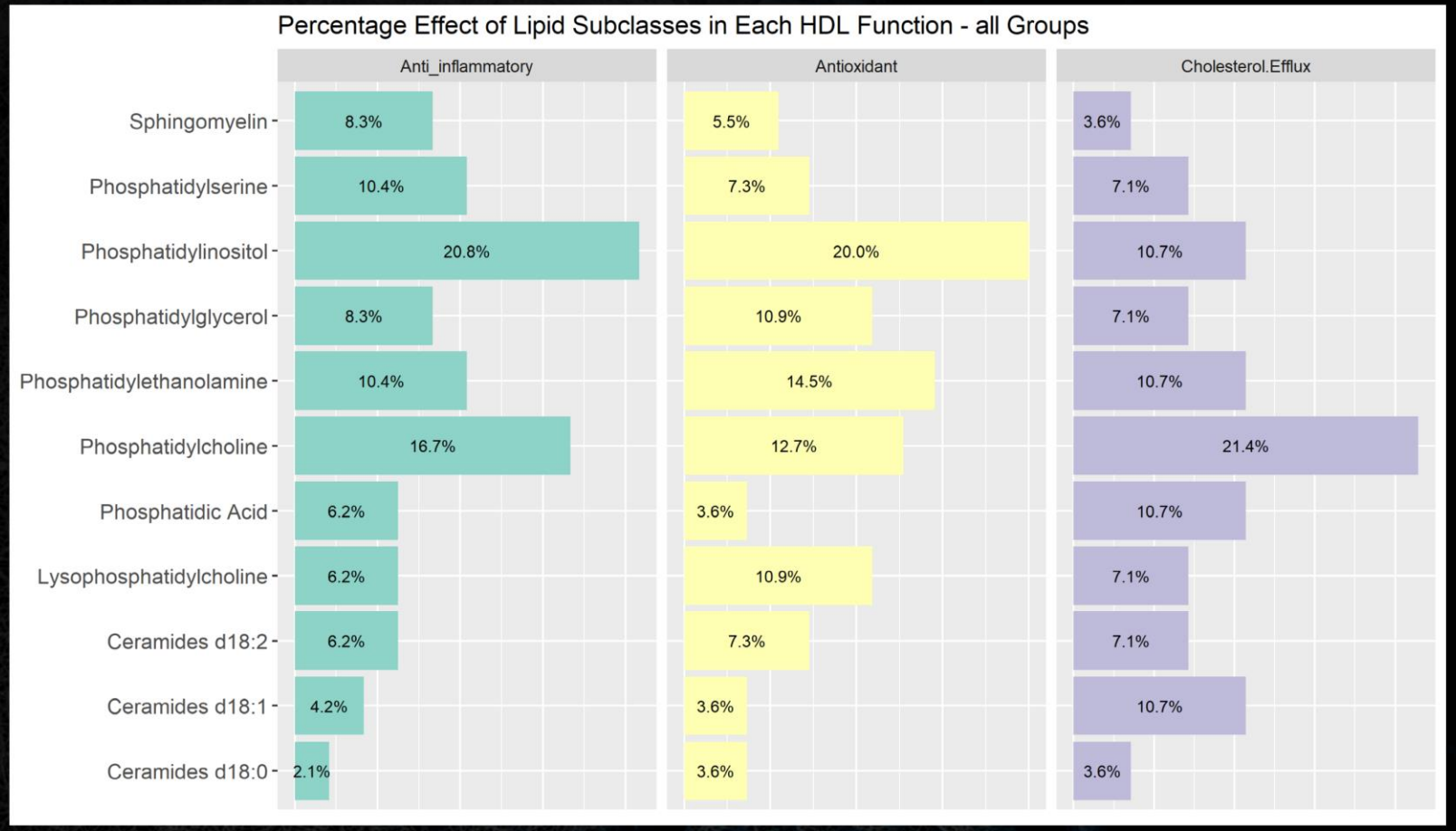
Anti-inflammatory Function

	RMSE	R-SQUARED	MAE
ADGFP	19.7	0.569	17.6
ADGFP-APOA1	2.06	0.644	1.77
ADGFP-APOC3	10.4	0.553	8.91
ADGFP-APOE3	11.3	0.508	10.2

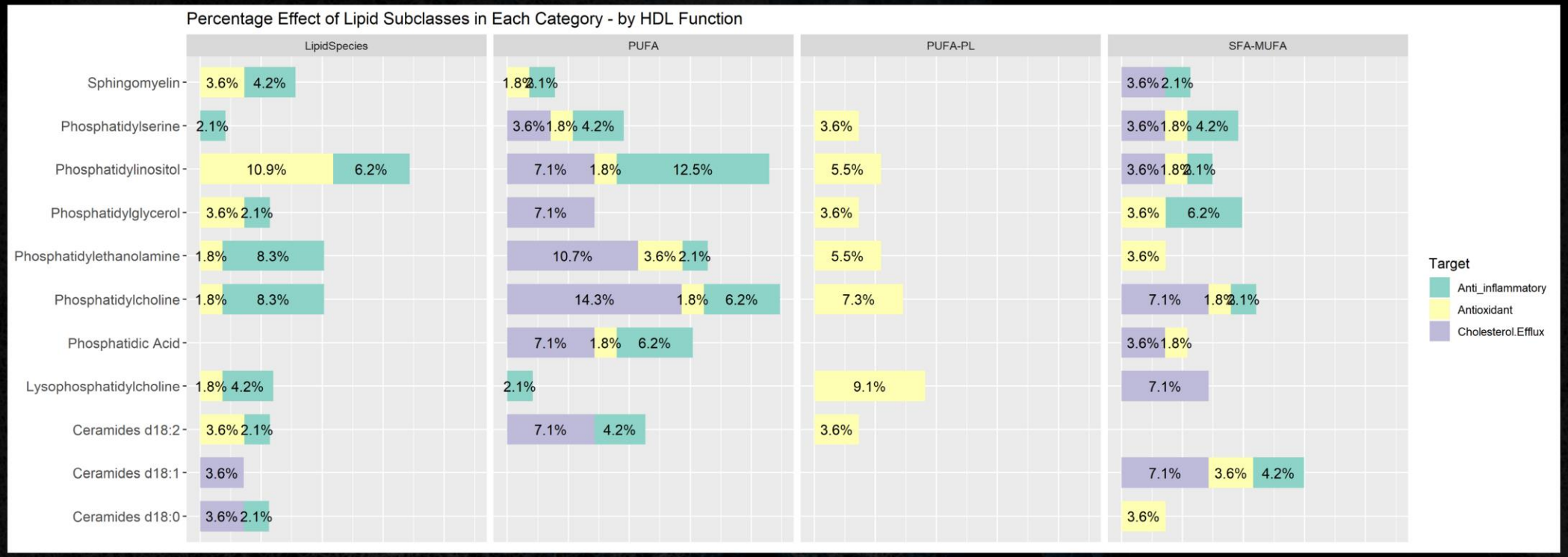
Cholesterol Efflux

	RMSE	R-SQUARED	MAE
ADGFP	0.907	0.497	0.831
ADGFP-APOA1	6.55	0.579	5.70
ADGFP-APOC3	0.655	0.528	0.560
ADGFP-APOE3	1.89	0.447	1.49

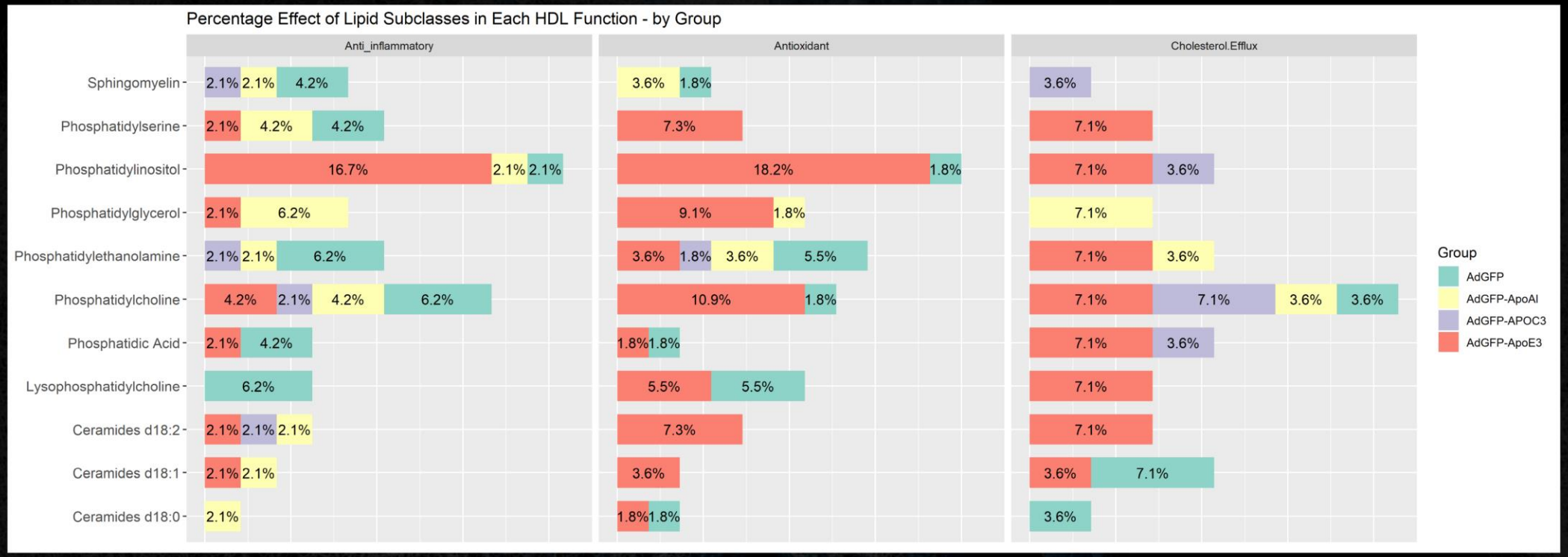
RANDOM FORESTS & VARIABLE IMPORTANCE



RANDOM FORESTS & VARIABLE IMPORTANCE



RANDOM FORESTS & VARIABLE IMPORTANCE



PARTIAL DEPENDENCE PLOTS

Partial dependence plots are a way to understand the marginal effect of a variable on the response signifying those ranges where an HDL function is either low or high.

Common variables found:
PI(38:4), PC(36:2), Cer(d18:2-24:1), PS(40:7), PS(40:5), Cer(d18:2-15:0)

	pi_38_4	lpc_16_0	lpc_16_1	lpc_16_2	lpc_18_0	Antioxidant
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	1.35	7.13	0.0934	0.00156	4.08	0.123
2	1.83	6.60	0.0913	0.00151	3.63	0.140
3	1.56	6.88	0.108	0.00108	3.65	0.123
4	1.58	6.87	0.0975	0.00138	3.79	0.121
5	1.72	7.02	0.103	0.00154	3.94	0.131
6	1.44	6.71	0.0923	0.00123	3.64	0.122

	pi_38_4	lpc_16_0	lpc_16_1	lpc_16_2	lpc_18_0	Antioxidant
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	1.35	7.13	0.0934	0.00156	4.08	0.123
2	1.35	6.60	0.0913	0.00151	3.63	0.140
3	1.35	6.88	0.108	0.00108	3.65	0.123
4	1.35	6.87	0.0975	0.00138	3.79	0.121
5	1.35	7.02	0.103	0.00154	3.94	0.131
6	1.35	6.71	0.0923	0.00123	3.64	0.122
7	1.83	7.13	0.0934	0.00156	4.08	0.123
8	1.83	6.60	0.0913	0.00151	3.63	0.140
9	1.83	6.88	0.108	0.00108	3.65	0.123
10	1.83	6.87	0.0975	0.00138	3.79	0.121

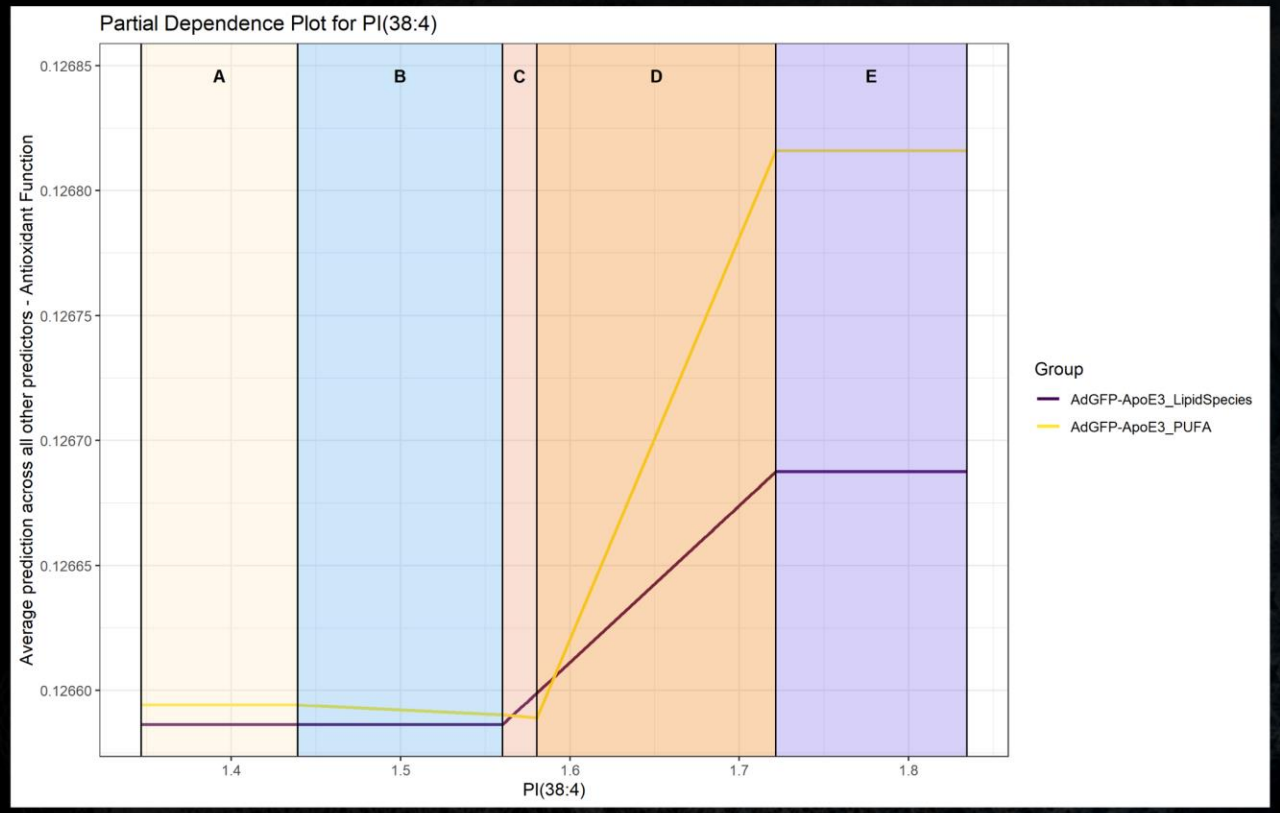
... with 26 more rows

	pi_38_4	lpc_16_0	lpc_16_1	lpc_16_2	lpc_18_0	Antioxidant	.fitted
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	1.35	7.13	0.0934	0.00156	4.08	0.123	0.124
2	1.35	6.60	0.0913	0.00151	3.63	0.140	0.133
3	1.35	6.88	0.108	0.00108	3.65	0.123	0.124
4	1.35	6.87	0.0975	0.00138	3.79	0.121	0.123
5	1.35	7.02	0.103	0.00154	3.94	0.131	0.131
6	1.35	6.71	0.0923	0.00123	3.64	0.122	0.124
7	1.83	7.13	0.0934	0.00156	4.08	0.123	0.124
8	1.83	6.60	0.0913	0.00151	3.63	0.140	0.134
9	1.83	6.88	0.108	0.00108	3.65	0.123	0.124
10	1.83	6.87	0.0975	0.00138	3.79	0.121	0.123

... with 26 more rows

	pi_38_4	marginal_prob	class
1	1.346670	0.1265865	AdGFP-ApoE3_LipidSpecies
2	1.439274	0.1265865	AdGFP-ApoE3_LipidSpecies
3	1.560210	0.1265865	AdGFP-ApoE3_LipidSpecies
4	1.580482	0.1265990	AdGFP-ApoE3_LipidSpecies
5	1.721690	0.1266875	AdGFP-ApoE3_LipidSpecies
6	1.834567	0.1266875	AdGFP-ApoE3_LipidSpecies

PARTIAL DEPENDENCE PLOTS

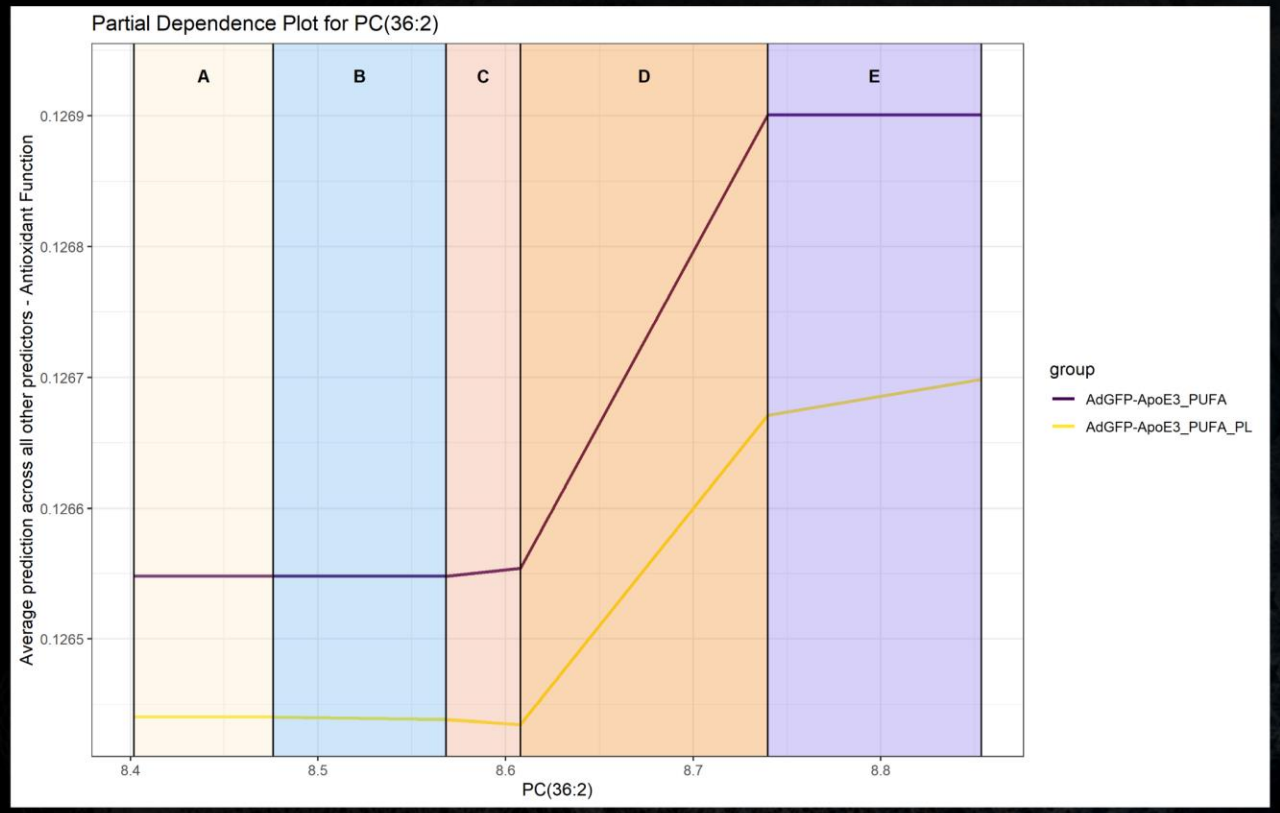


	Variable	A	B	C	D	E	RMSE	Group, Target
Slope aa	PI(38:4)	0	0	0.000617	0.000627	0	0.00643	ApoE3, Antioxidant Function
Slope bb		0	-3.31E-05	-5.92E-05	0.001608	0	0.00636	
Slope bb	PC(36:2)	0	0	0.000151	0.00263	0	0.006358592	ApoE3, Antioxidant Function
Slope dd		0	-1.95E-05	-9.58E-05	0.001792	0.000243	0.006338429	
Slope bb	PS(40:7)	0	-0.00315	0.165451	0.697345	0.435067	0.00636	ApoE3, Antioxidant Function
Slope dd		0	0	0.016011	3.015728	0	0.00634	
Slope aa	PS(40:5)	1163.151	0	-15597.5	12234	-7167.04	11.2912607	Control, Anti-infl. Function
Slope bb		812.4036	0	-38844.9	30468.32	-17156.2	10.38671931	
Slope aa	Cer(d18:2-24:1)	0	0	-0.87695	27.53457	0	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-1.89237	29.92889	2.632534	0.00634	
Slope aa	Cer(d18:2-15:0)	0	0	-26.3058	1508.966	126.4499	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-52.6116	2901.782	0	0.00634	

Interval	PI(38:4)		PC(36:2)		Cer(d18:2-24:1)	
	Start	End	Start	End	Start	End
A	1.34667	1.439274	8.401968	8.476051	0.0003632532	0.0003644271
B	1.439274	1.56021	8.476051	8.568228	0.0003644271	0.0003656905
C	1.56021	1.56021	8.568228	8.607894	0.0003656905	0.0003700237
D	1.580482	1.72169	8.607894	8.739737	0.0003700237	0.0003756203
E	1.72169	1.834567	8.739737	8.853487	0.0003756203	0.0003811283
Interval	**PS(40:7)**		**PS(40:5)**		**Cer(d18:2-15:0)**	
	Start	End	Start	End	Start	End
A	0.007209428	0.007810946	0.0004723024	0.000544434	4.419014e-07	4.647258e-07
B	0.007810946	0.008414396	0.000544434	0.0006034805	4.647258e-07	4.894282e-07
C	0.008414396	0.008957761	0.0006034805	0.0006280165	4.894282e-07	5.65457e-07
D	0.008957761	0.009017846	0.0006280165	0.0006592982	5.65457e-07	6.661882e-07
E	0.009017846	0.009075998	0.0006592982	0.0007058585	6.661882e-07	7.650416e-07

aa = LipidSpecies, bb = PUFA, cc = SFA_MUFA, dd = PUFA_PL

PARTIAL DEPENDENCE PLOTS

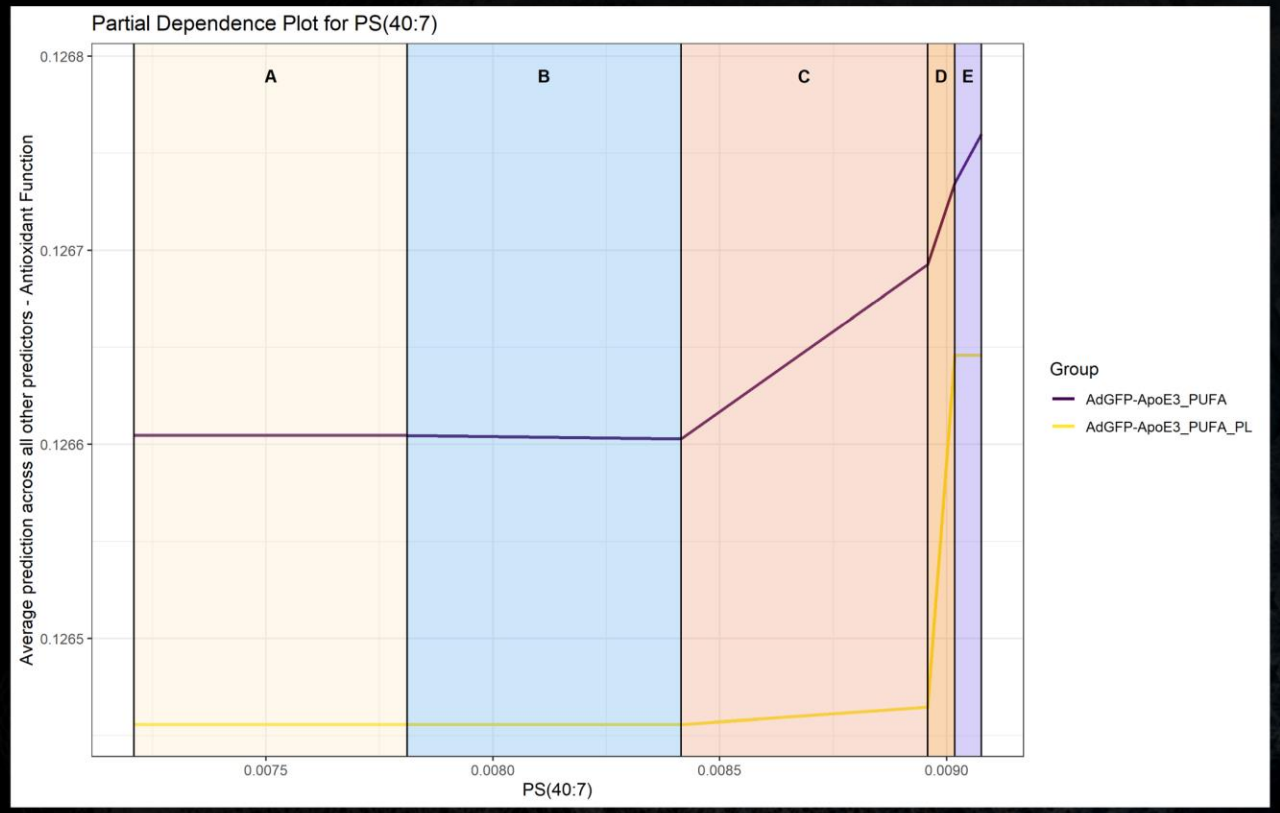


	Variable	A	B	C	D	E	RMSE	Group, Target
Slope aa	PI(38:4)	0	0	0.000617	0.000627	0	0.00643	ApoE3, Antioxidant Function
Slope bb		0	-3.31E-05	-5.92E-05	0.001608	0	0.00636	
Slope bb	PC(36:2)	0	0	0.000151	0.00263	0	0.006358592	ApoE3, Antioxidant Function
Slope dd		0	-1.95E-05	-9.58E-05	0.001792	0.000243	0.006338429	
Slope bb	PS(40:7)	0	-0.00315	0.165451	0.697345	0.435067	0.00636	ApoE3, Antioxidant Function
Slope dd		0	0	0.016011	3.015728	0	0.00634	
Slope aa	PS(40:5)	1163.151	0	-15597.5	12234	-7167.04	11.2912607	Control, Anti-infl. Function
Slope bb		812.4036	0	-38844.9	30468.32	-17156.2	10.38671931	
Slope aa	Cer(d18:2-24:1)	0	0	-0.87695	27.53457	0	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-1.89237	29.92889	2.632534	0.00634	
Slope aa	Cer(d18:2-15:0)	0	0	-26.3058	1508.966	126.4499	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-52.6116	2901.782	0	0.00634	

Interval	PI(38:4)		PC(36:2)		Cer(d18:2-24:1)	
	Start	End	Start	End	Start	End
A	1.34667	1.439274	8.401968	8.476051	0.0003632532	0.0003644271
B	1.439274	1.56021	8.476051	8.568228	0.0003644271	0.0003656905
C	1.56021	1.56021	8.568228	8.607894	0.0003656905	0.0003700237
D	1.580482	1.72169	8.607894	8.739737	0.0003700237	0.0003756203
E	1.72169	1.834567	8.739737	8.853487	0.0003756203	0.0003811283
Interval	**PS(40:7)**		**PS(40:5)**		**Cer(d18:2-15:0)**	
	Start	End	Start	End	Start	End
A	0.007209428	0.007810946	0.0004723024	0.000544434	4.419014e-07	4.647258e-07
B	0.007810946	0.008414396	0.000544434	0.0006034805	4.647258e-07	4.894282e-07
C	0.008414396	0.008957761	0.0006034805	0.0006280165	4.894282e-07	5.65457e-07
D	0.008957761	0.009017846	0.0006280165	0.0006592982	5.65457e-07	6.661882e-07
E	0.009017846	0.009075998	0.0006592982	0.0007058585	6.661882e-07	7.650416e-07

aa = LipidSpecies, bb = PUFA, cc = SFA_MUFA, dd = PUFA_PL

PARTIAL DEPENDENCE PLOTS

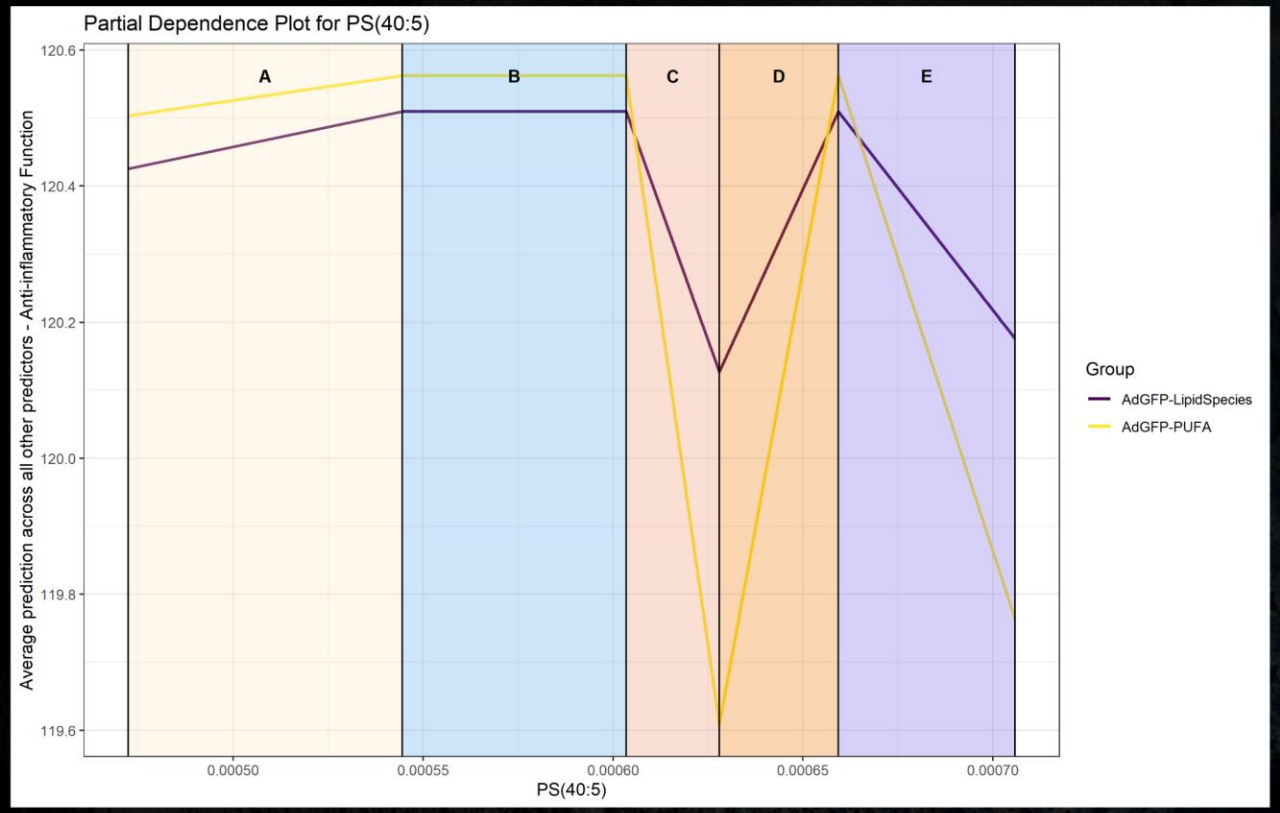


	Variable	A	B	C	D	E	RMSE	Group, Target
Slope aa	PI(38:4)	0	0	0.000617	0.000627	0	0.00643	ApoE3, Antioxidant Function
Slope bb		0	-3.31E-05	-5.92E-05	0.001608	0	0.00636	
Slope bb	PC(36:2)	0	0	0.000151	0.00263	0	0.006358592	ApoE3, Antioxidant Function
Slope dd		0	-1.95E-05	-9.58E-05	0.001792	0.000243	0.006338429	
Slope bb	PS(40:7)	0	-0.00315	0.165451	0.697345	0.435067	0.00636	ApoE3, Antioxidant Function
Slope dd		0	0	0.016011	3.015728	0	0.00634	
Slope aa	PS(40:5)	1163.151	0	-15597.5	12234	-7167.04	11.2912607	Control, Anti-infl. Function
Slope bb		812.4036	0	-38844.9	30468.32	-17156.2	10.38671931	
Slope aa	Cer(d18:2-24:1)	0	0	-0.87695	27.53457	0	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-1.89237	29.92889	2.632534	0.00634	
Slope aa	Cer(d18:2-15:0)	0	0	-26.3058	1508.966	126.4499	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-52.6116	2901.782	0	0.00634	

Interval	PI(38:4)		PC(36:2)		Cer(d18:2-24:1)	
	Start	End	Start	End	Start	End
A	1.34667	1.439274	8.401968	8.476051	0.0003632532	0.0003644271
B	1.439274	1.56021	8.476051	8.568228	0.0003644271	0.0003656905
C	1.56021	1.56021	8.568228	8.607894	0.0003656905	0.0003700237
D	1.580482	1.72169	8.607894	8.739737	0.0003700237	0.0003756203
E	1.72169	1.834567	8.739737	8.853487	0.0003756203	0.0003811283
Interval	**PS(40:7)**		**PS(40:5)**		**Cer(d18:2-15:0)**	
	Start	End	Start	End	Start	End
A	0.007209428	0.007810946	0.0004723024	0.000544434	4.419014e-07	4.647258e-07
B	0.007810946	0.008414396	0.000544434	0.0006034805	4.647258e-07	4.894282e-07
C	0.008414396	0.008957761	0.0006034805	0.0006280165	4.894282e-07	5.65457e-07
D	0.008957761	0.009017846	0.0006280165	0.0006592982	5.65457e-07	6.661882e-07
E	0.009017846	0.009075998	0.0006592982	0.0007058585	6.661882e-07	7.650416e-07

aa = LipidSpecies, bb = PUFA, cc = SFA_MUFA, dd = PUFA_PL

PARTIAL DEPENDENCE PLOTS

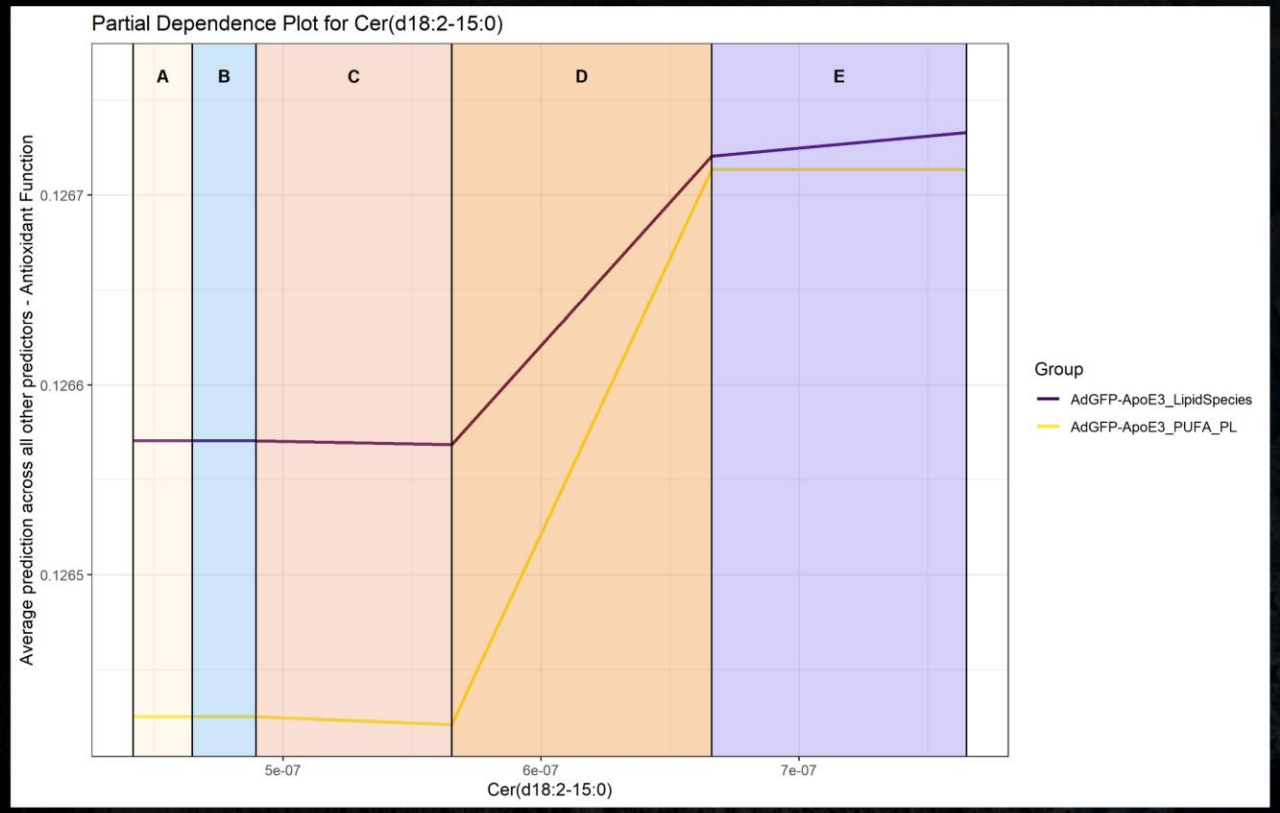


	Variable	A	B	C	D	E	RMSE	Group, Target
Slope aa	PI(38:4)	0	0	0.000617	0.000627	0	0.00643	ApoE3, Antioxidant Function
Slope bb		0	-3.31E-05	-5.92E-05	0.001608	0	0.00636	
Slope bb	PC(36:2)	0	0	0.000151	0.00263	0	0.006358592	ApoE3, Antioxidant Function
Slope dd		0	-1.95E-05	-9.58E-05	0.001792	0.000243	0.006338429	
Slope bb	PS(40:7)	0	-0.00315	0.165451	0.697345	0.435067	0.00636	ApoE3, Antioxidant Function
Slope dd		0	0	0.016011	3.015728	0	0.00634	
Slope aa	PS(40:5)	1163.151	0	-15597.5	12234	-7167.04	11.2912607	Control, Anti-infl. Function
Slope bb		812.4036	0	-38844.9	30468.32	-17156.2	10.38671931	
Slope aa	Cer(d18:2-24:1)	0	0	-0.87695	27.53457	0	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-1.89237	29.92889	2.632534	0.00634	
Slope aa	Cer(d18:2-15:0)	0	0	-26.3058	1508.966	126.4499	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-52.6116	2901.782	0	0.00634	

Interval	PI(38:4)		PC(36:2)		Cer(d18:2-24:1)	
	Start	End	Start	End	Start	End
A	1.34667	1.439274	8.401968	8.476051	0.0003632532	0.0003644271
B	1.439274	1.56021	8.476051	8.568228	0.0003644271	0.0003656905
C	1.56021	1.56021	8.568228	8.607894	0.0003656905	0.0003700237
D	1.580482	1.72169	8.607894	8.739737	0.0003700237	0.0003756203
E	1.72169	1.834567	8.739737	8.853487	0.0003756203	0.0003811283
Interval	**PS(40:7)**		**PS(40:5)**		**Cer(d18:2-15:0)**	
	Start	End	Start	End	Start	End
A	0.007209428	0.007810946	0.0004723024	0.000544434	4.419014e-07	4.647258e-07
B	0.007810946	0.008414396	0.000544434	0.0006034805	4.647258e-07	4.894282e-07
C	0.008414396	0.008957761	0.0006034805	0.0006280165	4.894282e-07	5.65457e-07
D	0.008957761	0.009017846	0.0006280165	0.0006592982	5.65457e-07	6.661882e-07
E	0.009017846	0.009075998	0.0006592982	0.0007058585	6.661882e-07	7.650416e-07

aa = LipidSpecies, bb = PUFA, cc = SFA_MUFA, dd = PUFA_PL

PARTIAL DEPENDENCE PLOTS

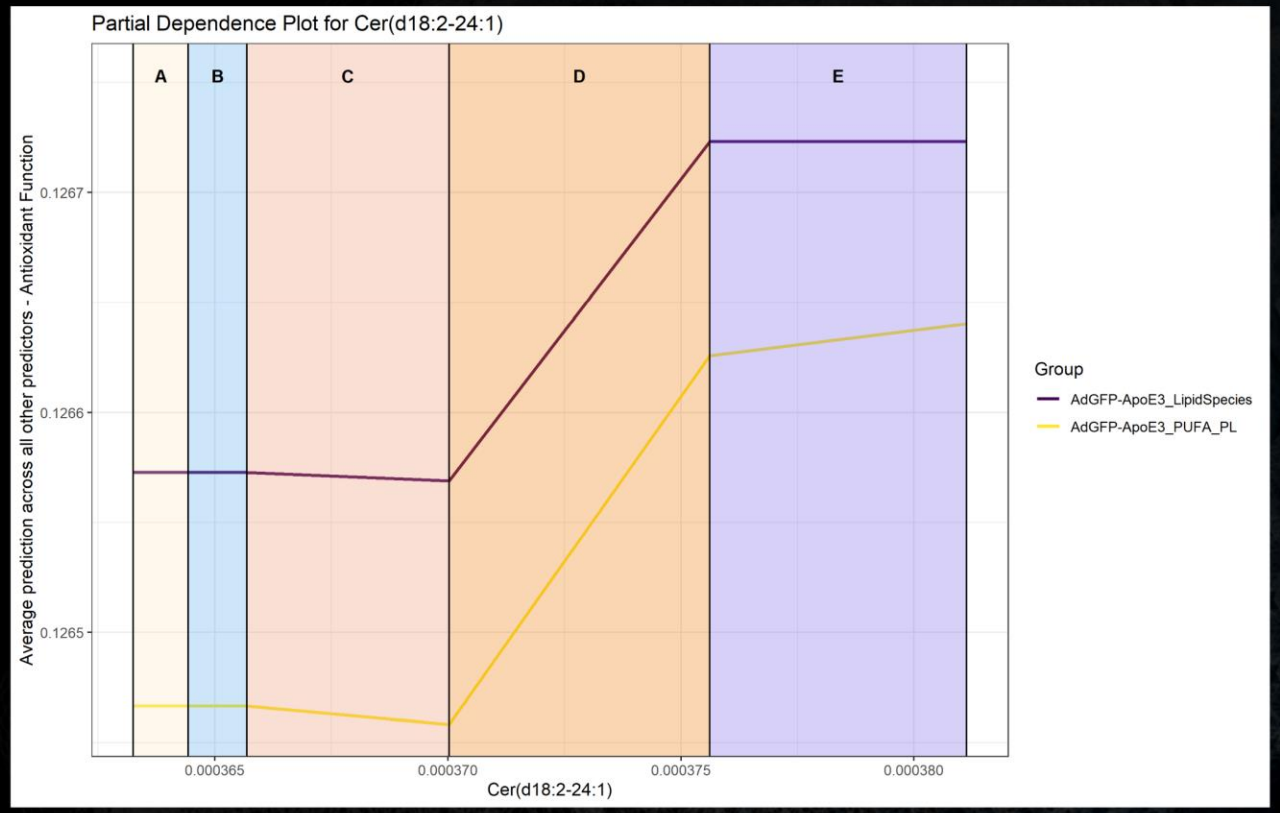


	Variable	A	B	C	D	E	RMSE	Group, Target
Slope aa	PI(38:4)	0	0	0.000617	0.000627	0	0.00643	ApoE3, Antioxidant Function
Slope bb		0	-3.31E-05	-5.92E-05	0.001608	0	0.00636	
Slope bb	PC(36:2)	0	0	0.000151	0.00263	0	0.006358592	ApoE3, Antioxidant Function
Slope dd		0	-1.95E-05	-9.58E-05	0.001792	0.000243	0.006338429	
Slope bb	PS(40:7)	0	-0.00315	0.165451	0.697345	0.435067	0.00636	ApoE3, Antioxidant Function
Slope dd		0	0	0.016011	3.015728	0	0.00634	
Slope aa	PS(40:5)	1163.151	0	-15597.5	12234	-7167.04	11.2912607	Control, Anti-infl. Function
Slope bb		812.4036	0	-38844.9	30468.32	-17156.2	10.38671931	
Slope aa	Cer(d18:2-24:1)	0	0	-0.87695	27.53457	0	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-1.89237	29.92889	2.632534	0.00634	
Slope aa	Cer(d18:2-15:0)	0	0	-26.3058	1508.966	126.4499	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-52.6116	2901.782	0	0.00634	

Interval	PI(38:4)		PC(36:2)		Cer(d18:2-24:1)	
	Start	End	Start	End	Start	End
A	1.34667	1.439274	8.401968	8.476051	0.0003632532	0.0003644271
B	1.439274	1.56021	8.476051	8.568228	0.0003644271	0.0003656905
C	1.56021	1.56021	8.568228	8.607894	0.0003656905	0.0003700237
D	1.580482	1.72169	8.607894	8.739737	0.0003700237	0.0003756203
E	1.72169	1.834567	8.739737	8.853487	0.0003756203	0.0003811283
Interval	**PS(40:7)**		**PS(40:5)**		**Cer(d18:2-15:0)**	
	Start	End	Start	End	Start	End
A	0.007209428	0.007810946	0.0004723024	0.000544434	4.419014e-07	4.647258e-07
B	0.007810946	0.008414396	0.000544434	0.0006034805	4.647258e-07	4.894282e-07
C	0.008414396	0.008957761	0.0006034805	0.0006280165	4.894282e-07	5.65457e-07
D	0.008957761	0.009017846	0.0006280165	0.0006592982	5.65457e-07	6.661882e-07
E	0.009017846	0.009075998	0.0006592982	0.0007058585	6.661882e-07	7.650416e-07

aa = LipidSpecies, bb = PUFA, cc = SFA_MUFA, dd = PUFA_PL

PARTIAL DEPENDENCE PLOTS



	Variable	A	B	C	D	E	RMSE	Group, Target
Slope aa	PI(38:4)	0	0	0.000617	0.000627	0	0.00643	ApoE3, Antioxidant Function
Slope bb		0	-3.31E-05	-5.92E-05	0.001608	0	0.00636	
Slope bb	PC(36:2)	0	0	0.000151	0.00263	0	0.006358592	ApoE3, Antioxidant Function
Slope dd		0	-1.95E-05	-9.58E-05	0.001792	0.000243	0.006338429	
Slope bb	PS(40:7)	0	-0.00315	0.165451	0.697345	0.435067	0.00636	ApoE3, Antioxidant Function
Slope dd		0	0	0.016011	3.015728	0	0.00634	
Slope aa	PS(40:5)	1163.151	0	-15597.5	12234	-7167.04	11.2912607	Control, Anti-infl. Function
Slope bb		812.4036	0	-38844.9	30468.32	-17156.2	10.38671931	
Slope aa	Cer(d18:2-24:1)	0	0	-0.87695	27.53457	0	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-1.89237	29.92889	2.632534	0.00634	
Slope aa	Cer(d18:2-15:0)	0	0	-26.3058	1508.966	126.4499	0.00643	ApoE3, Antioxidant Function
Slope dd		0	0	-52.6116	2901.782	0	0.00634	

Interval	PI(38:4)		PC(36:2)		Cer(d18:2-24:1)	
	Start	End	Start	End	Start	End
A	1.34667	1.439274	8.401968	8.476051	0.0003632532	0.0003644271
B	1.439274	1.56021	8.476051	8.568228	0.0003644271	0.0003656905
C	1.56021	1.56021	8.568228	8.607894	0.0003656905	0.0003700237
D	1.580482	1.72169	8.607894	8.739737	0.0003700237	0.0003756203
E	1.72169	1.834567	8.739737	8.853487	0.0003756203	0.0003811283
Interval	**PS(40:7)**		**PS(40:5)**		**Cer(d18:2-15:0)**	
	Start	End	Start	End	Start	End
A	0.007209428	0.007810946	0.0004723024	0.000544434	4.419014e-07	4.647258e-07
B	0.007810946	0.008414396	0.000544434	0.0006034805	4.647258e-07	4.894282e-07
C	0.008414396	0.008957761	0.0006034805	0.0006280165	4.894282e-07	5.65457e-07
D	0.008957761	0.009017846	0.0006280165	0.0006592982	5.65457e-07	6.661882e-07
E	0.009017846	0.009075998	0.0006592982	0.0007058585	6.661882e-07	7.650416e-07

aa = LipidSpecies, bb = PUFA, cc = SFA_MUFA, dd = PUFA_PL

CONCLUSIONS - FUTURE WORK

- Phosphatidylinositols were presented as a major factor with high predictive power regarding both Anti-inflammatory and Antioxidant functions (20.8% and 20% of total fatty acids respectively).
- Phosphatidylcholines have the highest participation when it comes to the explanation of Cholesterol Efflux with 21.4%, 18.7% in Anti-inflammatory function and a 12.7% of total lipid subclasses in Antioxidant function.
- Exploring common fatty acids individually, it is observed that the majority originates from the AdGFP-ApoE3 group with the antioxidant function as target indicating once more the dominant effect of apolipoprotein E3 over other apolipoproteins.

CONCLUSIONS - FUTURE WORK

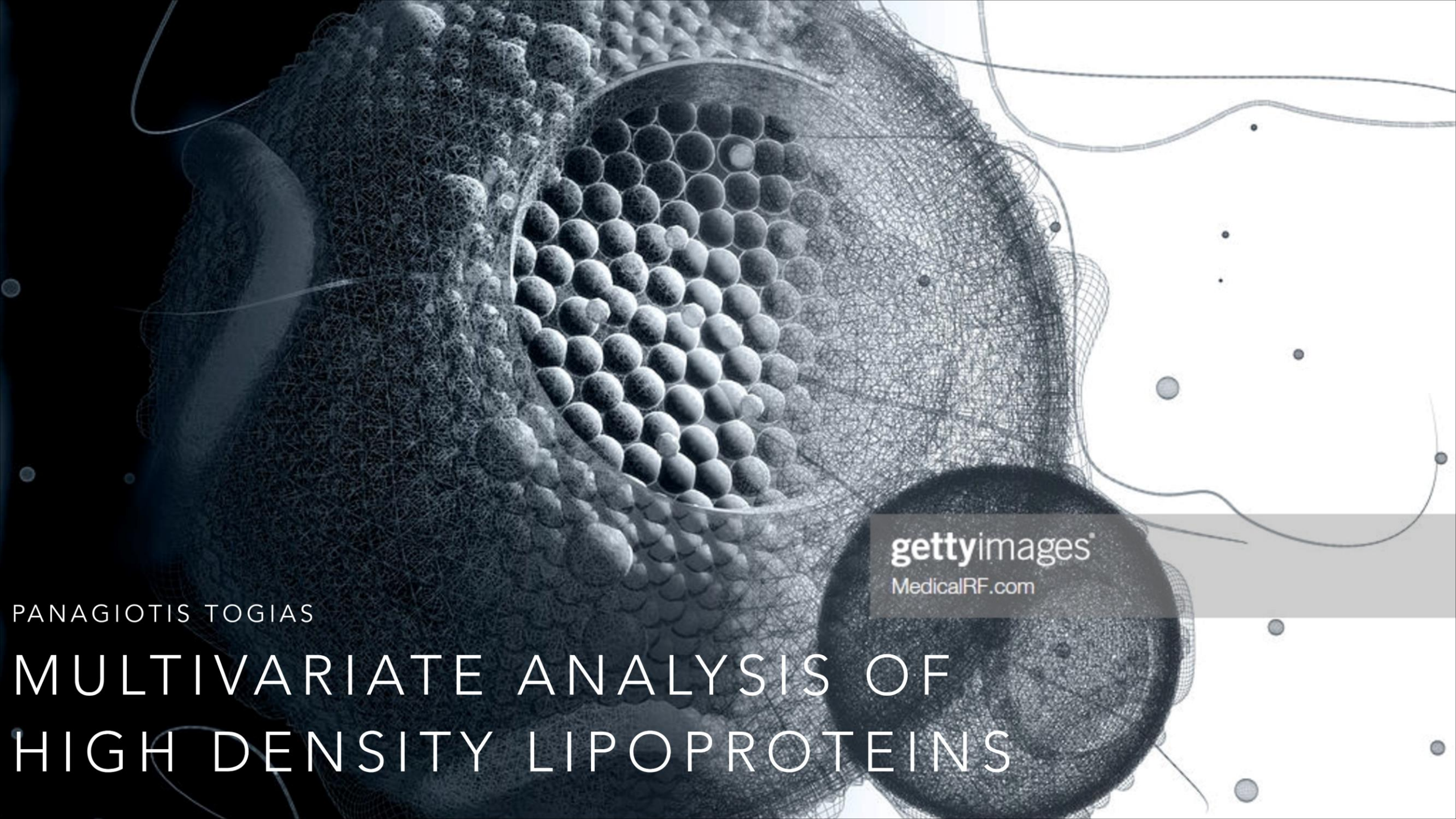
- Increase sample
- Clustering for better intervals and convert problem from regression to classification
- Test other ML methods for better classification
- Create text mining tools for knowledge extraction from databases and cross validate results
- Use of association rules to identify more underlying associations between variables and targets
- Construction of a composite index expressing all 3 functions

PANAGIOTIS TOGIAS

MULTIVARIATE DATA ANALYSIS LINKING STRUCTURAL AND FUNCTIONAL PROPERTIES OF HIGH DENSITY LIPOPROTEINS

UNIVERSITY OF PATRAS

MSC LIFE SCIENCES INFORMATICS



PANAGIOTIS TOGIAS

MULTIVARIATE ANALYSIS OF HIGH DENSITY LIPOPROTEINS