TOWARDS DETERMINATION OF NATURAL VARIATION OF BLOOD LIPIDS IN THE SINGAPOREAN POPULATION

HUSNA BEGUM

BSc (Hons), National University of Singapore

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

NUS GRADUATE SCHOOL FOR INTEGRATIVE SCIENCES AND ENGINEERING

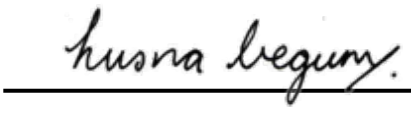
NATIONAL UNIVERSITY OF SINGAPORE

2013

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.



24th January 2013

HUSNA BEGUM

Acknowledgements

My journey throughout these 4 years of my PhD would not have been possible without the support of my supervisor, family and friends.

I would like to thank Markus for giving me constant support and sparking my passion and enthusiasm throughout the course of my project. I would like to thank my TAC, Limsoon and Steve for their useful feedback and suggestions on technical aspects of my study. I am grateful to NGS and their staff for their generous financial support and help involving academic matters.

I would like to thank Guanghou, Weifun and Aaron who gave me the opportunity to learn more on mass spectrometry, without which I could not have done this project. I would sincerely like to thank Bowen and YY who have relentlessly helped me with my statistical analysis from the time I embarked on this project. Special thanks to Phyllis who was always so cheerful and willing to help me with my sample processing.

To Federico, Lukas and Madhu, for amazing friendship and patience in helping me to put my thesis together, always encouraging me especially through these last few days of my PhD journey. I am grateful to my lab colleagues and friends, Anne, Amaury, Charmaine, Huimin, Jere, Jing

Yan, Jacklyn, Lissya, Pradeep, Shareef, Sudar, Xinrui, Xueli and Yuting, for making my life in the lab a crazy and productive one.

I am most grateful to my family, my parents and sisters who have always supported and encouraged me throughout my tough times, giving me the strength to pursue my dreams, no matter how insane they were.

Last but not least, to all my friends in Singapore and overseas, thank you for your good friendship and making my stay in Singapore a memorable one.

Table of Contents

Summary

Many unanswered questions remain on the impact of genome natural variation in healthy populations. The basis of the biological variation cannot be completely explained by genomics and proteomics. Lipidomics offers an avenue to complement these traditional technologies, studying lipid profiles across various conditions, and the recent advances in mass spectrometry complemented existing platforms in providing rapid and sensitive methods for lipids profiling.

Lipids have been found to be important in human health as their dysregulation has been reported in pathological conditions. It is, however, imperative to understand how lipids vary in normal physiological settings. As such, this study aims to explore the range of biological variation in human plasma of various lipids including glycerophospholipids, sphingolipids and sterol derivatives, in 360 healthy fasting Singaporeans. Targeted mass spectrometry using multiple reaction monitoring (MRM) was used to quantify over 100 individual lipid species. In a first step, variation that is introduced by sample pre-processing and instrumentation, i.e. technical variation, has been carefully considered and the importance of using quality control samples in study design was recognized. Issues involving the analysis of multiple batches by mass spectrometry and batch effect corrections were addressed prior to considering any biological variation. In this study, we identified

ethnicity dependent differences across lipid species, in particular significantly higher levels of PS and pPE and pPC in Chinese when compared to Indian and Malay populations. Ultimately, this work aims to quantify the biological variation of lipids in the healthy Singaporean population across three ethnicities (Chinese, Malay and Indian) and to understand the biological meaning of the underlying mechanisms.

List of Tables

Supplementary Tables

List of Figures

Supplementary Figures

List of abbreviations and keywords

a	Ester-linked
APCI	Atmospheric pressure chemical ionization
Cer	Ceramide
Chol	Cholesterol
CholE	Cholesterol ester
CID	Collision-induced dissociation
CM	Chylomicron
COV	Coefficient of variation
Di	Diunsaturated
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
FA	Fatty acyl
FT-ICR	Fourier transform ion cyclotron resonance
GC	Gas chromatography
GLs	Glycerolipid
GluCer	Glucosyl ceramide
GM3	Ganglioside mannoside 3
GPLs	Glycerophospholipid
HDL	High density lipoprotein
HPLC	High performance liquid chromatography
LC	Long carbon chain

LC-MS	Liquid chromatography mass spectrometry
LDL	Low density lipoprotien
LMW	Low molecular weight
LOD	Limit of detection
MALDI	Matrix-associated laser desorption ionization
MC	Moderate carbon chain
min	Minute
ml	Milliliter
Mono	Monounsaturated
MRM	Multiple reaction monitoring
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
OHChol	Hydroxycholesterol
p	plasmalogen
PA	Phosphatidic acid
PC	Phosphatidylcholine
PCA	Principal component analysis
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PL	Phospholipid
PS	Phosphatidylserine
PUFA	Polyunsaturated fatty acid
QC	Quality control
QCS	Quality control study

QqQ	Triple quadrupole
Q1	First quadrupole
Q2	Collision cell
Q3	Third quadrupole
Sat	Saturated
SC	Short carbon chain
SLE	Systemic lupus erythematosus
SM	Sphingomyelin
SNP	Single nucleotide polymorphism
SPs	Sphingolipid
STLs	Sterol
S/N	Signal to noise
TAG	Triacylglycerol
TLC	Thin layer chromatography
TOF	Time of flight
TV	Technical variation
TVS	Technical variation study
VLDL	Very low density lipoprotein
VS	Validation study

CHAPTER 1

INTRODUCTION

1.1 Metabolites

Metabolites are a class of biomolecules that include carbohydrates, amino acids, nucleic acids and lipids (Fahy et al., 2011). For a long time, these small molecules remained in the shadow of the ongoing 'omics' trend in biology of genomes and proteomes (Shevchenko and Simons, 2010). Unlike transcripts and proteins, the molecular identity of metabolites cannot be deduced from genomic information (Lei et al., 2011). It is the flux and biosynthesis of these low molecular weight (LMW) metabolites within a cell, tissue, or organism that generates a phenotype (Roberts et al., 2012).

1.1.1 Lipids

Among metabolites, lipids constitute the largest subset, including tens of thousands of distinct molecular species existing in cells, tissues and body fluids (Harkewicz and Dennis, 2011; Wenk, 2010). Defining a lipid is not as straightforward since there is no universally accepted definition to date. However, a lipid can be simply defined as a naturally occurring molecule

made up of fatty acids and its derivatives. Lipids are not genetically encoded, they are transient molecules that have a dynamic composition and compartmentalization, making them good markers of metabolic traits (Gross, 2006; Wenk, 2010).

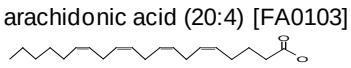
In recent years, there has been an increasing emphasis on lipid research. Lipids exist with great diversity in structure and function, and their roles are continuously being explored at a cellular, tissue and organism level (Wenk, 2010). The field of lipid research has extended from the study of prokaryotes to eukaryotes and the astonishing diversity of lipids within an organism and across different types of organisms has yet to be discovered.

1.1.1.1 Lipid chemistry

Amongst the various types of lipids, some of the major classes include fatty acyls (FA), glycerolipids (GLs), glycerophospholipids (GPLs), sphingolipids (SPs), sterol lipids (STLs) as shown in Figure 1-1 (Wenk, 2005). GPLs consist of a glycerol backbone made up by two FA chains attached at the sn-1 and sn-2 positions, and a phosphate group esterified to one of the glycerol hydroxyl groups (Hu et al., 2009). SPs contain a long-chain nitrogenous base as their core structure while STLs have a four ring core structure (Fahy et al., 2011).

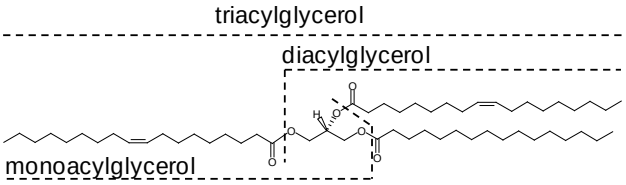
a. Fatty acyls (FA)

Fatty acids and conjugates [FA01]
Octadecanoids [FA02]
Eicosanoids [FA03]
Docosanoids [FA04]
Fatty alcohols [FA05]
Fatty aldehydes [FA06]
Fatty esters [FA07]



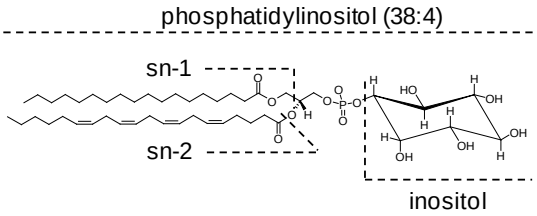
b. Glycerolipids (GL)

Monoradylglycerols (MG) [GL01]
Diradylglycerols (DG) [GL02]
Triradylglycerols (TG) [GL03]



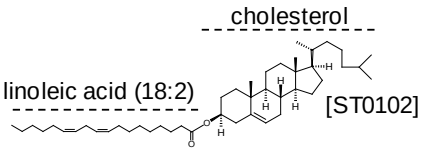
c. Glycerophospholipids (GPL)

Phosphatidic acids (PA) [GP10]
Phosphatidylcholines (PC) [GP01]
Phosphatidylserines (PS) [GP03]
Phosphatidylglycerols (PG) [GP04]
Phosphatidylethanolamines (PE) [GP02]
Phosphatidylinositols (PI) [GP06] & phosphoinositides [GP07-09]
Cardiolipins (CL) [GP12]



d. Sterol lipids (ST)

Sterols [ST01]
Steroids [ST02]
Secosteroids [ST03]
Bile acids and derivatives [ST04]



e. Sphingolipids (SP)

Sphingoid bases [SP01]
Ceramides (Cer) [SP02]
Phosphosphingolipids [SP03]
Phosphonosphingolipids [SP04]
Neutral glycosphingolipids [SP05]
Acidic glycosphingolipids [SP06]

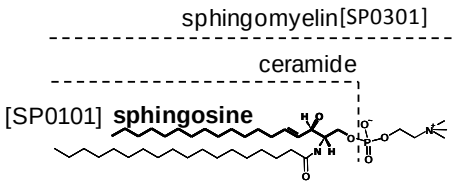


Figure 1-1: Lipid classes and their representatives. *Adapted from Wenk (Wenk, 2005).*

1.1.1.1.1 Glycerophospholipids

As the major group of lipids, GPLs consist of various combinations of different FA attached to the sn-1 and sn-2 positions of the glycerol backbone. Based on different polar head groups, GPLs can be divided into glycerophosphatidic acids (PA), glycerophosphatidylcholines (PC), glycerophosphatidylethanolamines (PE), glycerophosphatidylserines (PS), glycerophosphatidylglycerols (PG) and glycerophosphatidylinositols (PI) (Fahy et al., 2011). Lyso-GPLs which have one of the hydroxyl groups at the sn-1 or sn-2 positions of the glycerol backbone intact and the other one esterified to FA, are also included in the definition of GPLs (Hu et al., 2009).

Particularly, for some classes of GPLs (typically for PE and PC species, FA chains at the sn-1 position can also be attached as an ether-linkage including vinyl ether and alkyl ether, termed “plasmenyl-” and “plasmanyl-” lipids respectively, instead of the more common ester linkage (Hu et al., 2009). Approximately 18% of human phospholipid (PL) are ether linked lipids (Nagan et al., 1998). The choline and ethanolamine plasmalogens (plasmenylcholine and plasmenylethanolamine, respectively) contain a cis double bond adjacent to the sn-1 ether linkage, whereas the 1-alkyl-2-acyl PL plasmanylcholine and plasmanylethanolamine contain the sn-1 ether linkage but lack the cis-double bond (Vance and Tasseva, 2013).

GPLs are amphipathic and have both polar and non polar ends. The polar end is made up of the ionizable phosphate group while the non polar end is mainly made up of the attached FA chain tails. Additional structural diversity arises from the variation in carbon chain length and degree of saturation across the fatty acyl chains attached at the sn-1 and sn-2 positions.

1.1.1.1.2 Sphingolipids

SPs consist of a sphingoid base to which a FA is attached through an amide bond generating the ceramide (Cer) backbone to which a head group is attached at the primary hydroxyl (E.Vance, 2008). The head groups range from a simple hydrogen (for sphingosine and Cer) to more complex species, such as the phosphocholine moiety of sphingomyelin (SM) and the simple to complex glycans of glycosphingolipids like glucosyl ceramides (GluCer) and gangliosides (Hu et al., 2009). Typically, the amide-linked FA in Cer and SM are saturated or monounsaturated long carbon chains varying in length from 14 to 26 carbon atoms (Hu et al., 2009) but are predominantly found as 18 carbon chains in mammalian cells (Merrill, 2011). Like the GPLs, structural variation for SPs can arise from head group substitutions and differences in chain length and hydroxylation of the sphingoid base and fatty acyl chains.

1.1.1.1.3 Sterols

STLs consist of cholesterols (Chol) and derivatives [e.g. cholesterol esters (CholE)] where CholEs are derived from Chol through a long-chain FA esterification with the only hydroxyl group of Chol (Hu et al., 2009). Excess cholesterol is often converted to its oxygenated counterparts, oxycholesterols such as hydroxycholesterol (OH-Chol) (E.Vance, 2008).

1.1.1.2 Lipid Function

1.1.1.2.1 Glycerophospholipids

GPLs are found in most mammals and represent one of the most important backbone structures of cellular membranes (Hu et al., 2009). Their heterogenous distribution confers functionality to lipid membranes. GPLs participate in various biological activities including cell signaling, membrane anchoring and trafficking (E.Vance, 2008; Yorek, 1993).

1.1.1.2.2 Sphingolipids

SPs are found in almost all known plants, animals, viruses, and fungi (E.Vance, 2008). SPs are highly bioactive compounds that serve not only as components of biological structures such as membranes and lipoproteins, but also as regulators of cell proliferation, differentiation, cell–cell

and cell–matrix interactions, cell migration, intracellular (and extracellular) signaling, membrane trafficking, autophagy, and cell death (E.Vance, 2008; Merrill et al., 2005; Spiegel and Milstien, 2003).

1.1.1.2.3 Sterols

STLs are highly abundant in mammals and are important components of membrane lipids (Hu et al., 2009). One of the most eminent roles of STLs is its regulatory function related to cell signaling and membrane fluidity (Hu et al., 2009). STLs include cholesterol, cholesterol esters, hormones like estrogen and progesterone which participate in different biological processes (Tsai and O'Malley, 1994; Wenk, 2005).

1.1.1.3 Lipid Transport in the circulatory system

Lipids synthesized in liver and intestine need to be transported to specialized compartments where they play a role in essential body functions such as storage and degradation (E.Vance, 2008). Lipids are transported in circulation via soluble complexes that are made up of proteins (apolipoproteins) and lipids, known as lipoproteins (E.Vance, 2008). Human lipoproteins consist of a neutral lipid core (triacylglycerol (TAG) and CholE) surrounded by a single surface layer of polar lipid species (PL and Chol) (Jim Mann, 2001). Different classes of lipoproteins exist and they are often

characterized by the different protein and lipid compositions that define their densities. The different subtypes are chylomicrons (CM), very low density lipoproteins (VLDL), low density lipoproteins (LDL) and high density lipoproteins (HDL) (E.Vance, 2008). Lipoproteins are involved in forward lipid transport (delivery of lipids to peripheral tissues) and reverse lipid transport (movement of lipids from peripheral tissues to liver for catabolism) (E.Vance, 2008). Together, these processes play an important role in maintaining cholesterol homeostasis.

Lipid composition of these particles can be affected by numerous exogenous and endogenous factors, including diet, exercise, fasting, environmental influences, genetic background, age, ethnicity and disease states (Kathiresan et al., 2008; Kwiterovich, 2000; Li et al., 2003; Schreiner et al., 1996). Each lipoprotein has distinct physiological roles that have different adverse health consequences, when present in inappropriate amounts (too high or too low), such as atherosclerosis (Jim Mann, 2001). Cholesterol is the most characteristic component of the HDL lipidome as, in the form of HDL-cholesterol, it represents a major independent negative risk factor for cardiovascular disease (Kontush et al., 2013). Oxidized LDL has been shown to be proatherogenic, and thought to be an important step in the subsequent progression to atherosclerosis (Engelmann et al., 1994). Due to the complex interplay between these factors that contribute to the unique composition of these particles, it is often difficult to study the effects of these components in an unbiased manner.

1.2 Lipidomics

Lipidomics is an emerging field aiming at systems scale detection, characterization and quantification of lipids (Adibhatla et al., 2006). It has taken a forward trend in complementing the field of genomics and proteomics (Weckwerth, 2003; Young and Wallace, 2009). and thus aid in providing a complete picture of a living organism (Vuckovic, 2012). Lipidomics is developing as an independent discipline at the interface of lipid biology, technology and medicine (Wenk, 2005). It is the comprehensive understanding of the influence of all lipids on a biological system with respect to cell signaling, membrane architecture, transcriptional and translational modulation, cell-cell and cell-protein interactions, and response to environmental changes over time (Watson, 2006).

There has been recent advances in analytical platforms used in lipidomics research. Several techniques currently employed in lipidomics include chromatography, mass spectrometry (MS) and nuclear magnetic resonance (NMR) (Adibhatla et al., 2006).

1.2.1 Chromatography

Thin layer chromatography (TLC) (Dittmer and Lester, 1964), gas chromatography (GC) (James et al., 1952) and high performance liquid chromatography (HPLC) (Giddings, 1965) are various types of

chromatographic separation methods that have been used in lipid research for a long time and conditions have been well adapted for the analysis of various classes of lipid compounds.

LC is a popular technique due to its high selectivity and efficiency (Wenk, 2005). It is an easily automated method with quantification potential, and both reverse- and normal-phase conditions have been developed for lipid analyses (Bou Khalil et al., 2010).

1.2.2 Mass spectrometry

MS methods have emerged as the most important and principal detection techniques for lipidomic analysis, often coupled to chromatographic separation methods such as HPLC and GC (Gross and Han, 2011). Lipidomics is not only restricted to phospholipid analysis and MS is widely applied to almost all lipid classes. MS show high sensitivity and selectivity, and new developments in ionization technique, mass analyzer technology, resolving power and mass accuracy have led to much improvements in the detection and analysis of lipids (Lei et al., 2011). Further information on MS will be given in Section 1.3.

1.3 Mass spectrometry in lipidomics research

MS has been widely used for the detection and quantification of PL (Watson, 2006). Subsequent sections will outline the mass spectrometer and current MS platforms used in lipidomics research.

1.3.1 The mass spectrometer

A mass spectrometer consists of three major parts: the ion source, the mass analyzer and the detector (Ekman, 2009).

1.3.1.1 Ion source and ionization methods

The role of the ion source is to create gas phase ions. Analyte atoms, molecules, or clusters are transferred into gas phase and ionized (Ekman, 2009). Several types of ionization methods exist including electrospray ionization (ESI), matrix-associated laser desorption ionization (MALDI) and atmospheric pressure chemical ionization (APCI).

1.3.1.1.1 Electrospray ionization

ESI is a ‘‘soft’’ ionization technique (i.e., little or no molecular fragmentation occurs during ionization) (Fenn JB, 1989). Small charged droplets are formed which eventually break up into smaller droplets, increasing availability of the surface area (Brouwers, 2011) (Figure 1-2).

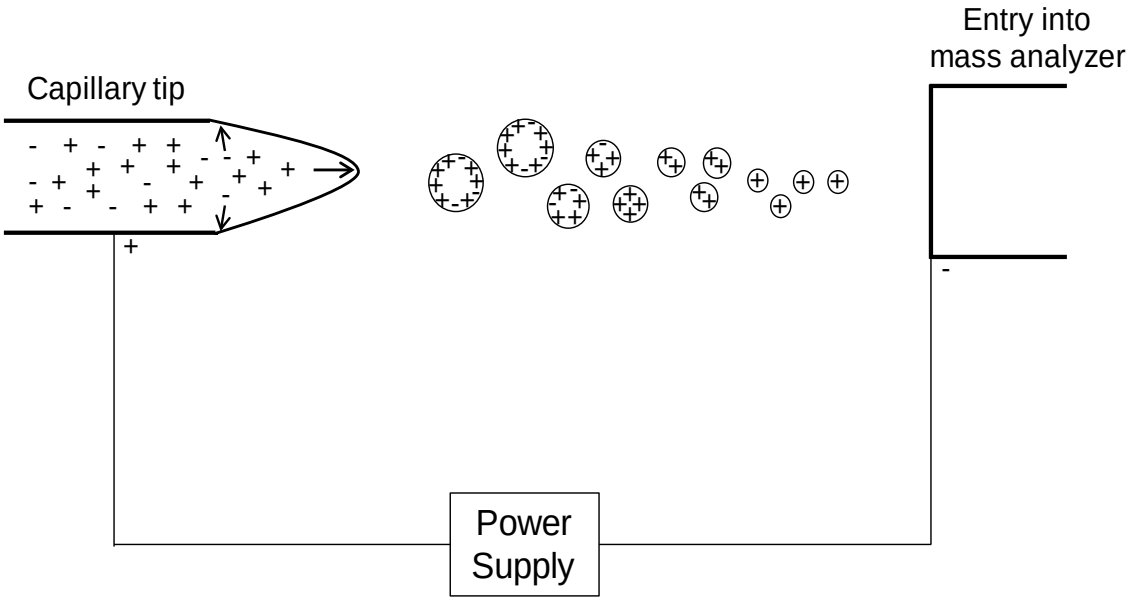


Figure 1-2: Schematic diagram showing the principle of electrospray ionization in positive-ion mode. *Adapted from Han and Gross (Han and Gross, 2005).*

ESI has recently been playing a forward lead in mass spectrometry analysis of lipids, however it has several limitations. For example, ESI substantially depends on the solvent system used for its ionization as it requires sufficient conductivity (Brouwers, 2011). Furthermore, separation of ions depends on solvent viscosity as it affects charge separation. Solvents used for normal phase separations are often made up of organic solvents like

chloroform whereas reverse phase is often made up of water or aqueous buffers mixed with organic solvents like acetonitrile (ACN). As such, normal phase solvents are more suited for ESI due to their higher organic content enabling better sensitivity (Masood et al., 2012).

ESI is also susceptible to matrix effects due to the sample. For example, triacylglycerols can be a potential matrix for the analysis of GPLs in human plasma samples due to their high abundance. While matrix effects can affect ionization and analysis, selective lipid extraction methods and solid phase extraction (SPE) methods can be used for processing the samples to ensure lower matrix effects. The most significant disadvantage of ESI might be its high susceptibility for ion suppression and variability of ion suppression especially when it comes to the analysis of PL (Han and Gross, 2005). Different samples may have different matrix effects due to the differences in biological matrix. This issue can be addressed partly by addition of internal standards similiar to the lipids measured, such that they undergo similar ion suppression effects as the members of the same lipid class within each biological matrix. Standardized sample pre-processing can help to reduce variability introduced by this way. This greatly affects the analysis of low abundant PL species (Brouwers, 2011).

1.3.1.1.2 Atmospheric pressure chemical ionization

A less commonly used ionization technique is chemical ionization (APCI). This method is used for the study of ionic and polar compounds that ionize as monocharged ions and have low molecular weights, up to about 1500Da (Edmond de Hoffmann, 2007). It requires a higher liquid flow rate than ESI. A major advantage of the high temperatures used in APCI is prevention of heating chamber contamination by analytes, which may lead to memory effects because of slowly desorbing analytes and inconsistent ion source performance due to building up of contaminants (Brouwers, 2011). Contrarily, thermal drying imposes a disadvantage due to increased (thermal) decomposition of labile compounds, which may be reflected in the abundance of fragment ions in the mass spectrum (Brouwers, 2011). The greatest advantage of APCI over ESI is that it is able to actively generate ions from neutral molecules, allowing for low to medium polarity analytes to be analyzed. APCI is inherently less susceptible to ion suppression by matrix compounds than ESI (Bruins, 1998).

1.3.1.2 Mass analyzers

The mass spectrometers currently employed in lipidomics include time-of-flight (TOF), ion trap, triple quadrupole (QqQ) and fourier transform ion cyclotron resonance (FT-ICR) (Bou Khalil et al., 2010). Here, I will focus only on QqQ-MS, the method used for this study.

1.3.1.2.1 QqQ-MS

A linear quadrupole consists of four precisely matched parallel metal rods where the direct current (dc) and radio frequency(rf) potentials are applied between the two opposing pairs of these electrodes to produce a high frequency oscillating electric field (Yost and Enke, 1978). Mass separation is thus made possible by the oscillatory motions of ions in this electric field. Ions of a specific *m/z* value pass through the quadrupole rods only with specific values of the dc and rf potentials (Yost and Enke, 1978). A mass spectrum is usually obtained by changing both the dc and rf potentials while keeping their ratio constant (Chhabil, 2007). At specific dc and rf potentials, a specific range of *m/z* values have stable trajectories within the quadrupole field to reach the detector while all other ions are lost because they follow unstable trajectories (Chhabil, 2007).

In QqQ instruments, three quadrupoles are arranged sequentially. While the first (Q1) and third (Q3) quadrupoles are operated in the mass-selective mode using dc and rf potentials, the second quadrupole (q2) is operated with only the rf potential (Yost and Enke, 1978) (Figure 1-3).

Tandem MS can be performed with QqQ instruments. The second quadrupole, which nowadays often is a hexapole or octapole, is not used for *m/z* selection or scanning (Ekman, 2009), but serves as a collision cell containing nitrogen gas at a pressure, such that an ion entering the quadrupole undergoes one or several collisions (Edmond de Hoffmann,

2007). This can induce ion fragmentations that provide chemical information (Bou Khalil et al., 2010). With such a setup, low-energy collision induced dissociation (CID) can be performed (Ekman, 2009). This technique is efficient in producing fragment ions, but reproducibility between instruments is limited since the spectra obtained heavily depend on the selected collision gas (such as nitrogen) and its pressure on the collision energy, as well as on other parameters (Ekman, 2009).

1.3.1.2.1.1 Scan modes and shotgun approach

There are three main available scan modes using tandem MS to analyze lipids, namely precursor ion scan, neutral loss scan and product ion scan (Edmond de Hoffmann, 2007). Multiple reaction monitoring (MRM) is one of the most powerful methods of quantitative targeted analysis (Watson, 2006) (Figure 1-3).

MRM consists of selecting a fragmentation reaction where both the first and second analyzers are focused on selected masses (Edmond de Hoffmann, 2007). Ions selected by the first mass analyzer are only detected if they produce a given fragment by a selected reaction (Edmond de Hoffmann, 2007). The absence of scanning allows one to focus on the precursor and fragment ions over longer times, increasing the sensitivity and selectivity (Ekman, 2009). Tandem mass spectrometers are able to analyze a large

number of MRM transitions with dwell times as low as 5-10ms, acquiring cycle time of 10-15 data points per transition allows a reproducible quantitation.

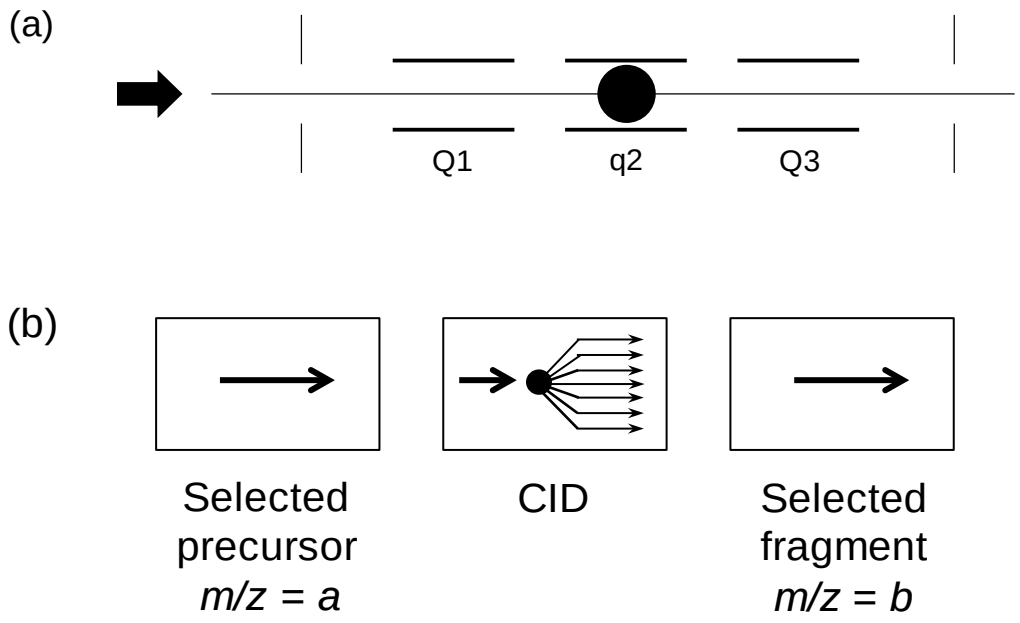


Figure 1-3: Diagram of a triple quadrupole instrument. (a) The first and the last (Q1 and Q3) are mass spectrometers. The central quadrupole, q2, is a collision cell made up of a quadrupole using rf only. The quadrupole mass spectrometers are symbolized by upper case Q, and the rf-only quadrupoles with a lower case q. (b) MRM approach. CID stands for collision-induced dissociation, as occurs when an inert gas is present in the collision cell. *Adapted from Hoffmann (Edmond de Hoffmann, 2007).*

Advancements in ESI-MS have allowed the development of shotgun lipidomics (also known as direct infusion) (Han and Gross, 2005). Shotgun lipidomics is one of the commonly used approaches for rapid global analysis of lipids in crude biological extracts (Han and Gross, 2005). Intact total lipid extracts are infused directly into MS via ESI, and lipid species are subsequently identified and quantified by tandem MS (Isaac, 2011). A cocktail

of non-naturally occurring internal standards representative of the different lipid classes is added to the lipid extracts for quantification purposes (Isaac, 2011). The internal standards are used to correct for multiple experimental factors such as possible variations in MS parameters, which can affect ionization and ion count for accurate quantification (Isaac, 2011).

In shotgun lipidomics, the composition of the sample entering the instrument is constant, so ratios of internal standards to compounds of interest are constant throughout the analysis, making quantification straightforward (Isaac, 2011). While prior chromatographic separation methods like LC are used to separate the different lipid classes before ionization, the shotgun approach provides a good tool for the analysis of thousands of molecular species in a short time period (Han and Gross, 2005).

The main drawbacks of this method are potential difficulties in resolving isobaric compounds, especially in the absence of tandem MS techniques, and a risk of ion suppression that may lead to decreased sensitivity for the analysis of minor lipid species (Isaac et al., 2007). Furthermore, compounds with high *m/z* ratios (>1000) like gangliosides may give rise to multiply charged ions, and this can reduce ion intensity due to different charge distribution.

1.3.2 MS lipid analysis

1.3.2.1 Terms

Signal-to-noise ratio (S/N) describes the uncertainty of an intensity measurement and provides a quantitative measure of signal quality by quantifying the signal intensity relative to noise (Gross, 2011). Noise is often defined as the background that is coming from the electronics of the instrument, presenting itself both in between signals and on the signals (Gross, 2011). The reduction of noise is proportional to acquisition time or number of single spectra that are averaged (Gross, 2011). Thus averaging a larger number of spectra can result in a lower noise (Ekman, 2009). A S/N of three or more is often used as a guideline for detection limit for a reliable signal (U.S. Department of Health and Human Services, 2001).

1.3.2 Glycerophospholipid, sphingolipid and sterol analysis by MS

The importance of ESI in detection of PL was first demonstrated by Han and Gross (Han and Gross, 1994). Membrane PL were detected with enhanced sensitivity and did not require prior chromatographic separation. ESI-MS enables high sensitivity and can be used for molecules without any ionizable site (for e.g. PE and PC) by losing or gaining of protons, or by formation of sodium, potassium, ammonium, chloride, acetate or other adducts (Edmond de Hoffmann, 2007). Furthermore, ESI-MS allows for rapid and efficient analysis of lipid classes, subclasses, and individual molecular

species directly from crude extracts with a dramatically higher S/N ratio in comparison to other mass spectrometric approaches such as MALDI coupled to TOF MS (Han and Gross, 2005).

Targeted lipidomics using ESI has been shown to be useful in the analysis of cellular lipids (Han and Gross, 2005). It is best used for quantification of a select number of molecular lipid species that can be controlled by appropriate lipid standards to minimize possible errors arising from changes in lipid ionization efficiency (Ekroos et al., 2010). MRM is based on a targeted approach which has previously been used successfully for quantification of SPs (Merrill et al., 2005) and GPLs (Chan et al., 2008; Han and Gross, 2005; Wenk et al., 2003) from various biological sources.

Many lipids naturally carry charges and this makes them ideal for ionization and analysis via mass spectrometry. PE molecular species contain a polar head group and can be analyzed by ESI-MS in either the positive or the negative ion mode (Han et al., 2005). It can be measured via monitoring precursor of *m/z* 196 from deprotonated PE molecular ions in the negative ion mode (Brugger et al., 1997).

PA has the simplest polar head group, and is possible to observe $[M-H]^-$ and $[M+H]^+$ ions upon ESI, though the most abundant ions are by collision-induced decomposition of $[M-H]^-$ ions (Hsu and Turk, 2000b). It can be measured by monitoring precursor of *m/z* 153 for glycerol containing phospholipids in the negative ion mode.

PI contain a phosphodiester of the six-carbon sugar inositol as a polar head group (Pulfer and Murphy, 2003). Tandem mass spectrometry of PI $[M-H]^-$ ions showed a similar behaviour to that of PA in the loss of the sn-1 and sn-2 FA substituents as neutral carboxylic acids and ketene, and the presence of carboxylate anions as major product ions resulting in the detection of several ions unique for PI including monitoring precursor of *m/z* 241 for inositol containing phospholipids (Hsu and Turk, 2000a).

PG serve as precursors for more complex PL, including cardiolipins found in mitochondria (Pulfer and Murphy, 2003). There have not been many detailed studies using ESI-MS to analyze PG beyond the observation that they form abundant negative ions which is expected for acidic PL carrying neutral polar head groups (Pulfer and Murphy, 2003). They can be detected by monitoring precursor of *m/z* 153 for glycerol containing phospholipids in the negative ion mode.

Ionization of PS during ESI results in the formation of positive and negative ions, however, negative-ion formation tends to dominate for this PL class (Koivusalo et al., 2001). Collisional activation of the $[M-H]^-$ ion results in the formation of ions that correspond to the loss of serine [M-88], leading to the detection of PS by scanning neutral loss of *m/z* 88 from the parent mass in negative ion mode (Pulfer and Murphy, 2003).

PC are the most abundant lipid species in eukaryotic cells, and are characterized by the presence of a quaternary nitrogen atom whose positive

charge is neutralized by the negative charge of the phosphate group (Pulfer and Murphy, 2003). PC are always ionized, where the quaternary nitrogen atom readily forms an abundant $[M+H]^+$ ion by ESI because the phosphate anion can be protonated during the electrospray process (Pulfer and Murphy, 2003). Collision-induced decomposition of the positive protonated molecule ion $[M+H]^+$ yields the expected phosphocholine ion at *m/z* 184, which is typical of all phosphocholine-containing lipids.

ESI of SM follows closely the related choline-containing GPLs, PC where *m/z* 184 is the only ion observed following collisional activation, which corresponds to the protonated phosphocholine species (Pulfer and Murphy, 2003). It can be measured in the positive ion mode. However, at low collisional energy, we can also observe the loss of water (*m/z* 18) in the positive ion mode which acts as a more accurate detection method for SM using shotgun approach.

Gas chromatography (GC) approaches and HPLC with ultraviolet detection have commonly been used for STLs analysis (Honda et al., 2008). Recently, LC-APCI based tandem MS has been introduced as more sensitive, specific and rapid quantification method for analyses of STLs (Honda et al., 2008). Targeted approaches like MRM have been used to analyze Chol with reasonable sensitivity (Briche et al., 2002; Raith et al., 2005; Sevanian et al., 1994).

Despite advancements APCI and ESI, ESI is still the preferred choice when it comes to the analysis of PL. Studies have shown that most PL

classes showed a better sensitivity with ESI when compared to other ionization methods (A.Å. Karlsson, 1996; Byrdwell, 1998; Ismaiel et al., 2008).

1.4 Variation and Lipid Profiles

1.4.1 Technical Variation

1.4.1.1 Instrument and process variation

As mentioned before, MS has emerged as one of the more popular techniques used in detecting and identifying lipids (Watson, 2006). A large diversity in the types of mass spectrometers exists, some of them better suited for certain applications (Edmond de Hoffmann, 2007). The choice of instrumentation for lipidomics studies is important and is largely dependent on the study objectives and designs (Bou Khalil et al., 2010).

When looking at technical variation, instrumentation and upstream sample processing can be considered as major sources of variation. Despite the well-known importance of instrument variation, understanding of upstream process variation is equally important (Ekroos et al., 2010). Proper sample storage, handling and preparation are crucial in the analytical process to avoid artificial formation, modification, or loss of lipid species (Ekroos et al., 2010) as it can lead to an unwanted reservoir of variation. Several factors can affect

the process variation involved in human plasma processing, some of which are listed below:

1) **Sample collection** – how the blood is drawn, which type of anticoagulant is used (EDTA, heparin, citrate) (Bando et al., 2010; Bernini et al., 2011).
2) **Processing plasma** - how plasma is processed from the blood e.g. centrifugation, mixed before aliquoting etc (Bernini et al., 2011).
3) **Time of day** – time of collection and effect of circadian rhythm (Altmaier et al., 2008; Dallmann et al., 2012).
4) **Fasting status** – fed or fasting state, duration of fasting (Hammad et al., 2010; Krug et al., 2012).
5) **Number of freeze-thaw cycles** - depends on the number of aliquots and how many times each tube was used prior to lipid analysis (Bernini et al., 2011; Suhre and Gieger, 2012; Teahan et al., 2006).
6) **Sample storage** – time taken from sample processing to storage at -80˚C and duration of long-term storage (Bernini et al., 2011). Use of appropriate antioxidants to prevent lipid oxidation (generation of F2-isoprostanes, malondialdehyde etc) during storage (Breusing et al., 2010).

It is important to control as many variables as possible in order to minimize process variation, though standardizing every factor is difficult. Another source of technical variation stems from the instrument used for analysis, i.e. MS (Ekroos et al., 2010). As MS is widely used for lipids, it is

important to be able to understand the performance specifications and limitations of the instrument before being able to use it for quantitation and detection (Crutchfield et al., 2010; Edmond de Hoffmann, 2007). A different set of factors influence variation including instrumentation parts (e.g. sample loop, injection needle, probe, capillaries, and drift of operation of electronic components) and sample analyses (matrix effects, ionization) (Benton et al., 2012; Dunn et al., 2012). As such, the study design plays a critical role in controlling or accounting for these variables as much as possible (Suhre and Gieger, 2012).

Typically, the greatest issue during the course of a MS analysis is a change in sensitivity (normally a decrease) as sample components aggregate in the electrospray ion source (affecting ion transmission from atmospheric to vacuum regions of the MS) (Dunn et al., 2012). Matrix components from samples are the primary reason for degradation in analytical performance (Dunn et al., 2012). Furthermore, uncontrolled fluctuations in signal intensity due to the lipid extraction and analytical processes can affect data processing and interpretation (Dunn et al., 2012). Fortunately, non-naturally occurring lipid standards can be spiked into the samples at different stages of sample processing, which facilitates the correction of signal intensity, as previously described (Bou Khalil et al., 2010). As different classes of GPLs and SPs have different ionization patterns, it is essential to spike an internal standard unique for each lipid class measured. This facilitates data normalization under the assumption that a given lipid standard undergoes similar fluctuations as members of that lipid class (Benton et al., 2012; Wenk, 2005).

A recent review by Suhre and Gieger (Suhre and Gieger, 2012) has highlighted the important factors to be considered in the study design when looking at biological variation. An overview of this study's workflow is shown in Figure 1-4.

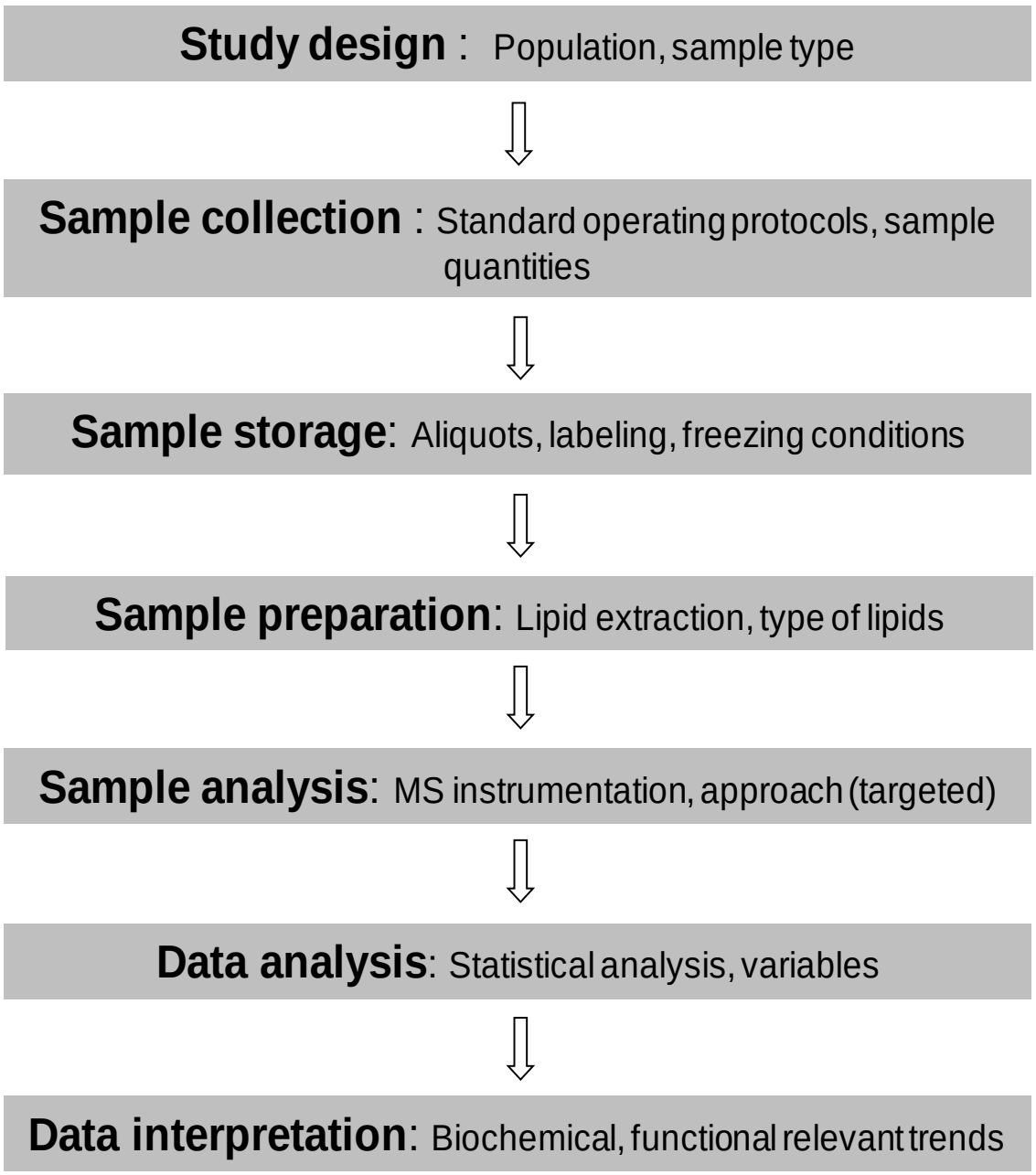


Figure 1-4: Workflow considerations in this study.

1.4.1.2 Quality control samples

One measure of controlling technical variation is through the employment of proper quality control samples (QC) (Gika et al., 2012). Generally, QC samples are similar in metabolic and sample matrix composition as the biological samples under study (Dunn et al., 2011). This serves as a rationale comparison to other samples in the study when looking at sample intensity and instrument performance. QC samples can be of 2 types: 1) pooled QC samples, essentially a mixture of either all or a small subset of the samples in the study that would be uniformly used throughout the study and 2) external QC samples, such as commercially available QC samples or samples obtained from other studies used for the same purpose (Dunn et al., 2012; Gika et al., 2012). There are advantages and disadvantages for both types of QC samples. For example, pooled QC samples are very similar in composition to the biological samples under study and will be more comparable (Benton et al., 2012; Dunn et al., 2012). External QC may be cleaner in matrix background but have different composition of lipids and amounts, thus making their comparison with samples from a given study set more difficult (Gika et al., 2012).

Study design plays an important role in determining the type of QC samples to be chosen. Studies involving a small controlled sample population, where collection of samples can be completed prior to processing and analysis, can utilize the pooled QC approach (Dunn et al., 2012). For larger

studies where sample collection takes a longer time period and sample processing as well as analysis is performed continuously, external QC may be a better option. QC samples are vital in any study design as they can be used to account for both process variability as well as instrument variability. QC samples that are extracted and stored with the biological samples can serve as a platform to look for processing variation between different batches of samples or can be used to monitor instrument variation and analytical reproducibility within a batch analysis, sometimes termed as an analytical block (Kamleh et al., 2012). This can help to determine if a particular batch analysis can be accepted or rejected based on the percentage of coefficient of variation (COV) of the QC samples, typically acceptable up to around 15-20% of the QC mean (Kamleh et al., 2012). QC samples can also be used for signal correction between analytical blocks (Kamleh et al., 2012). The frequency of QC sample injections also matter and the number of QC samples to be analyzed would depend on the length of each analytical block (Dunn et al., 2012). This is especially useful when there are many samples having staggered analysis times over longer time periods, where instrument sensitivity and detection may not be constant throughout (Dunn et al., 2012).

The incorporation of QC samples in study design is becoming more prominent and is considered as a reliable standard for controlling technical variation (Gika et al., 2012; Kamleh et al., 2012).

1.4.1.3 Technical variation for measured lipids

In recent years, several groups have attempted to establish the range of technical variation in lipid measurements using MS (Crews et al., 2009; Ekroos et al., 2010; Sandra et al., 2010). Despite different lipid extraction procedures and MS types harnessed for analysis, the general reported technical variation is within 20% whereas some lipid classes show a variation of up to 30% (Chau et al., 2008; Ekroos et al., 2010). Clearly, higher intensity lipids tend to have a lower COV and are thus less variable across repetitive measurements as compared to lower intensity lipids which are close to background noise (Horwitz, 1982).

1.4.2 Lipids as biomarkers

Lipid metabolism is tightly regulated and has an impact on various physiological functions including membrane composition, signal transduction, energy storage and biological messenger functions (Wenk, 2005). Aberrant lipid metabolism often leads to onset of pathologies, and thus the precise balance of lipids and their effectors can serve as biomarkers (Fernandis and Wenk, 2007; Wymann and Schneiter, 2008). Most importantly, lipid biomarkers can be placed into biological context by means of their metabolic pathways (Wiest and Watkins, 2007). However, biomarker studies often face problems of non-uniformized sample collection, processing and analysis

(Teahan et al., 2006). Reproducibility of analytical platform used is also important for reliable detection of biomarkers (Gao et al., 2005).

There is an increasing amount of clinical studies being performed in biofluids, looking at the potential role of lipids in various pathological conditions including neurological (Blennow et al., 2010; Schwarz et al., 2008), immunological disorders (Brinkmann, 2007; Rivera and Olivera, 2007; Wheeler et al., 2008), diabetes (Lin et al., 2006) and cardiovascular conditions (Stojiljkovic et al., 2002). In the general human population, genetic, environmental and lifestyle factors significantly differ, including physical activity, mental status as well as microflora (both in and on man), thus the biological variability will be high (Mamas et al., 2011). If these factors are not carefully considered, it can lead to difficulty in reproducing identified biomarkers from published studies across various pathological states (Wiest and Watkins, 2007).

While the search for better therapeutic interventions based on lipid biomarkers are necessary (Bou Khalil et al., 2010), it is imperative to first understand and quantify the natural variation of lipids among healthy individuals before differentiating a disease state from a healthy individual (Krug et al., 2012; Stoop et al., 2010). For example, if a person suspected of systemic lupus erythematosus (SLE) gives plasma for diagnostic testing, the panel of lipids measured can be identified as significantly variable if the values are outside the range of normal biological variation for those lipids. Thus, it is vital to first establish the natural variation of lipids in plasma across a large

well-represented cohort of healthy individuals in any population prior to future lipid biomarker development (Figure 1-5). Biological variation in human plasma is further discussed in Section 1.4.3.

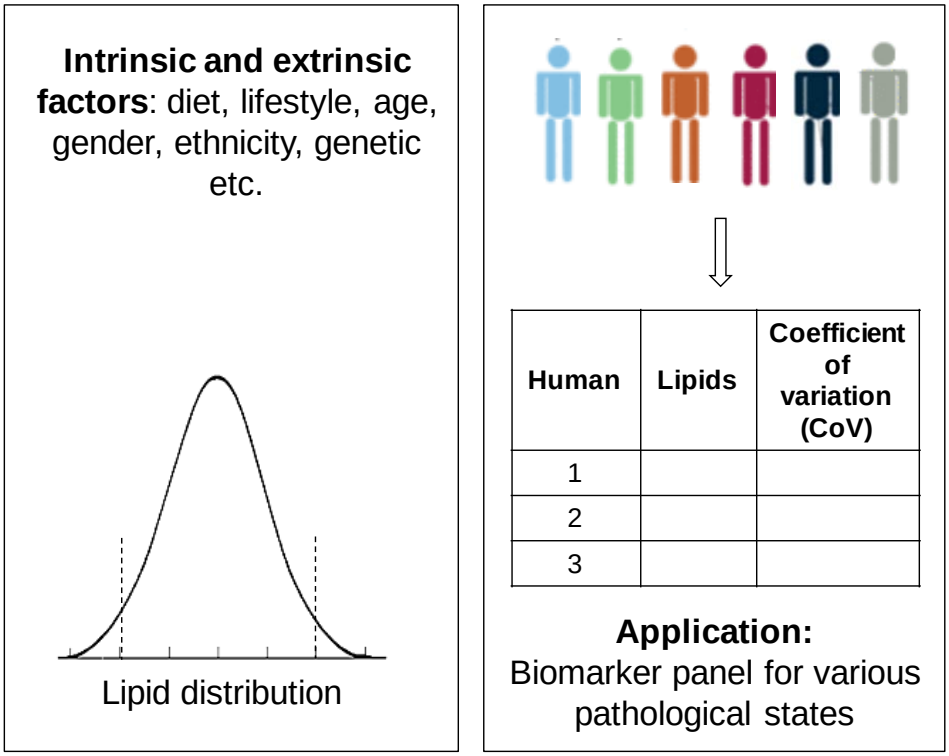


Figure 1-5: Conceptual understanding of the biological variation across individuals in a population (left panel) and its application in subsequent biomarker studies (right panel).

1.4.3 Biological variation

1.4.3.1 Biological variation in human plasma

Lipids are metabolites which are the endpoints of the molecular pathways regulated by proteins (Bruce et al., 2009). With the advent and improvement of modern day technology such as MS and NMR, there is an increasing interest to identify and quantify these lipids across various complex biological materials including human plasma.

A recent comprehensive profiling of the human plasma lipidome by Quehenberger et al. has shed some light on the magnitude of structural diversity of these lipids (Quehenberger et al., 2010). The ability to distinguish biological, i.e., “true”, variation from technical variation is a key concern in metabolic profiling studies, regardless of the analytical platform and the sample type (Lindon et al., 2005; Teahan et al., 2006). Characterizing the technical range of variation due to factors like sample collection is important to prevent misinterpretation of data (Bando et al., 2010). It is essential to define the degree of technical variation before measuring absolute quantities of these lipids (Masson et al., 2011; Molloy et al., 2003) in order to obtain a ‘‘natural,’’ stable, and invariant metabolic profile that is typical of a given subject, even if not necessarily unique in healthy human individuals (Assfalg et al., 2008).

1.4.3.2 Factors affecting biological variation

The lipidome is affected by various physiological factors like genetic drift, age, dietary variation, estrus cycle (Bell et al., 1991; Bollard et al., 2001; Gavaghan et al., 2002), gender, ethnicity (Lawton et al., 2008), feeding state and circadian rhythm (Dallmann et al., 2012). This exarcebates the difficulty in distinguishing between pathophysiological responses and biological variation (Bollard et al., 2005).

The dynamic changes that a lipidome undergoes was coined plasticity (Houshyani et al., 2012). Lipid levels are known to be perturbed on a day-to-day basis (Dallmann et al., 2012; Minami et al., 2009). In mice, some lysolipids display remarkable circadian patterns with up to 2-fold differences in their levels (Minami et al., 2009). Whereas modern mass spectrometers provide linear outputs over several orders of magnitude (linear dynamic range), the biologically relevant dynamic range is lipid specific, varying from two- to three-fold for abundant membrane lipids to 10- to 100- fold in extreme cases (such as mediator lipids) (Wenk, 2010). Furthermore, lipids are also found at very different basal concentrations and have distinct temporal dependencies (Wenk, 2010). This may vary depending on cell type or organism of interest.

Studies on biological variation of healthy populations have been reported. A recent study by Rutherford et al. reported higher prevalence of

lower HDL-Chol levels in the healthy Filipino women population when compared to the US population, highlighting ethnic differences in lipid profiles (Rutherford et al., 2010). Similarly, effects of age (Yu et al., 2012), gender and ethnicity were studied (Hicks et al., 2009) by MS in plasma samples of a healthy representative US population consiting of Caucasians, Hispanics and African-American where significant changes with lipid metabolism was observed with age (Lawton et al., 2008).

Genetic variation can also have a significant impact on lipid levels (Demirkan et al., 2011; Hicks et al., 2009; Teslovich et al., 2010). Other studies which have been predominantly done on European populations (Hicks et al., 2009; Kathiresan et al., 2008) and East Asians (Chinese, Koreans and Filipinos), South Asians and African Americans (Teslovich et al., 2010) additionally confirm ethnicity dependent differences in lipid metabolism. Several studies have shown that single-nucleotide polymorphisms (SNP) in genes involved in lipid metabolism affect lipid levels through catabolism, anabolism and trafficking of lipids (Hicks et al., 2009).

1.4.2.2.1 PUFA lipids and biological variation

Many studies have arguably considered dietary intake as an important aspect of variation in lipid profiles across different populations, including supplementation and dietary polyunsaturated fatty acids (PUFA) (Calzada et al., 2010; Hodge et al., 2007b; Rudkowska et al., 2013). PUFA

refer to the class of fatty acids with multiple unsaturations in the aliphatic tail (Tanaka et al., 2009). PL containing short-chain PUFA are endogenously synthesized while PL containing long-chain PUFA can either be synthesized from two essential fatty acids (oleic acid and palmitic acid) or obtained exogenously via dietary sources (Tanaka et al., 2009) . Depending on the position of the double bonds, long-chain PUFA can be either classified as n-3 or n-6 (E.Vance, 2008). PUFA can be incorporated by other lipids in their chains and thus studying their effect on lipid profiles is important (Tanaka et al., 2009). n-3 and n-6 PUFA are commonly used in dietary intervention epidemiological studies, where lipid profiles of various body fluids and organs (plasma, serum, red blood cells, and adipose tissue) are studied over a period of time (Tanaka et al., 2009). For example, the consumption of dietary flavonoids associated with a fruit and vegetable rich diet has been shown to increase plasma levels of n-3 PUFA (Toufektsian et al., 2011).

Studies have shown a strong correlation between intake of PUFA with plasma PL FA composition (Hodge et al., 2007b) and changes in LDL composition (PL and CholE) in healthy individuals (Calzada et al., 2010). However, several controversies exist on the influence of diet on lipid profiles and this could be due to absence of standardized study designs and reporting methods to assess dietary intake (Tanaka et al., 2009).

Some of the studies performed till date on biological variation in healthy individuals is summarized in Table 1-1. It can be seen that more biological variation studies have been carried out in recent years. Based on

literature, it is recognized that there is a need to characterize biological variation from adult healthy individuals, particularly in populations of Asian ethnicity. Multi-ethnic societies like the Singaporean population provide an interesting cohort to study due to their cohabitation and unique genetic backgrounds (Suo et al., 2012).

Subject population	Biofluid	Factors studied	Lipids measured	Population size	Reference
Italy, Finland, USA	Plasma, rbc, platelets	Dietary fats	Cholesterol, glycerophospholipid, sphingolipid	62	Dougherty et al (1987)
Germany	Serum	GWAS	Glycerophospholipid, sphingolipid	284	Gieger et al (2008)
USA (Caucasian, African-American, Hispanic)	Plasma	Age, ethnicity, gender	Cholesterol, fatty acids	269	Lawton et al (2008)
Italy	Plasma, rbc	GWAS, dietary intake	PUFA	1616	Tanaka et al (2009)
Netherlands, Italy, Sweden, Scotland, Croatia	Plasma, serum	GWAS	Sphingolipid	4400	Hicks et al (2009)
Philippines, USA	Plasma	Diet, BMI	Cholesterol, triglycerides	1896	Rutherford et al (2010)
Switzerland	Plasma, saliva	Circadian clock	Fatty acids	10	Dallmann et al (2012)
Germany, UK	Serum	Age	Phosphatidylcholine, sphingomyelin	2886	Zhonghao et al (2012)

Table 1-1: Summary of some human biological variation studies performed in recent years. *Abbreviations:* rbc is red blood cell, GWAS is genome-wide association studies.

1.4.3.3 Biological variation in the healthy Singaporean population

The Integrative Omics platform (iOmics) was initiated in 2010 and involves three main teams, genomics, transcriptomics and lipidomics. This platform was established to provide a descriptive analysis of the natural physiological variation between healthy individuals of Singaporean population, comprising of equal ratios of the Chinese, Indian and Malay ethnicities. The participants were selected based on the absence of major metabolic disease, as described in Chapter 2. This is the first study looking at biological variation in a unique Asian population, representative of most of Southeast Asian countries. The work presented in this thesis is based on the lipidomics platform, to look at the natural variation of targeted lipids across the three ethnic groups.

1.5 Aims of this thesis

Based on the importance in establishing true biological variation, I decided to first quantify technical variation by employing several methodological considerations like QC before analyzing biological variation. Here, an ESI-based QqQ targeted approach (MRM) is used to quantify 138 lipid species across 10 lipid classes of GPLs, SPs and STLs known to exist in the human plasma (Quehenberger et al., 2010) in 360 healthy individuals in the Singapore population. The following aims are covered in this thesis:

1) Technical variation from sample processing and instrumentation.

2) (a) A small scale pilot study (separate from the main study) looking at effect of multiple time-point collection on lipid profile.

(b) Biological variation for 360 healthy fasting individuals that is representative of the Singaporean population is analyzed.

CHAPTER 2

MATERIALS AND METHODS

2.1 Chemical and reagents

Chloroform and methanol were purchased from Fisher Scientific (UK) and were of LC grade. Chloroform and methanol were purchased in bulk to ensure same production batch to reduce variability due to solvent composition (Dunn et al., 2012). Piperidine was purchased from Sigma Aldrich (USA). Milli-Q water was obtained from Arium 611VF purification system (Sartorius, Germany).

2.2 Materials

Axygen MCT-175-C Ultra clear eppendorf tubes were purchased from Axygen Scientific (USA). Cryotubes (2 ml and 5 ml) were purchased from Greiner Bio-one (Germany). BD Vacutainer® plastic plasma tubes with ethylenediaminetetraacetic acid (EDTA) as anti-coagulant were purchased from Biomed Diagnostics (Singapore). All tubes were purchased in bulk to ensure same production batch. All lipid standards were purchased from AVANTI Polar Lipids (USA) and Echelon Biosciences (USA) (for GPL and SP

standards) and CDN Isotopes Inc. (Canada) and Medical Isotopes Inc (USA) (for ST standards). Nitrogen gas for running the mass spectrometer was purchased from Cryoexpress Singapore (Singapore).

Triple quadrupole instrument ABI 4000 (Applied Biosystems, USA) coupled with Agilent 1100 Autosampler system (Agilent Technologies, USA) was used for measurement of all GPLs and SPs using ESI. Triple quadrupole instrument ABI 3200 (Applied Biosystems, USA) coupled with Agilent 1200 series HPLC system (Agilent Technologies, USA) was used for analysis of STLs using APCI. Nitrogen was used as a collision gas in all instruments.

2.3 Ethics statement

All healthy human plasma samples were collected in accordance to ethical guidelines and protocols approved by the National Health Group Institutional Review Board (IRB 10-434), Singapore. All participants who were approached and agreed to participate in this study were required to provide written consent. Participants could withdraw, refuse participation or discontinue at any point during the study without further prejudice.

2.4 Sample collection and processing

2.4.1 Pilot study

All participants were of Singaporean nationality consisting of 4 Chinese, 2 Malay and 2 Indian (in round 1) and 4 Chinese,4 Indian and 1 Malay for round 2 (validation round). Ethnicity of each subject was ensured by obtaining verbal confirmation on information of their grandparents' and parents' ethnicity. Subjects voluntarily participated in this study and did not have any existing major health conditions like diabetes and cardiovascular disease. Female subjects were ensured not to be pregnant at time of sampling. Samples were collected by trained phlebotomists. All participants had fasted overnight and blood was taken after a minimum of 10 hours of fasting. Only consumption of water was allowed during fasting period. A standardized breakfast was given to all participants within 30 mins of the first blood draw. Blood was drawn again two hours post breakfast. Participants were then given a standardized lunch and the last time-point for blood collection was done two hours post lunch. An overview of the pilot study for both rounds is shown in Figure 2-1.

10 ml of whole blood samples were collected into BD Vacutainer® plastic plasma tubes with EDTA as anti-coagulant by venipuncture. Tubes were inverted three to five times to allow adequate mixing of blood with the anticoagulant. Plasma was processed by spinning 10 ml of the collected blood

at 2200xg for 15 mins at 4˚C using a swing-out bucket rotor centrifuge (Allegra 6R Centrifuge, Beckman Coulter, USA). The top plasma layer was then removed by carefully transferring the supernatant to fresh 5 ml cryovials. 500 µl of plasma near the plasma-blood interface layer was left to avoid any possible contamination by platelets. Subsequently, the plasma was aliquoted to 100 µl and 1 ml aliquots in fresh tubes. The plasma was initially frozen to –20°C before transferring to –80°C within 45 mins of blood collection. All aliquots were stored at -80°C. Sample aliquots were ensured to undergo same number of freeze-thaw cycles to ensure consistency in sample processing.

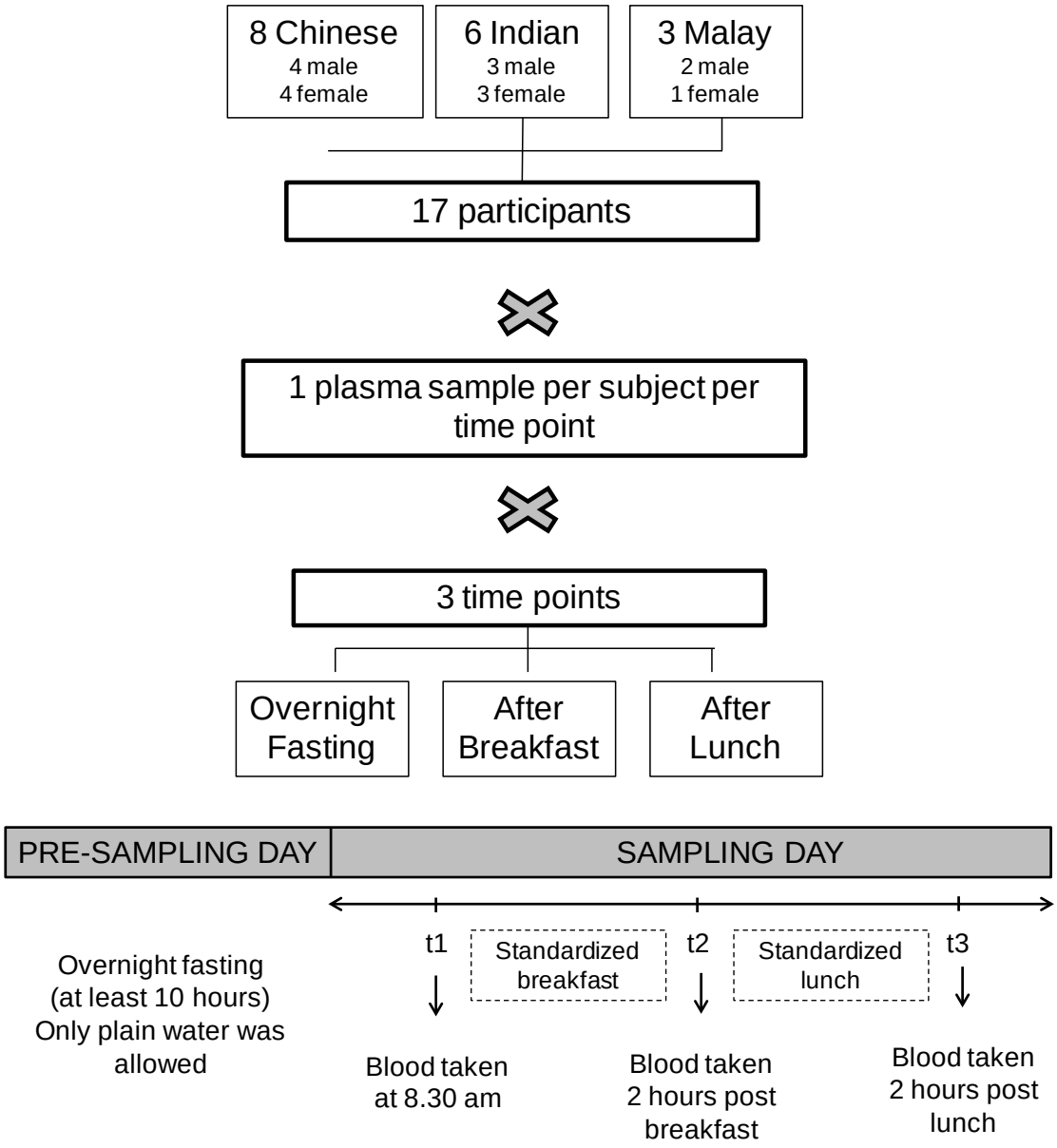


Figure 2-1: Overview of sample collection procedure used in the pilot study.

2.4.2 Main study

All participants were of Singaporean nationality, consisting of equal proportions of Chinese, Malay and Indian ethnicities. Samples were collected from healthy Singaporeans defined by absence of major clinical diseases including diabetes and stroke. Participants were selected from the Singapore Consortium of Cohort Studies. Their age group was restricted to 40 to 69, with an average age of 49 to 53 across the three ethnic groups. Higher compliance rates of females to participate in this study resulted in a 1.7:1 female:male ratio across all three ethnicities. Samples were collected at a fixed venue by trained phlebotomists over a period of seven months. Participants were required to fast overnight for at least eight hours and samples were collected between 8 to 10 am in the morning. Blood collection was done as previously described in Section 2.4.1.

Chol, HDL and LDL was clinically measured in processed plasma. Briefly, plasma was processed by spinning blood at 3000rpm for 10 mins at room temperature. Chol, HDL and LDL levels were then measured by colorimetry method, as routinely used in clinics. These measurements were made by Lee Kian Ng (National University Hospital).

Background information of the participants was collected using the Singapore Consortium of Cohort Studies official questionnaire. The comprehensive 50 page questionnaire included information on the

participants' lifestyle factors (smoking), personal and family medical history, physical activity and social background. An intensive food frequency questionnaire was also included to gauge participants' dietary habits. These information was not analyzed in this study.

2.4.2.1 Technical variation studies

Three technical variation studies were performed – technical variation study (TVS), validation study (VS) and quality control study (QCS).

In TVS, 10 samples were chosen from the 360 samples and four sample aliquots from each individual were stored and processed together to account for process variation. Each of these aliquots was analyzed by MS at two different days with two replicates each (total four times) to account for instrument variation. This resulted in 16 MS analyses for each individual sample and subsequent analyses were done based on process replicates or instrument replicates. Technical variation data was analyzed based on process, instrument and subject variation, and represented by coefficient of variation (COV) vs intensity plots. Process variation is the COV calculated based on the measurements across four aliquots of the same sample, i.e. an average of all MS measurements for these 4 aliquots, averaged across the 10 subjects. Instrument variation is COV calculated based on the MS measurements for each aliquot across the two days, i.e. eight measurements in the mass spectrometer, averaged across the 10 subjects. Subject variation is the COV calculated based on an average of all MS measurements per

individual subject, averaged across the 10 subjects in this study. Process variation includes instrument variation due to the calculation of COV measurements across the aliquots run on different days. Data is represented as COV vs intensity plots for all the lipid classes.

In VS, 18 samples were randomly chosen from the main study and re-analyzed to give an additional measure of technical variation. In QCS, pooled samples from the main study were used as QC samples. QC samples were analyzed 20 times within the same day to assess intraday variation for all GPLs and SPs. As the same sample was analyzed repeatedly, this study gives an indication of instrument variation within a day, to assess its suitability for large scale studies involving multiple batches. Data from both these studies are represented in the form of boxplots and barplots.

2.5 Sample Handling

All samples were thawed and aliquoted into 1 ml in cryovials and 100µl in eppendorf clear tubes. The former was stored for future independent validation analysis and the latter was used for analysis in this study. One freeze-thaw cycle was defined as thawing a sample at 4°C for more than 30 mins. The number of freeze-thaw cycles were recorded and kept consistent within samples. All samples were frozen and stored at –80°C until lipid extraction.

Detailed information on study design will be outlined in the technical variation chapter (Chapter 3).

2.6 Lipid extraction

Pre-chilled solvents were used throughout lipid extraction. Lipids were extracted according to a modified Bligh and Dyers extraction method (Bligh and Dyer, 1959). 100 μl of thawed plasma was used for lipid extraction as previously described (Guan et al., 2006; Shui et al., 2007). Briefly, 900 μl of ice cold chloroform-methanol 1:2 (volume to volume (v/v)), were added to the plasma and vigorously vortexed for 1 min followed by shaking at 4˚C at 1200 rpm for 30 min using thermomixer (Eppendorf, Germany). 300 μl ice cold chloroform and 300 μl of ice cold Milli-Q water were added to break phases. The samples underwent 2 min of vigorous vortexing and centrifugation at 12000 rpm for 3 mins. The lower organic phase, containing majority of lipids, was transferred to a fresh tube and kept on ice. The aqueous phase was re-extracted with 500 μl of ice cold chloroform, vortexed again for 2 mins and centrifuged at 12000 rpm for 2 mins. The organic fraction was collected again and combined with the first extraction. Pooled samples were then dried under vacuum using a miVac Duo Concentrator (Genevac Ltd, UK). Samples were stored at −80°C until further analysis by MS. An overview of the lipid extraction procedure is shown in Figure 2-2.

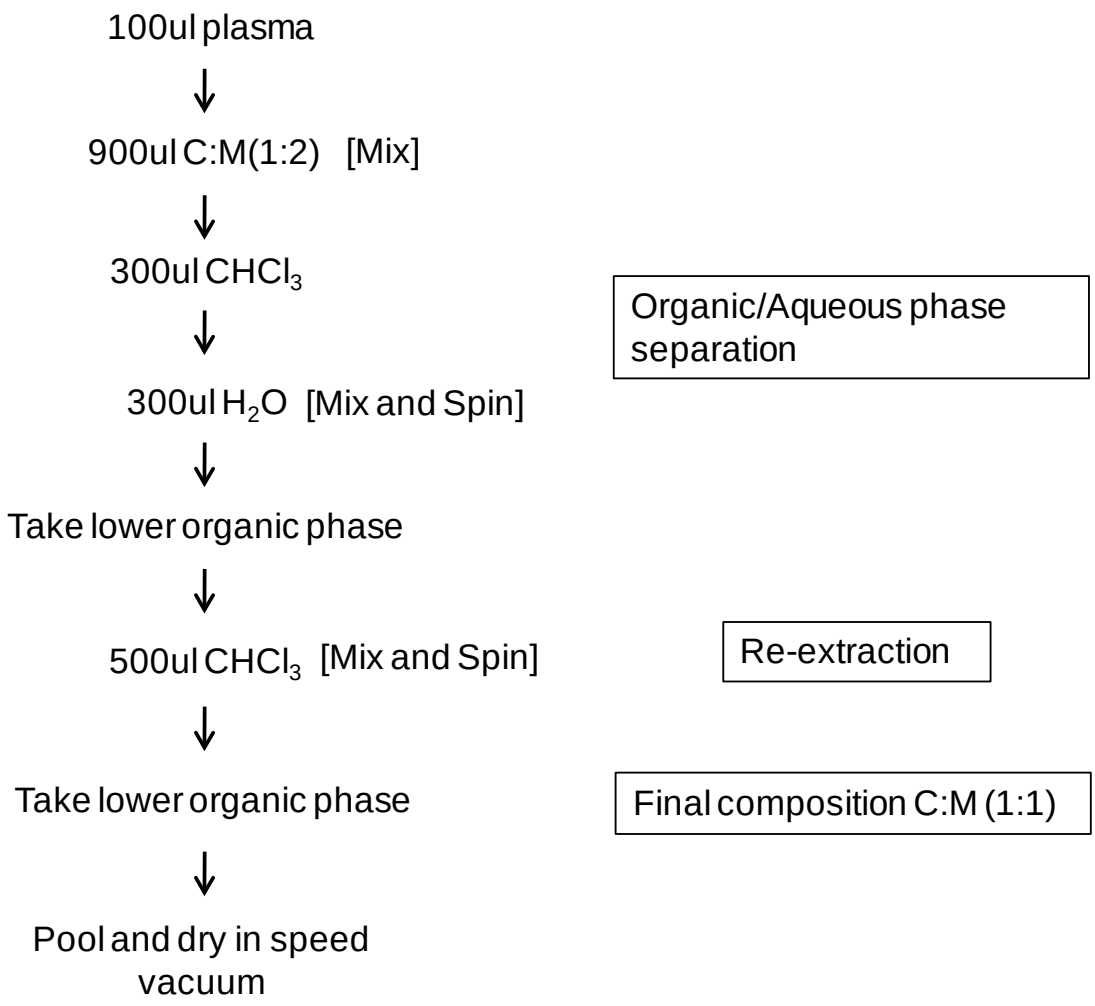


Figure 2-2: Overview of lipid extraction procedure used in this study.*Abbreviations:* $CHCl_3$ is chloroform, C:M is chloroform:methanol, H_2O is water.

2.7 Lipid analysis

Dried lipid extracts were re-suspended in 200 μl of chloroform: methanol (1:1 v/v). 20 μl of resuspended lipid extract was spiked with 20 μl of 2x internal standard mixture that was prepared separately, containing representative standards for each lipid class purchased commercially (Avanti Polar Lipids and Echelon Biosciences, USA). A summary of the standards used in this study is shown in Table 2-1. The internal standard mixture consisted of PA (1,2-dimyristoyl-glycero-3-phosphate or DMPA; final concentration 0.1 μg/ml), PE (1,2-dimyristoyl-glycero-3-phosphoethanolamine or DMPE; final concentration 4 μg/ml), PG (1,2-dimyristoyl-glycero-3-phospho- (1'-rac-glycerol) or DMPG; final concentration 0.1 μg/ml), PI (2-dioctanoyl-glycero-3-phosphoinositol or C8PI; final concentration 1 μg/ml), PS (dimyristoyl-glycero-phosphoserine or DMPS; final concentration 0.1 μg/ml), PC (1,2-dimyristoyl-glycero-3-phosphocholine or DMPC; final concentration 20 μg/ml), SM (lauroyl sphingomyelin or LSM; final concentration 12 μg/ml), Cer (N-heptadecanoyl-D-erythro-sphingosine or C17Cer; final concentration 0.5 μg/ml) and GluCer (D-glucosyl-β1-1‘-N-octanoyl-D-erythro-sphingosine or C8GluCer; final concentration 0.5 μg/ml) species.

20μl of sample was injected using an Agilent 1100 Autosampler system (Agilent Technologies, USA) for MS analysis using a triple quadrupole instrument ABI 4000 (Applied Biosystems, USA). Lipids were ionized by ESI and each individual lipid molecular species was analyzed using MRM

measuring over 100 unique transitions. The samples were first analyzed in negative ion mode measuring transitions corresponding to PA, PE, PG, PI, PS and ganglioside mannoside 3 (GM3) and afterwards PC, SM, Cer and GluCer specific transitions were analysed in the positive ion mode. Chloroform: methanol (1:1) with 2% of 300mM piperidine was used as the mobile phase for lipids ionized in the negative ion mode (piperidine was used as a base to favour the formation of negative ions) and chloroform: methanol (1:1) was used as the mobile phase for lipids ionized in positive ion mode (methanol as a source of protons). Piperidine is a well known base that is used to enhance the signal of polar PL measured by ESI in the negative ion mode (Lytle et al., 2000). Instrumentation parameters for ionization were previously established (Shui et al., 2007). Scan time was fixed at two mins for each ion mode and dwell time was 20ms for negative ion mode and 18ms for positive ion mode respectively. All samples in each batch were analyzed in negative ion mode before switching polarity for positive ion mode analysis. Two chloroform:methanol (1:1) blanks were run between each sample for negative ion mode, and three chloroform:methanol (1:1) with 1% piperidine blanks were run between each sample for positive ion mode.

STLs analysis was done as previously described using APCI LC-MS (Shui et al., 2011a). Briefly, 20 μl of resuspended lipid extract was spiked with 20 μl of 2x internal standard mixture that was prepared separately, containing representative standards for each lipid class purchased commercially (CDN Isotopes Inc, Canada and Medical Isotopes Inc, USA).

The internal standard mixture consisted of cholesterol d6 (final concentration 15 μg/ml), cholesterol ester (final concentration 15 μg/ml), 7-ketone (final concentration 0.5 μg/ml), 7β-hydroxycholesterol (final concentration 0.5 μg/ml), 24-hydroxycholesterol d7 (final concentration 0.5 μg/ml) and zymosterol (final concentration 0.5 μg/ml) species.

Samples were initially separated on Agilent Eclipse XDB-C18 reverse phase column (5 μm, 4.6 x 150 mm) (Agilent Technologies, USA) before injection into the mass spectrometer. Typically, 10 μl of the sample was injected for analysis. The inlet system consisted of a Agilent 1200 series HPLC system (Agilent Technologies, USA). Chloroform-methanol (1:1 v/v) was used as the mobile phase for isocratic elution at a flow rate of 0.5 ml/min with an analysis time of 10.5 mins. The column elutes were measured using heated nebulizer in positive ion mode through a ABI 3200 triple quadrupole instrument (Applied Biosystems, USA). The source temperature was maintained at 500°C. The individual molecular STL species were identified using MRM approach. Dwell time was 30ms. One C:M(1:1) blank was run every 10 samples analyzed for STs in the positive ion mode. Chol and its derivates were measured based on monitoring precursor of *m/z* 161 for Chol, *m/z* 167 for CholE, *m/z* 159 for 7 dehydroChol, 7β-hydroxyChol (7β-OHChol) and 24-hydroxyChol (24-OHChol), and *m/z* 175 for 7-ketoChol.

A summary of the instrument parameters used in this study is shown in Table 2-2. Baseline signal was stable throughout run and varied

within 15% for all analyses. Ion source and instrument parts (needle, needle seat, tubings) were thoroughly cleaned and sonicated with 100% methanol, followed by chloroform: methanol (1:1) prior to and at the end of each batch analysis (total 16 batches for GPLs and SPs; 2 batches for STLs) to ensure minimal contamination between each batch analysis. The instruments were equilibrated with blanks (same as the mobile phase) and 3-5 pooled QC samples were run to ensure system stability prior to sample analysis. Samples were randomized to ensure that all experimental samples were affected to a similar extent, avoiding any bias in subsequent statistical analysis (Want et al., 2010).

A list of all the lipids measured in this study by MRM and their respective collision energies (CE) is given in Supplementary Table 6-1. A sample chromatogram from sterol analysis is given in Supplementary Figure 6-8.

	LIPID	Molecular weight	Mode	Q1	Q3	1:1 Conc of sample:standard	Company
						ug/ml	
1	DMPA	614.767	ESI Neg	591.6	227.3	0.1	AVANTI
2	DMPG	688.845	ESI Neg	665.4	153.0	0.1	AVANTI
3	DMPE	635.853	ESI Neg	634.7	196.1	4.0	AVANTI
4	DMPS	701.844	ESI Neg	678.1	591.6	0.1	AVANTI
5	diC8 PI	608.59	ESI Neg	585.5	241.1	1.0	AVANTI
6	DMPC	677.933	ESI Pos	678.7	184.0	20.0	AVANTI
7	L- SM	646.922	ESI Pos	647.9	184.0	12.0	AVANTI
8	C17 Cer	551.927	ESI Pos	552.6	264.4	0.5	AVANTI
9	C8 Glu Cer	587.44	ESI Pos	588.8	264.4	0.5	AVANTI
10	d6 C18 Cholesterol ester	392.7	APCI Pos	375.3	167.0	15.0	CDN ISOTOPES INC
11	Cholesterol	389.69	APCI Pos	375.3	161.0	15.0	CDN ISOTOPES INC
12	7-Ketocholesterol-25,26,26,26,27,27,27-d_7	407.69	APCI Pos	408.3	175.0	0.5	CDN ISOTOPES INC
13	7-beta hydroxycholesterol	409.7	APCI Pos	392.3	159.0	0.5	CDN ISOTOPES INC
14	24-hydroxycholesterol	402.66	APCI Pos	374.3	159.0	0.5	MEDICAL ISOTOPES INC
15	Zymosterol	389.67	APCI Pos	372.3	220.2	0.5	MEDICAL ISOTOPES INC

Table 2-1: Summary of the list of standards used in this study.

	Positive ion mode	**Negative ion mode**	**Sterols**
Curtain Gas	20L/min	20L/min	10L/min
Ion spray voltage	5500V	-4500V	5500V
Temperature	300˚C	300˚C	500˚C
Ion source gas 1	30L/min	40L/min	60L/min
Ion source gas 2	40L/min	30L/min	35L/min
Flow rate	0.25ml/min	0.25ml/min	0.50ml/min

Table 2-2: Summary of instrument parameters used in this study.

2.8 Data analysis and bioinformatics

2.8.1 Data extraction and pre-processing

Data for each GPL and SP lipid species were extracted using an in-house developed MATLAB algorithm (Bowen Li, National University of Singapore) (Math Works, USA) to obtain signal intensity. Extracted data in text format was imported into MS Excel (Microsoft, USA) and signal intensity was normalized to the representative spiked internal standards to obtain relative measurements as previously described (Brugger et al., 1997). GM3 levels were normalized to PI internal standard. For STLs species, signal intensities were extracted according to their retention time using an in-house developed MATLAB algorithm (Bowen Li, National University of Singapore) (Math Works, USA) and imported into MS Excel (Microsoft, USA). Relative lipid concentrations were calculated by normalization to representative spiked internal standards.

2.8.2 Technical variation study

COV were measured on raw data normalized to spiked internal standard. COV is calculated as the ratio of standard deviation to the average intensity, multiplied to 100 to get the values in percentage, as previously described (Ekroos et al., 2010). Instrument, process and subject COV are calculated as described in Chapter 3.

2.8.3 Biological variation study

2.8.3.1 Pilot study

One-way analysis of variance (ANOVA) test of significance (two-tailed, $p<0.05$) with Bonferroni adjustment was used to analyse GPLs and SPs which were significant across the three timepoints, i.e. overnight fasting, two hours post breakfast and two hours post lunch for both pilot and validation rounds.

2.8.3.2 Main study

Batch effect correction and normalization was done using pooled QC samples and is further described in Chapter 3. The effectiveness of the batch effect correction method was represented by principal component analysis (PCA) plots where the largest variance was explained by the first principal component. Cer and SM (*m/z* 184) are excluded from biological variation analysis.

QC normalized lipids were tested for significant difference across ethnicity using student's t-test (two-tailed, $p<0.05$) in MS Excel (Microsoft, USA). For global overview of pairwise comparisons of all lipid species across ethnicity, p-values obtained using student's t-test (two-tailed, $p<0.05$) were ranked across the three comparisons (Chinese vs Malay, Chinese vs Indian and Indian vs Malay). These ranks for each lipid species were added to give a

rank score. The rank score was then sorted and the median rank score with median absolute deviation was obtained. Difference between the combined rank and combined median rank score was calculated. The lipids were ordered based on the highest rank difference, with highest rank difference being most significant. The data is represented in a scatter plot plotted as log p-value (primary y-axis) and combined rank score (secondary y-axis, trendline) against lipid species (x-axis) using MS Excel (Microsoft, USA).

For hierarchical clustering analysis, z-score of each lipid was first calculated after excluding lipids significant due to gender (t-test, two-tailed, $p<0.05$). Remaining lipids were clustered by hierarchical clustering using Pearson correlation distances (uncentered) with average-linkage implemented in the open ware clustering software Cluster3.0 (Eisen et al., 1998). The calculated dendrogram was visualized using the open ware software Java TreeView (Saldanha, 2004) with the colours red (upregulation), green (downregulation) and black (no change). Patterns in identified clusters were further analysed by MS Excel (Microsoft, USA).

Lipid species were analyzed for ethnic differences according to carbon chain length, degree of unsaturation and ester/plasmalogen (a/p) species using student's t-test (two-tailed, $p<0.05$) in MS Excel (Microsoft, USA) and significant differences were represented by density plots with 95% confidence interval of the mean of each ethnic population using R (Team, 2008) (by Bowen Li (National University of Singapore). All lipid species with an ester bond (usually at sn-1 position) are denoted as a lipid species, while

all lipid species with an ether bond or plasmalogen vinyl ether bond are denoted as p lipid species. Density plots are plotted as density (frequency) against normalized lipid intensity. Gender differences determined by a significant p-value (t-test, two-tailed, $p<0.05$) were analysed separately for male and female, represented by bar charts.

QC normalized lipids were also tested for significant difference across ethnicity using linear regression with type II ANOVA (two-tailed, $p<0.05$) in R (Team, 2008), taking into account age and gender (by Bowen Li (National University of Singapore)), according to carbon chain length, degree of unsaturation and a/p species. Statistical significance across ethnicity is reported in tabular format.

Statistical validation for observed ethnic differences was done using permutation test in R (Team, 2008) (by Bowen Li ,National University of Singapore). An arbitrary cutoff of $p>0.005$ was considered significant (i.e. five out of 1000 permutations). A p-value of 0.001 indicates either zero or one out of 1000 permutations is likely to produce the a significant ethnic dependent difference due to random chance.

Clinical measurements of HDL and LDL were correlated with p lipid species (pPE and pPC) using Pearson correlation in MS Excel (Microsoft, USA). Data is represented by scatter plots with cumulative distribution (y-axis) against pearson correlation (x-axis).

CHAPTER 3

TECHNICAL VARIATION STUDIES

3.1 Study design methodology

While mass spectrometry has been proven to be a useful tool for profiling lipid metabolites, there has been an increasing need to address issues on accurate determination and quantification of lipids in large sample numbers (Dunn et al., 2012). Nowadays, there is a shift from few lipid species in a limited sample number to larger epidemiological scale studies (such as the HUSERMET Consortium (Dunn et al., 2011; Zelena et al., 2009)), potentially requiring many hours of continuous sample analysis (Bijlsma et al., 2006; Zelena et al., 2009). Such studies can only be carried out if high analytical reproducibility can be assured so that batches can be effectively compared across each other, irrespective of the time they were analyzed (Dunn et al., 2012).

Data variation can arise from several sources, including experimental sources (e.g., due to the analytical or sampling processes), biological and genetic variations (Gika et al., 2012). Currently, literature is

focusing on guidelines that need to be followed to ensure reproducibility of data and minimize technical variation (Dunn et al., 2012; Hammad et al., 2010; Vuckovic, 2012). This is an important part of any study design so that true biological variation can be effectively observed (Lindon et al., 2005). In the first part of this chapter, the factors taken into consideration for the design will be outlined (Figure 3-1), while the following sections will describe the technical and biological variation studies performed.

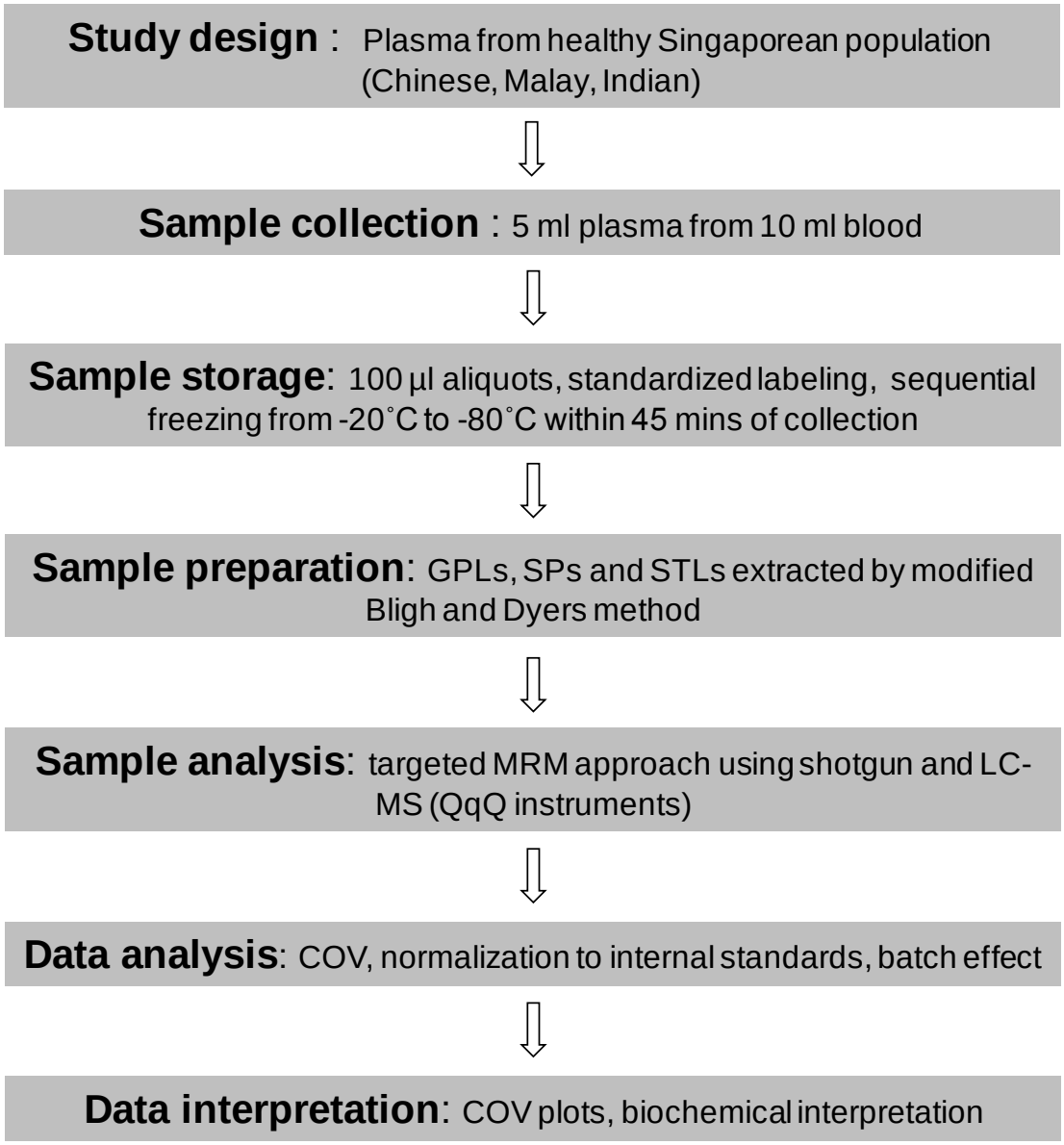


Figure 3-1: Schematic diagram representing the methodological approach followed in this study.

3.1.1.1 Results and Discussion

To start with, it was necessary to standardize sample collection and processing. Three main parameters were considered for further investigation. Firstly, samples were collected after fasting state as dietary effects greatly impact lipidomics studies (Teahan et al., 2006). All participants were required to give blood samples between 8-10 am after a minimum of 10 hours overnight fasting (Altmaier et al., 2008). This was also to account for variation caused by circadian rhythm. Previous literature has shown that the time of the day on which the samples were taken influences inter-individual sample variation from healthy volunteers (Radich et al., 2004; Whitney et al., 2003). Furthermore, recent literature highlighted how circadian rhythm affected metabolic pathways (Dallmann et al., 2012; Minami et al., 2009).

Secondly, the method to obtain efficient and optimal yields of plasma was optimized. It was eventually optimized to spin once at 2200xg for 15 mins at 4˚C. The choice of collection tubes and anticoagulant, BD Vacutainer EDTA tubes, was based on those widely used in published large-scale human studies (Hammad et al., 2010; Hicks et al., 2009). Thirdly plasma was stored at -80˚C which is the universally accepted method for long-term storage of human biological samples (U.S. Department of Health and Human Services, 2001). These samples were transferred to long term storage cryotubes and were initially frozen to -20˚C within 45 mins of collection. Subsequently, they were transferred to -80˚C for longer storage. This was especially important for lipid analysis as quick freezing minimizes

degradation ensuring a more robust sample analysis (Vuckovic, 2012). Participants were required to give additional blood samples if red blood cells lysis occurred during plasma processing. Lysed plasma was discarded and not used for subsequent analysis.

As several studies would be carried out, all processed plasma was separated into 100 µl aliquots for subsequent lipid extraction and analysis. Every sample underwent the same number of freeze-thaw cycles for a similar time period. Several studies looked at the effect of freeze-thaw cycles and showed spectral variations depending on storage conditions (Bernini et al., 2011; Hammad et al., 2010; Teahan et al., 2006). In this study, all samples analyzed did not exceed three freeze-thaw cycles prior to processing and analysis, which was in accordance to recommended guidelines (U.S. Department of Health and Human Services, 2001).

3.2 Overview of technical variation studies

After optimizing the study design, the next step was to look at technical variation caused by sample processing and instrumentation used for analysis. Three studies were carried out in order to assess technical variation in 360 samples collected from the Singaporean population. An overview of the studies performed is shown in Figure 3-2.

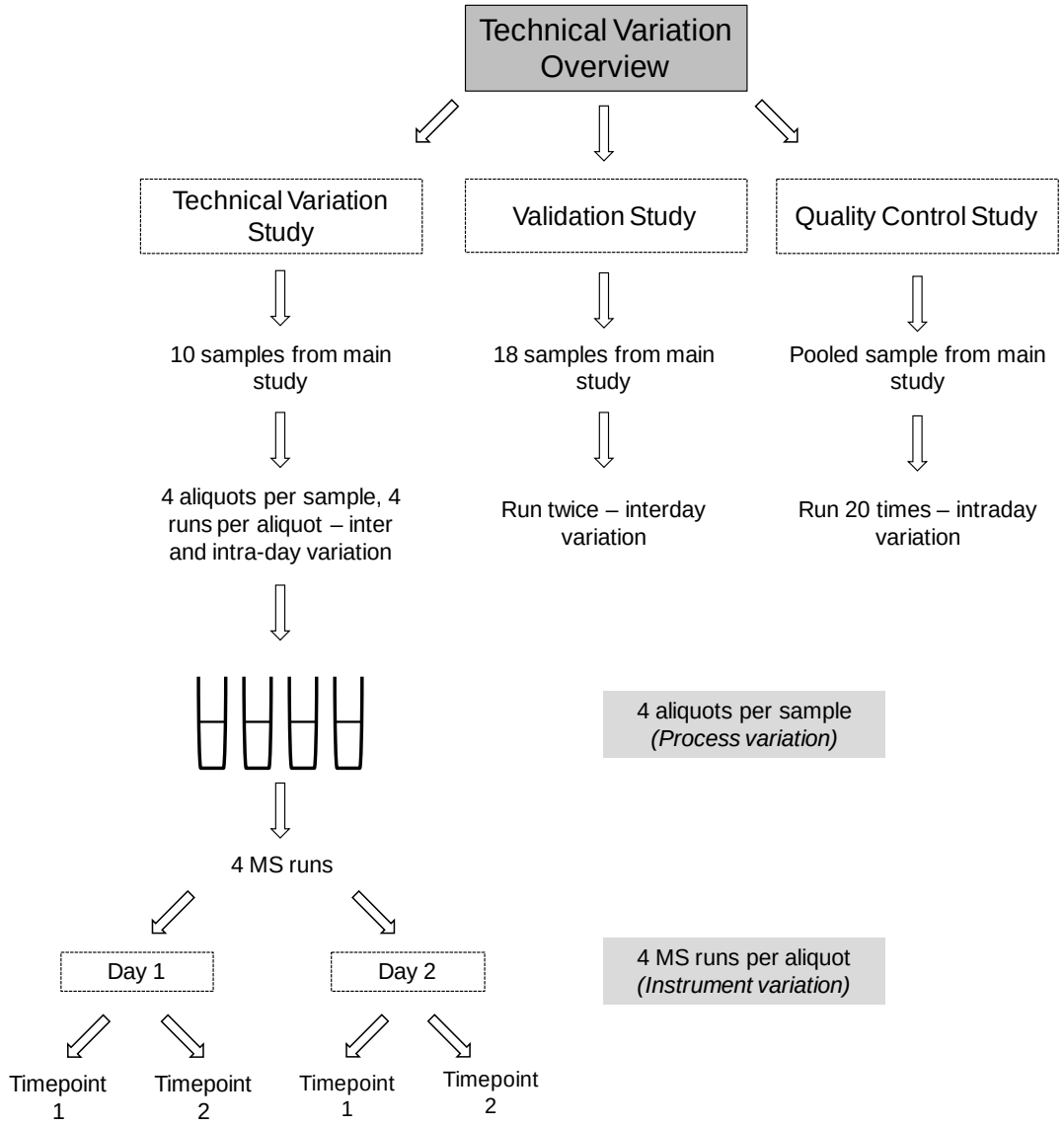


Figure 3-2: Overview of the 3 technical variation studies carried out.

3.2.1 Results

3.2.1.1 Technical variation study (TVS)

In this section, technical variation data obtained from 10 individuals (four Chinese, three Indian and three Malay individuals) from the main study are shown. Two main categories of variation were assessed: process variation and instrument variation, as previously described in Chapter 1 and 2. An overview of the study design is shown in Figure 3-2.

The measurement of choice was COV as previously described. COV is a widely used statistical measure of variation between comparison groups and was first described by Horwitz (Horwitz, 1982). Our technical variation data were analyzed based on process, instrument and subject variation, and represented by COV vs intensity plots. Data is presented for all the lipid classes as shown in Figure 3-3

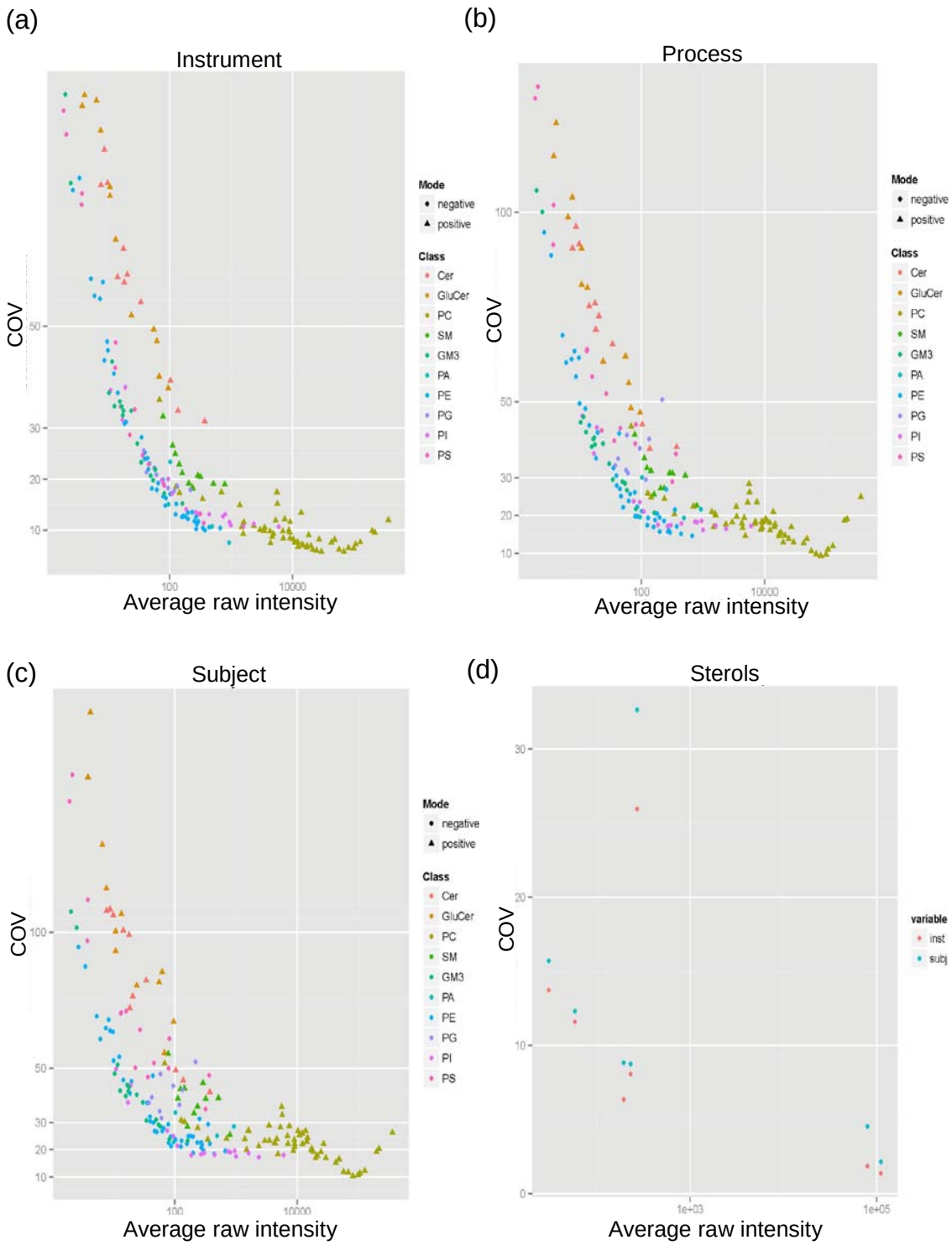


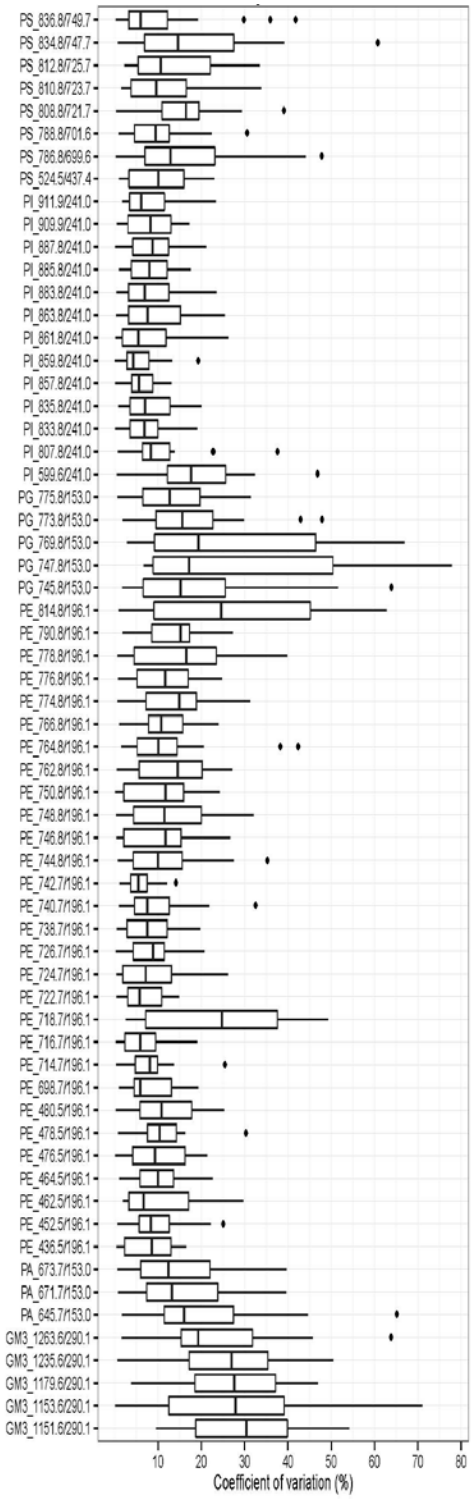
Figure 3-3 COV vs log intensity plots for GPLs and SPs for (a) Instrument variation (b) Process variation (c) Subject variation, for process, instrument and subject variation for (d) STLs.Negative and positive indicate the MS ion mode in which the lipid classes were analyzed. *Abbreviations:* inst is instrument variation, subj is subject variation.

From Fig 3-3 COV vs log intensity plots for the various lipid classes show a dynamic range of COV. As process COV includes instrument variation, the COV is higher than instrument COV alone. The difference in COV between instrument and process variation is small, indicating that major source of technical variation is coming from instrument rather than sample processing. Some classes intrinsically exhibit a higher COV (like Cer,GluCer and PA) whereas other classes like PE exhibit a lower COV. The COV for STLs is much lower than GPLs and SPs, which could be due to the methodology used for analysis, i.e. LC-MS vs shotgun approach. The difference in COV between instrument and process variation is small, indicating that major source of technical variation is coming from instrument rather than process.

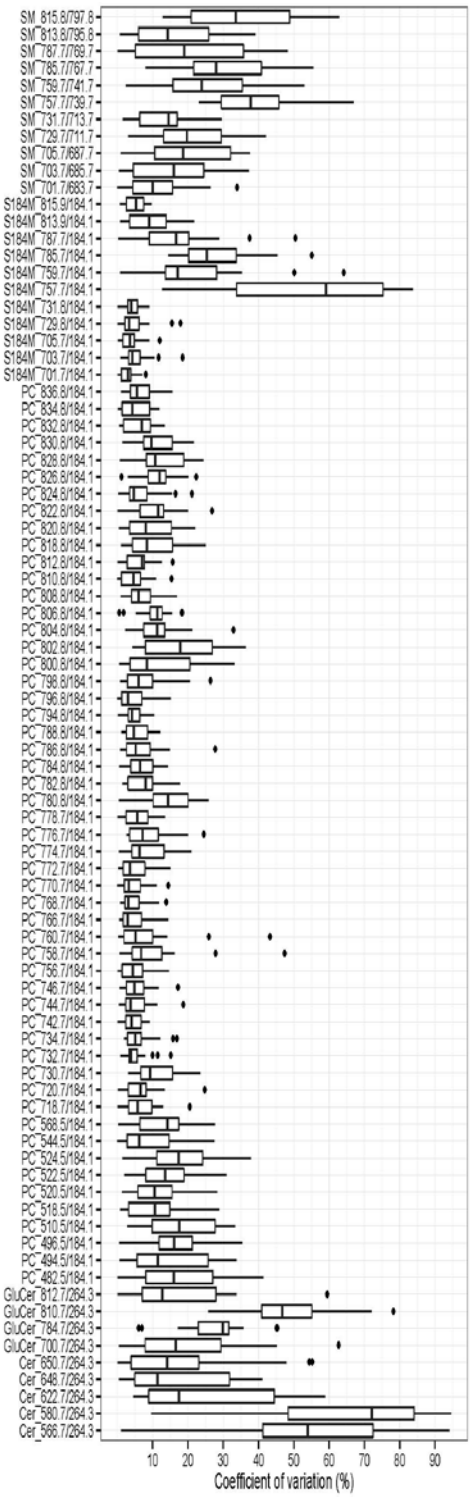
3.2.1.2 Validation study (VS)

As TVS was analyzed separately from the main study, technical variation from the main study was assessed by analyzing 18 randomly chosen samples as previously described. Lipids to be measured were revised by removing lipids that showed unstable measurements and a high degree of variation in COV. Data from this revised list of lipids are represented in the form of boxplots as shown in Figure 3-4.

(a)


PS_836.8/749.7
PS_834.8/747.7
PS_812.8/725.7
PS_810.8/723.7
PS_808.8/721.7
PS_788.8/701.6
PS_786.8/699.6
PS_524.5/437.4
PI_911.9/241.0
PI_909.9/241.0
PI_887.8/241.0
PI_885.8/241.0
PI_883.8/241.0
PI_863.8/241.0
PI_861.8/241.0
PI_859.8/241.0
PI_857.8/241.0
PI_835.8/241.0
PI_833.8/241.0
PI_807.8/241.0
PI_599.6/241.0
PG_775.8/153.0
PG_773.8/153.0
PG_769.8/153.0
PG_747.8/153.0
PG_745.8/153.0
PE_814.8/196.1
PE_790.8/196.1
PE_778.8/196.1
PE_776.8/196.1
PE_774.8/196.1
PE_766.8/196.1
PE_764.8/196.1
PE_762.8/196.1
PE_750.8/196.1
PE_748.8/196.1
PE_746.8/196.1
PE_744.8/196.1
PE_742.7/196.1
PE_740.7/196.1
PE_738.7/196.1
PE_726.7/196.1
PE_724.7/196.1
PE_722.7/196.1
PE_718.7/196.1
PE_716.7/196.1
PE_714.7/196.1
PE_698.7/196.1
PE_480.5/196.1
PE_478.5/196.1
PE_476.5/196.1
PE_464.5/196.1
PE_462.5/196.1
PE_452.5/196.1
PE_436.5/196.1
PA_673.7/153.0
PA_671.7/153.0
PA_645.7/153.0
GM3_1263.6/290.1
GM3_1235.6/290.1
GM3_1179.6/290.1
GM3_1153.6/290.1
GM3_1151.6/290.1
10 20 30 40 50 60 70 80
Coefficient of variation (%)

(b)


SM_815.8/797.8
SM_813.8/795.8
SM_787.7/769.7
SM_785.7/767.7
SM_759.7/741.7
SM_757.7/739.7
SM_731.7/713.7
SM_729.7/711.7
SM_705.7/687.7
SM_703.7/685.7
SM_701.7/683.7
S184M_815.9/184.1
S184M_813.9/184.1
S184M_787.7/184.1
S184M_785.7/184.1
S184M_759.7/184.1
S184M_757.7/184.1
S184M_731.6/184.1
S184M_729.6/184.1
S184M_705.7/184.1
S184M_703.7/184.1
S184M_701.7/184.1
PC_836.8/184.1
PC_834.8/184.1
PC_832.8/184.1
PC_830.8/184.1
PC_828.8/184.1
PC_826.8/184.1
PC_824.8/184.1
PC_822.8/184.1
PC_820.8/184.1
PC_818.8/184.1
PC_812.8/184.1
PC_810.8/184.1
PC_808.8/184.1
PC_806.8/184.1
PC_804.8/184.1
PC_802.8/184.1
PC_800.8/184.1
PC_798.8/184.1
PC_796.8/184.1
PC_794.8/184.1
PC_788.8/184.1
PC_786.8/184.1
PC_784.8/184.1
PC_782.8/184.1
PC_780.8/184.1
PC_778.7/184.1
PC_776.7/184.1
PC_774.7/184.1
PC_772.7/184.1
PC_770.7/184.1
PC_768.7/184.1
PC_766.7/184.1
PC_760.7/184.1
PC_758.7/184.1
PC_756.7/184.1
PC_746.7/184.1
PC_744.7/184.1
PC_742.7/184.1
PC_734.7/184.1
PC_732.7/184.1
PC_730.7/184.1
PC_720.7/184.1
PC_718.7/184.1
PC_568.5/184.1
PC_544.5/184.1
PC_524.5/184.1
PC_522.5/184.1
PC_520.5/184.1
PC_518.5/184.1
PC_510.5/184.1
PC_496.5/184.1
PC_494.5/184.1
PC_482.5/184.1
GluCer_812.7/264.3
GluCer_810.7/264.3
GluCer_784.7/264.3
GluCer_700.7/264.3
Cer_650.7/264.3
Cer_648.7/264.3
Cer_622.7/264.3
Cer_580.7/264.3
Cer_566.7/264.3
10 20 30 40 50 60 70 80 90
Coefficient of variation (%)

(c)

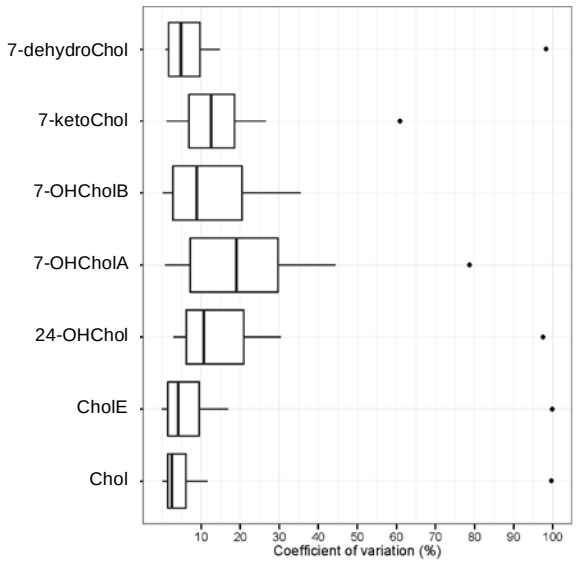


Figure 3-4: Boxplot representations of COV for technical variation of validation batch samples for lipids analyzed in (a) negative ion mode (PE,PA,PI,PG,PS,GM3) (b) positive ion mode (PC, SM, Cer, GluCer) (c) STLs. Boxes represent lower quartile (25th percentile), median and upper quartile (75th percentile). Linear boundaries represent the lowest and highest value for each lipid. Dots are outliers (lipid values that are more than 1.5 times the interquartile range).

The COV is within 20%for most GPLs and SPs, though SPs exhibit a higher median COV. Some classes still exhibit a high COV, particulary for low abundant species like GM3, Cer and PA. STLs show lower COV when compared to GPLs and SPs, similar to TVS. Furthermore, the higher abundant species like Chol and PC show both a lower COV and a lower degree of variation in COV. Those low abundant species with high COV were excluded from biological variation analysis.

3.2.1.3 Quality control study (QCS)

Pooled samples from the main study were used as QC samples to assess intraday variation, as previously described. The data obtained are represented by COV plots as shown in Figure 3-5.

As we can see from Figure 3-5, most lipids show a COV within 20%. Some classes exhibit a higher COV like the GM3 and Cer, whereas some classes exhibit a wider dynamic range of COV like PE, PS and SM. PC species have the lowest COV measurements, and this could be due to their high abundance in human plasma.These findings were also observed in the TVS and VS. The findings of these 3 studies will be discussed in section 3.3.

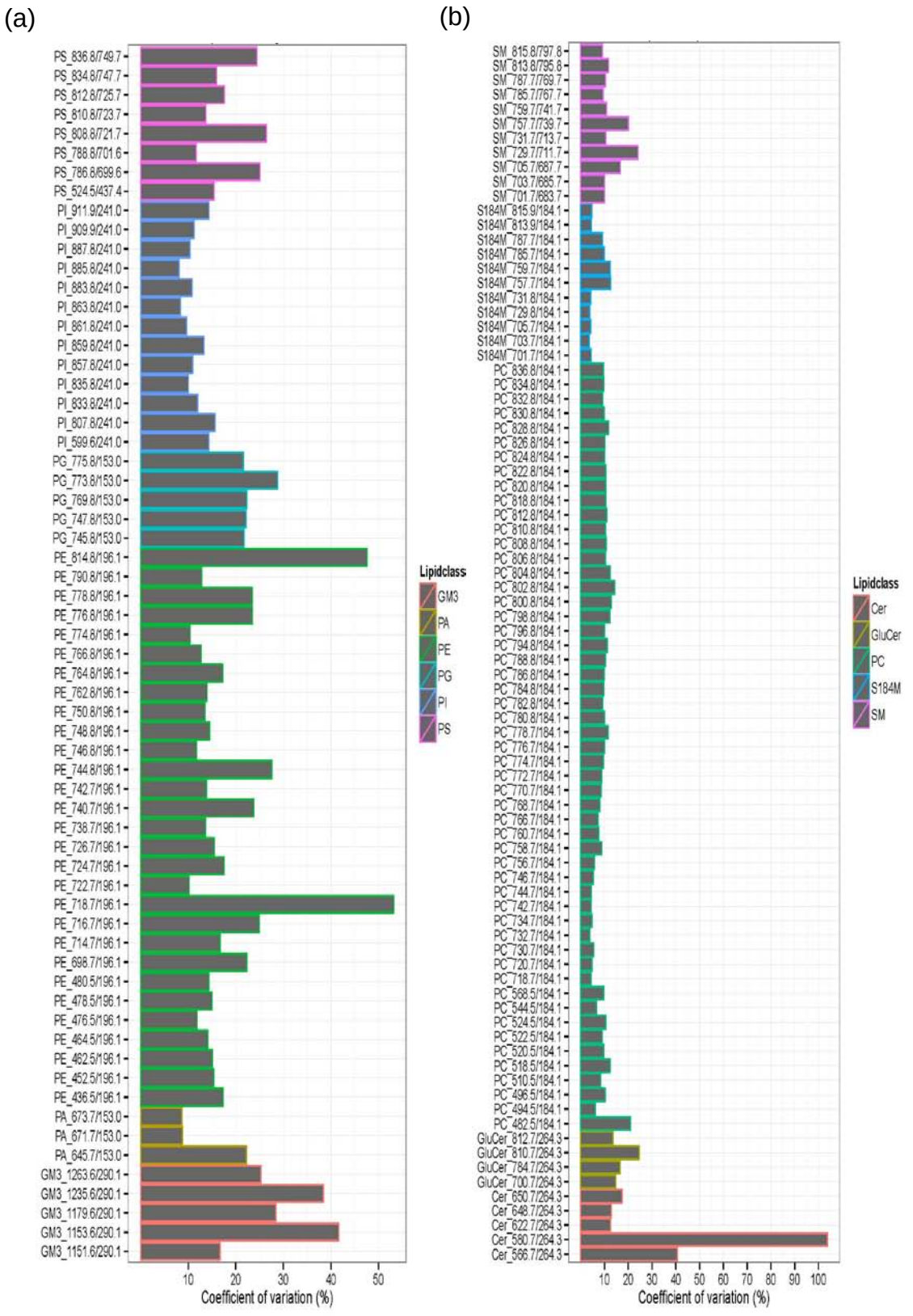


Figure 3-5: COV plots for QCS showing the COV for GPLs and SPs analyzed in (a) negative ion mode and (b) positive ion mode.

3.3 Discussion

Conceptual understanding of various technical aspects of any study contributing to variation is important. Technical variation influences the robustness of an experiment, right from initial stages of sample acquisition to processing and final data analysis (Bando et al., 2010; Gika et al., 2012). In this study, we performed a large scale human plasma sample collection, thus it was imperative to be able to understand the various sources of variation in the study design and to try to account for them as much as possible.

All 3 studies (technical variation study, validation study and quality control study) showed similar trends in COV across each lipid class (Figures 3-4 to 3-6), indicating a consistent performance when analyzed using MS. As we can see from Figure 3-4, all the COV vs log intensity plots follow a general trend of high COV at low mean intensity and low COV at high mean intensity. This can be explained considering the limit of detection (LOD) or threshold of signal in the mass spectrometer. When the lipid species such as Cer have a low S/N ratio (Quehenberger et al., 2010), it fluctuates simply due to difficulties of the instrument in establishing a signal distinguishable from background noise (Trotzmuller et al., 2011). Noise is contributed from a variety of sources, mainly from the instrument itself (Ekroos et al., 2010). Thus, these lipid species have a high COV (Horwitz, 1982). On the other hand, lipid species such as PC having high S/N ratio are easier to distinguish from the background noise. This results in more stable measurements, hence a lower COV (Quehenberger et al., 2010; Sato et al., 2010).

The addition of piperidine as a base enhances the ionization of lipid species measured in the negative ion mode, greatly improving the signal of certain lipid classes like PE (Lytle et al., 2000; Shui et al., 2007). It can be seen that PC and PE measured from the positive and negative ion modes, respectively, have a larger mean intensity when compared to other lipid classes. Recent literature showed that the majority of GPLs in plasma consist of PE and PC, in accordance to what we observe. (Quehenberger et al., 2010). Furthermore, it has been reported that PE and PC are dominant GPL collectively accounting for 75 mol% of total GPL mass in eukaryotic membranes and human plasma (Han and Gross, 2005; Quehenberger et al., 2010). In this study, adducts of lipid species with cations or anions were not analyzed. No acid or salts were added to solvents used for lipid extraction and sample analysis. Despite this, chlorinated adducts (due to chloroform) and sodiated adducts due to sample processing may be present and detectable in the negative and positive ion modes respectively, especially for predominant lipid species like PC and SM (Shui et al., 2007). However, this study is based on relative quantitation and not absolute quantitation, thus the total measurement of different adducts is not necessary. Furthermore, it is estimated that the ratio of each molecular ion to its adduct is consistent throughout all the analyses, allowing us to reliably report differences between the samples. It is also important to note that PE and PC have multiple subclasses and also play diverse important roles in cellular function (E.Vance, 2008). Other lipid classes like PI and PS exist in modest amounts in eukaryotic membranes and human plasma (Han and Gross, 2005; Quehenberger et al., 2010).

While PE and PC exist in high amounts in the human plasma, they are detected at intensities spanning a few orders of magnitude within their lipid classes (Figures 3-4 to 3-6). Differences in acyl chain composition and degree of unsaturation of PE and PC species contribute to differences in ionization efficiencies (Han and Gross, 2005). Most PC and PE have reasonably high intensities, some species exhibit lower intensities. Lower intensities could result from lower concentrations in the human plasma (Han and Gross, 2005). In the case of lyso species, they have one FA chain and usually exist in lower amounts in the human plasma as observed in our analyses (Quehenberger et al., 2010). Thus, the stable and reliable detection of these lipids may be more difficult.

Some lipid classes exhibited a variable range of COV despite giving rise to a narrow range of intensity measurement in the MS, such as the GM3. This could be due to the incorrect lipid standard used for normalization. Here, PI standard was used to normalize GM3 but this may not be the best method as a proper GM3 standard may give rise to more accurate measurement of GM3 species. Due to the constraints in this study, PI was the best available standard for measuring GM3.

A similar trend was observed for STLs although not as distinct as GPLs and SPs. STLs are one of the most abundant lipid classes present in the human plasma (Quehenberger et al., 2010). Only 7 species were measured (Figure 3-4d), thus the COV *vs* log intensity plot is unable to show a clear trend. When looking at each individual species measured, Chol and

CholE have very high mean intensities with low COV range, whereas sterol intermediate 7-dehydroChol which has a low intensity exhibits a high COV range. This can be seen by the raw data which exhibits unstable signals, i.e. a low S/N ratio, indicating difficulty in distinguishing 7-dehydroChol from background noise. However, this is in line with published literature where Chol and CholE are between the most abundant STLs in human plasma (Quehenberger et al., 2010). The lower COV for STLs when compared to GPLs and SPs could be due to prior chromatographic separation of STLs (using LC) resulting in better sensitivity (Nygren et al., 2011).

Based on these 3 technical variation studies, 2 decisions were made before progressing to the main study; firstly, SM analysis in the main study was focused on the loss of water transition (*m/z* 18) rather than *m/z* 184 though the COV was higher. This was because the intensity counts were more accurate and less likely to be coming from PC as previously reported (Liebisch et al., 2004; Pulfer and Murphy, 2003). Secondly, Cer species were excluded from further analysis due to unstable intensities and high COV. Certain lipid species with high COV above 30% were removed and the remaining lipid species were analyzed using the existing method with no further changes.

As there are a large number of samples to be analyzed for the main study, we had to account for batch effect due to analysis of samples in multiple batches. As seen by the TVS study, multiple batches were run and COV obtained was above 30% for some lipids. The employment of QC

samples and batch effect correction methods was recognized and subsequently implemented in the VS and QCS. Clearly, there is an improvement in the COV, with most lipids exhibiting a COV of less than 20%. Furthermore, the difficulty in obtaining reproducible data for large scale projects using MS has been recognized (Dunn et al., 2012). Hence, study design plays a critical role in these cases. Data from the QCS helped in the study design and multiple QC samples were incorporated into each batch. The concept of batch variation and normalization is further discussed in section 3.4.

3.4 Batch variation and normalization

With increasing advancements in technology, we are able to better measure and characterize a large number of lipids with higher precision and accuracy (Watson, 2006). Despite rapid advancements in technological capabilities, often the disadvantages that this may bring have been overlooked so far. While the dimensions of complexity in data acquisition and analysis increase, the basic question of solving the problem of batch effects remains. Batch effects are sub-groups of measurements that have qualitatively different behavior across conditions and are unrelated to the biological or scientific variables in a study (Leek et al., 2010). For example, batch effects may occur if a subset of experiments was analyzed on different days, if two technicians were responsible for different subsets of the experiments, and if two different lots of reagents or instruments were used

(Leek et al., 2010). This can lead to systematic differences between measurements of different batches (Luo et al., 2010).

In these days, the focus on larger population studies is becoming increasingly popular in both genomics and metabolomics studies due to its ability to detect meaningful trends and associations between various groups of comparison (Gieger et al., 2008; Hicks et al., 2009). In some cases, significant associations are difficult to replicate in independent populations. It is being recognized that significant trends and associations that have been reported could in fact be spurious association simply due to the fact that they may be affected by other factors involved in the study design (Suhre and Gieger, 2012). Specific examples have been documented in published studies (Petricoin et al., 2002; Spielman et al., 2007) in which the biological variables were extremely correlated with technical variables, which subsequently led to serious concerns about the validity of the biological conclusions (Akey et al., 2007; Baggerly et al., 2004).

In order to address the issue of batch effect, normalization procedures are often necessary prior further analyses (Leek et al., 2010). Normalization is a data analysis technique that adjusts global properties of measurements for individual samples so that they can be more appropriately compared (Leek et al., 2010) While data normalization may sometimes exacerbate technical artifacts in the dataset, one has to note that different procedures may work differently based on the study design and type of dataset involved. There is still no universally accepted method of batch effect

correction and normalization procedures for lipidomics dataset. One of the well studied methods of normalization using large-scale data applies a ratio-based method based on inclusion of calibration samples (like QC) in each batch (Luo et al., 2010). The number of calibration samples is recommended to be around three to five samples per batch (Kamleh et al., 2012; Luo et al., 2010). In this study we normalized our dataset using QC samples (QC) consisting in a pooled plasma sample mix (from five samples from the main study) representative of the sample population being studied. The pooled plasma samples were extracted and the reconstituted lipid extracts were pooled, aliquoted and frozen at minus 80 degrees in vials to be analyzed separately for each batch. This was to ensure that each QC vial underwent the same number of freeze-thaw cycles for every batch analysis. During each batch analysis, QC samples were analyzed three times, at the beginning, middle and end of the batch analysis. QC samples were used to normalize all the samples across different batches. An overview of the normalization procedure used in this study is shown in Figure 3-7. The application of this normalization procedure to the main study samples and effectiveness of the batch normalization procedure was checked by looking at the PCA plot before and after normalization and is shown in Chapter 4.

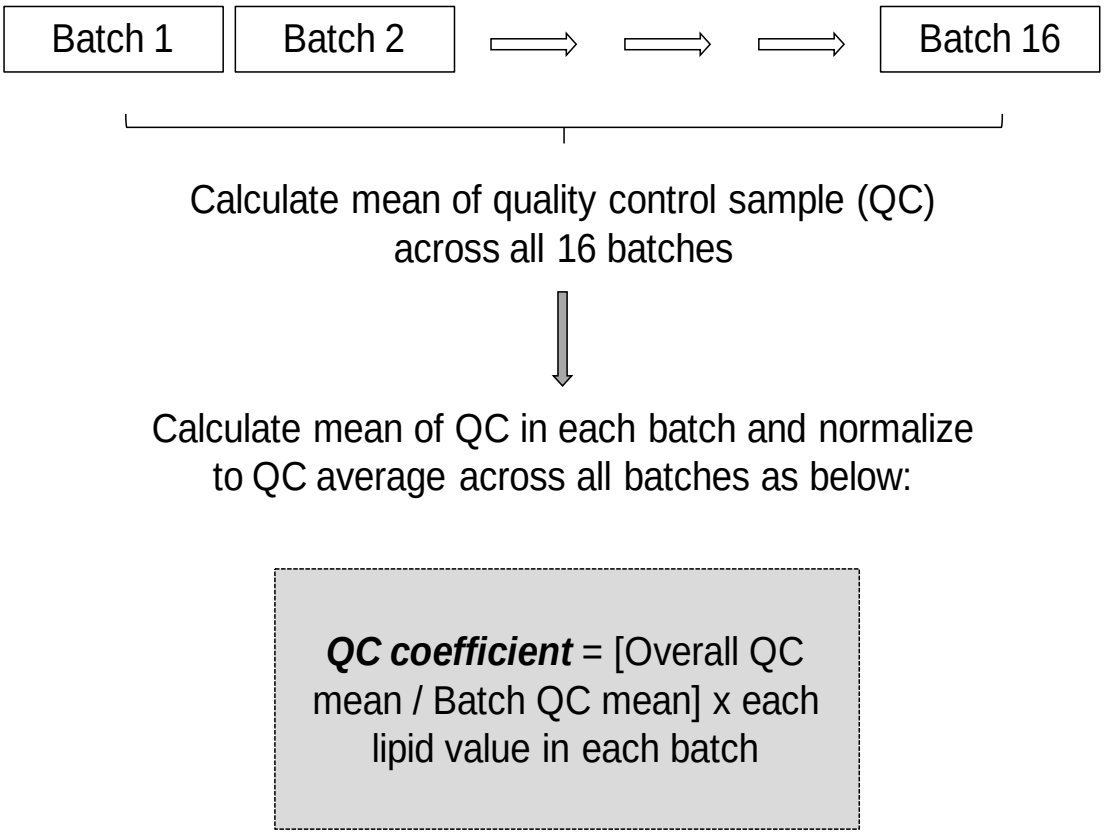


Figure 3-6: Outline of the normalization procedure used in this study.

CHAPTER 4

BIOLOGICAL VARIATION IN THE SINGAPOREAN POPULATION

4.1 Pilot study

Taking multiple samples from a subject would allow discrimination between inter- and intra-individual variations and provide a more robust baseline (Teahan et al., 2006). However, due to logistical constraints, the main study of 360 subjects was only collected at one time-point, i.e. after overnight fasting as discussed in section 4.2.

The results of small scale pilot study, aimed at investigating the effect of multiple sampling from the same individual, for example, the effect of time-point collection after a standardized meal, are discussed below.

4.1.1 Six PE and one PC lipid species show time-dependent changes for both pilot and validation studies

One-way ANOVA (two-tailed, $p<0.05$) with Bonferroni correction was used to analyze all the measured GPLs and SPs and identify lipid species that significantly changed across three time-points, i.e. overnight fasting, two hours post breakfast and two hours post lunch. The abundance of six PE (containing two lyso PE) and one lyso PC lipid species identified to be significantly different across the three time-points in both pilot and validation rounds were plotted as shown in Figure 4.1.

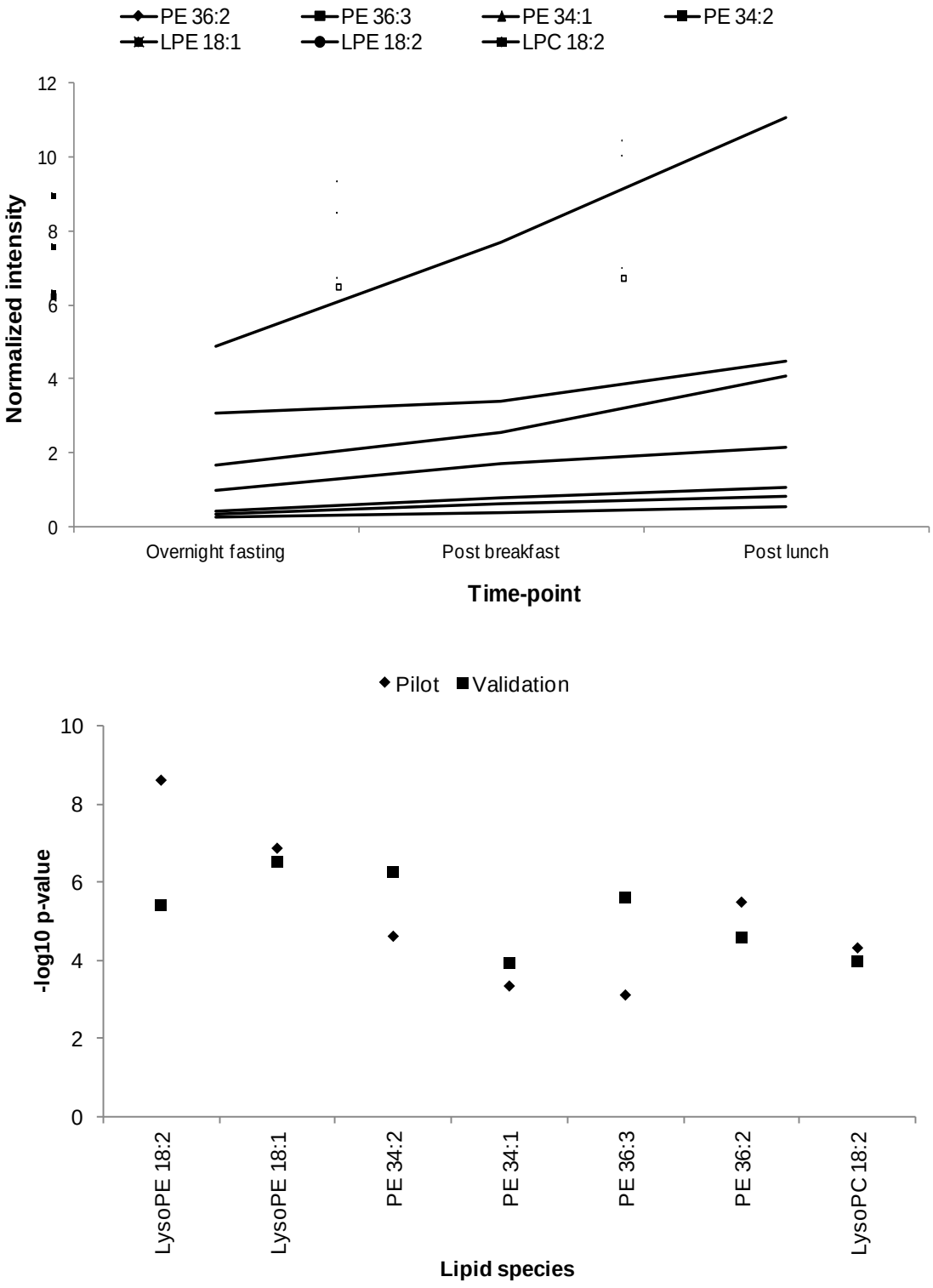


Figure 4-1: Line plot showing linear trend (top panel) and scatter plot showing –log10 p-value (bottom panel) of the seven significant lipids from one-way ANOVA analysis across the three time-points for both pilot and validation studies.

A time-dependent increasing trend was observed in all subjects, irrespective of gender and ethnicity, and consistent with meal intake (data not shown). Although the observed changes could be due to differences in their lifestyle factors (like dietary habits), according to published reports, a single day of dietary standardization is sufficient to provide all of the normalization required within strict controls as the subjects remained stable in their metabolic space after 24 hours (Winnike et al., 2009). Thus, it is possible that significant lipid changes observed in this study could be metabolically important.

The significant lipid species identified were lyso PE 18:1 and 18:2, PE 34:1, PE 34:2, PE 36:2, PE 36:3 and lyso PC 18:2. These species are commonly found in circulating lipoproteins (Dashti et al., 2011). The majority of these lipids are esters which consist of at least one C18 FA chain carrying one to three double bonds (e.g. oleic acid and linoleic acid) (Quehenberger et al., 2010). Linoleic acid is an essential FA which cannot be synthesized by mammalian cells hence, need it has to be obtained by the diet, primarily from plant oils (E.Vance, 2008). Since most mammalian PE and PC species have C18 attached at the sn-1 position (Quehenberger et al., 2010), it is likely that the changes we observed on lysoPL could be due to differences in phospholipase activity, PLA_2 acting only at the sn-2 position. Since PE and PC species tend to be incorporated into PL membrane of VLDL, we can hypothesize that VLDL acts as a courier to deliver these essential FAs to peripheral tissues (Dashti et al., 2011; E.Vance, 2008).

PE 34:2, PE 36:2 and lyso PC 18:2 have been reported to be present in the VLDL fraction of plasma lipoproteins (Dashti et al., 2011). Indeed, these lipids could be exhibiting this linear increase due to meal intake. Our findings are similar to a previously published study where a postprandial increase of PE and PC species was observed in VLDL with highest values at two and four hours after meal intake (Sweeney et al., 1996). However, the reported findings were on p lipid species whereas our significant lipid species were all a lipid species. These differences could be due to the different statistical analysis methods used. The choice of considering lipid species that exhibited a time dependent trend across all three time-points was used in order to make our analysis as robust as possible. As changes in p lipid levels are rapid due to their high turnover rates (Magnusson and Haraldsson, 2011), it could have been missed in our study as we report differences across all three time-points instead of between any two time-points. Here, as a targeted approach was used, we report a subset of the FA composition present in these lipid species that were significantly changed across the three time-points which has not been reported previously.

4.2 Main study

4.2.1 Normalization by QC helps to reduce batch effect

As described earlier in section 3.4, QC normalization was applied to all the of the main study. In order to see the effectiveness of this normalization method, PCA plots were drawn to show the batch distribution (Figure 4-2). As shown in Figure 4-2, QC normalization was successfully applied for most lipids measured in both negative and positive ion modes, particularly for positive ion mode which had distinct batch clusters prior to QC normalization. After normalization, there was no distinct batch separation for batches measured in each ion mode, with overlapping PCA clusters for all measured batches. However, some lipid classes showed less distinct batch separation when compared to others, e.g. PE vs GluCer.

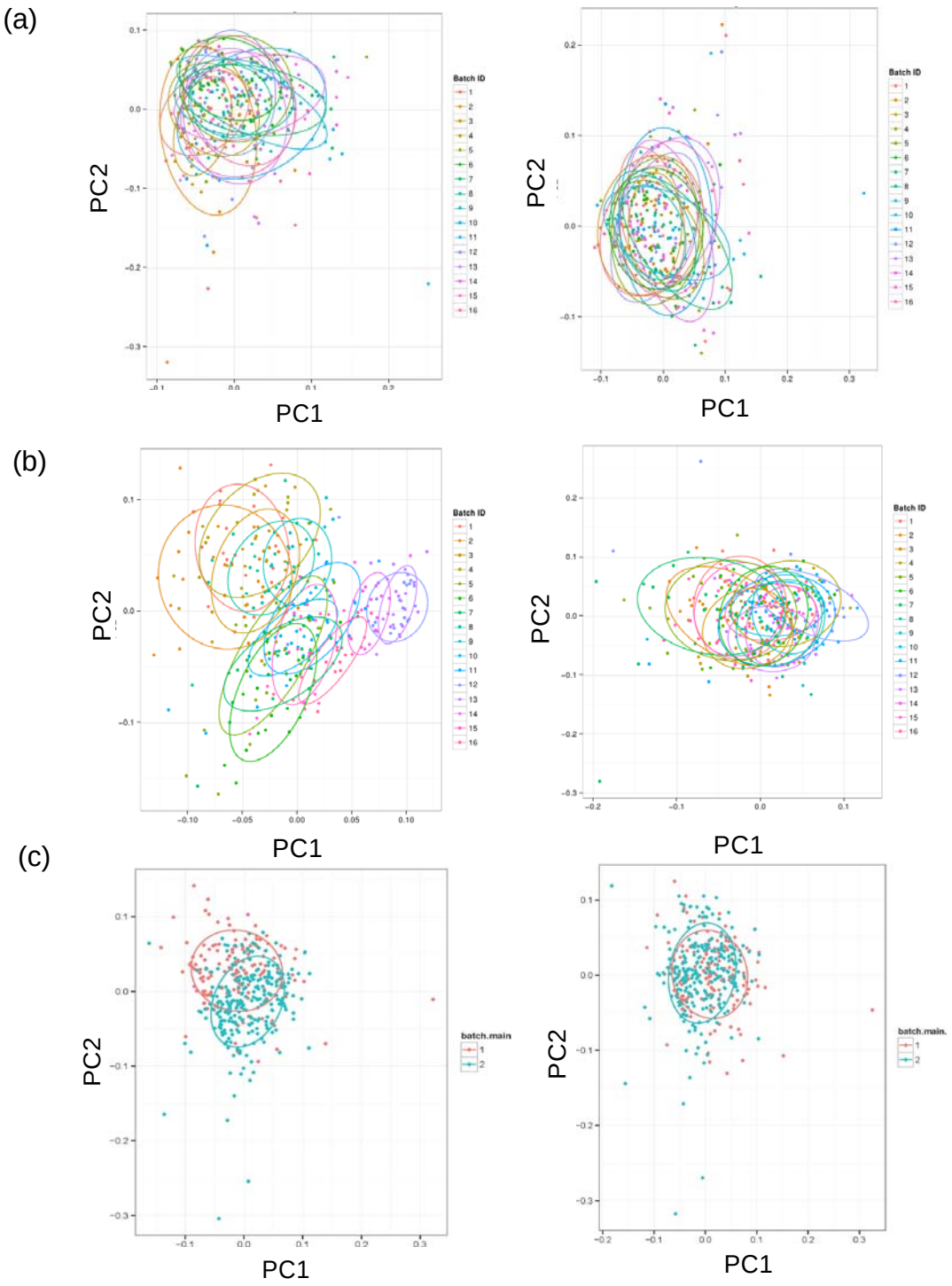
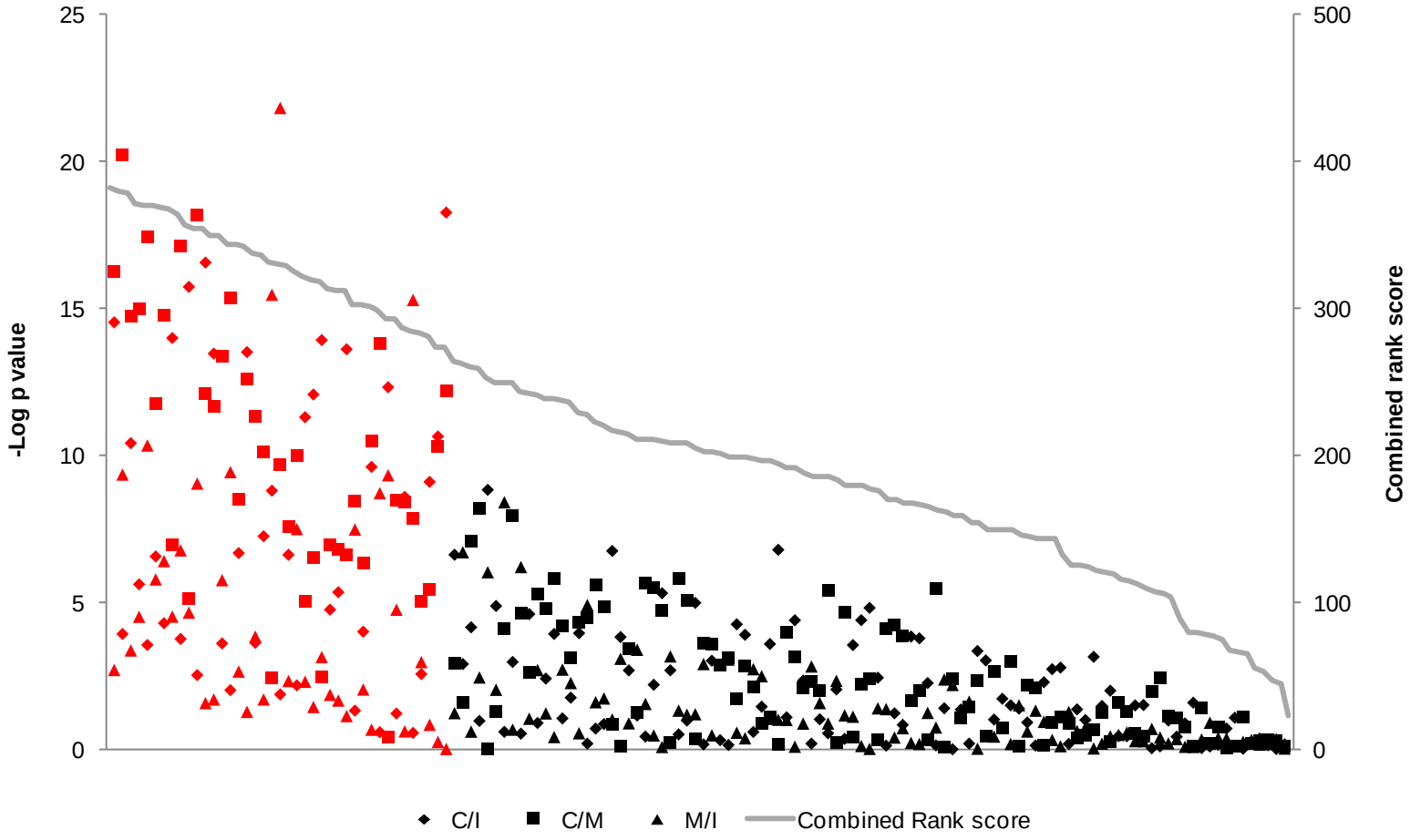


Figure 4-2: PCA plot showing batch effect before (left panel) and after (right panel) QC normalization for (a) lipids measured in negative ion mode (PE,PA,PI,PS,PG,GM3) (b) lipids measured in positive ion mode (PC,SM, GluCer) (c) STLs.

4.2.2 Targeted lipidomics analysis of various glycerophospholipid, sphingolipid and sterol species in the healthy Singaporean population

We used a shotgun targeted approach using MRM to look at 132 lipids across 9 different classes of GPLs (PE, PA, PI, PS, PG and PC) and SPs (GM3, SM, GluCer) as previously described. We also analyzed seven different STLs using LCMS approach. In order to obtain a global profile of all the lipids analyzed in this study, QC normalized lipids were tested for significant difference using pairwise comparisons between the different ethnic groups using student's t-test (two-tailed, $p<0.05$) as previously described. Combined rank score is shown as a trendline (Figure 4-3(a)). The distribution of all the lipid species measured in this study as well as the distribution of the 41 lipid species with highest rank scores above the median absolute deviation are likely to be more significant (Figure 4-3(b)).

(a)


25
20
15
10
5
0
-Log p value
500
400
300
200
100
0
Combined rank score
C/I
C/M
M/I
Combined Rank score

(b)

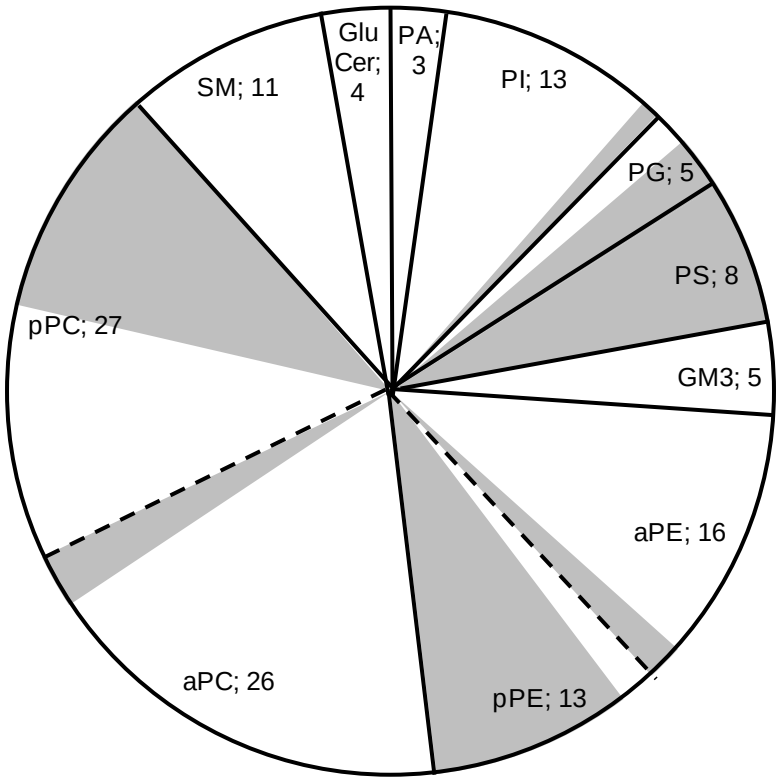
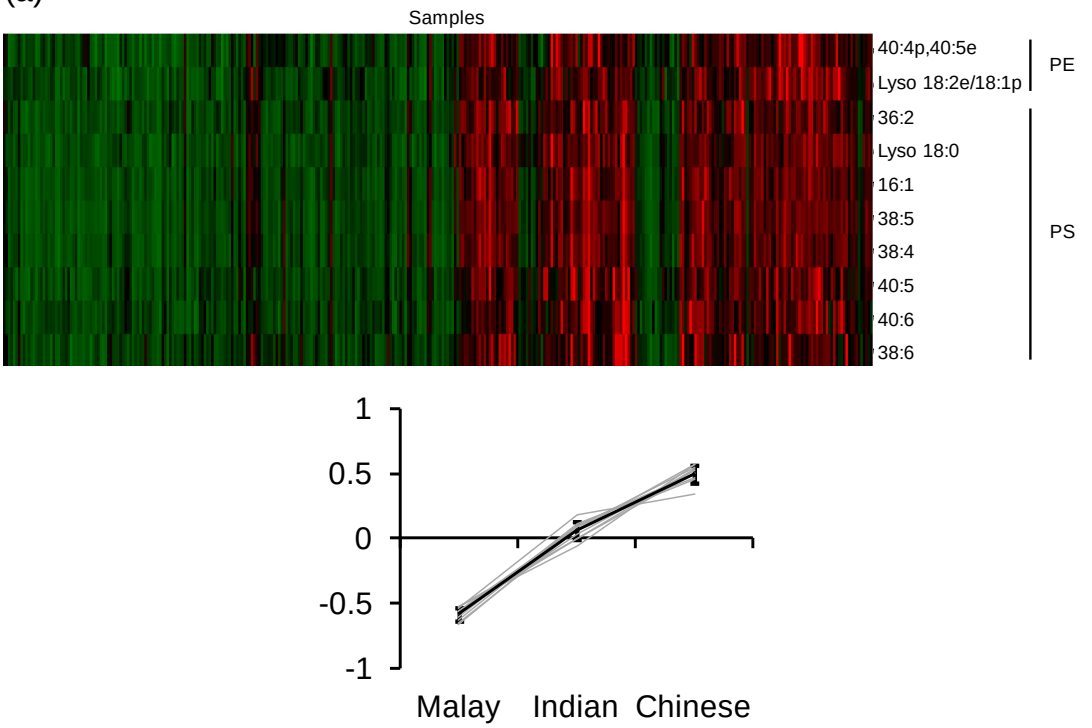


Figure 4-3: Global lipid profile across Malay, Indian and Chinese ethnicity in the Singaporean population for GPLs and SPs (a) Scatter plot showing the log p value (primary y-axis) and combined rank score (secondary y-axis) vs lipid species (x-axis). Lipid species with rank scores above the median absolute deviation are represented in red (b) Pie chart illustration of the distribution of all lipid species measured and the top 41 ranked lipids (highlighted in grey). Abbreviations a denotes ester-linked lipid species and p denotes plasmalogen lipid species.

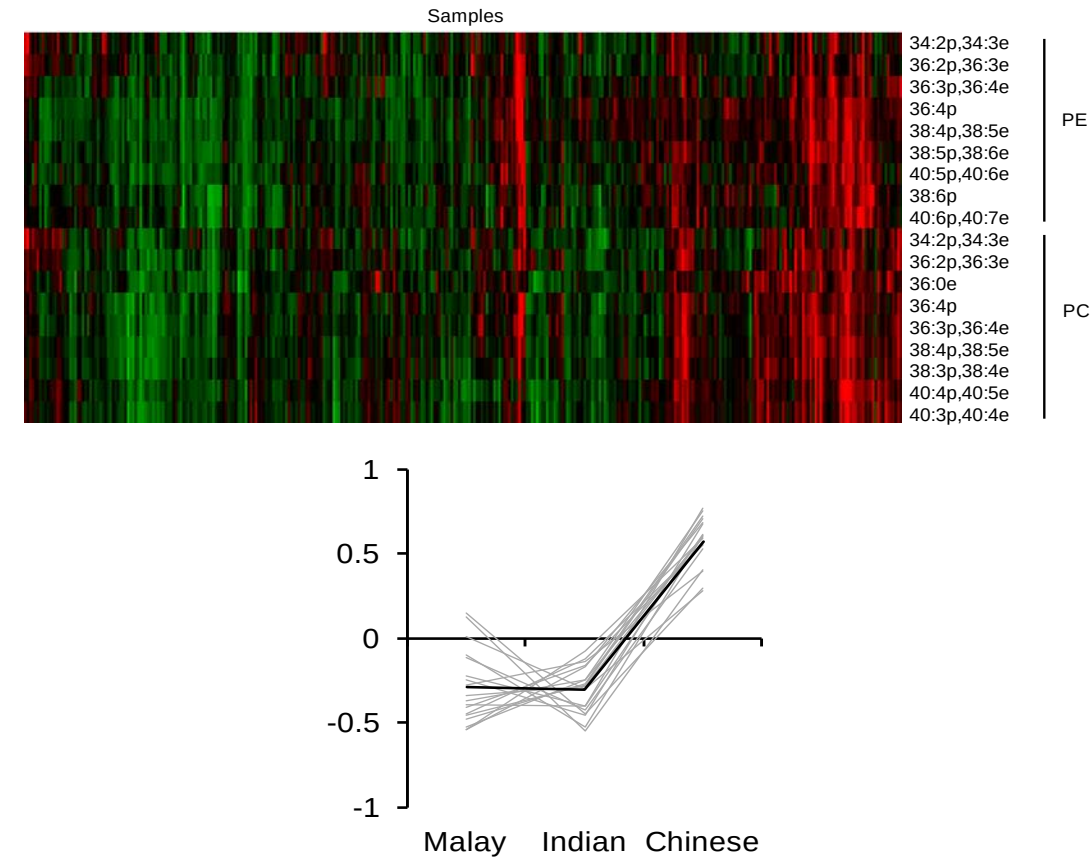
Most lipid classes did not show any ethnic associated significant differences. More than half of the significant lipids are represented by the plasmalogen GPLs, i.e. pPE and pPC species (Figure 4-3(b)). Very few ester GPLs, i.e aPE and aPC were significant. Notably, PS lipid species formed one-fifth of the significant lipids. In order to better see the significant differences across individual lipid classes, bar charts were drawn and shown in Supplementary Figure 6-1. Carbon chain length and degree of unsaturation was classified as total number of carbon atoms and double bonds across both FA chains respectively.

Gender dependent significant lipids were removed and the remaining lipids were analyzed as z-score values. Hierarchical clustering was done with these z-score values using Cluster 3.0 (Eisen et al., 1998) and visualized with Java TreeView (Saldanha, 2004), as previously described. Nine clusters were separated and analyzed to see the mean intensity across the three ethnicities in each cluster. The results of three main clusters are shown in Figure 4-4. The rest of the six clusters are shown in Supplementary Figure 6-2.

(a)


Samples
40:4p,40:5e
Lyso 18:2e/18:1p
PE
36:2
Lyso 18:0
16:1
38:5
38:4
40:5
40:6
38:6
PS
1
0.5
0
-0.5
-1
Malay
Indian
Chinese

(b)


Samples
34:2p,34:3e
36:2p,36:3e
36:3p,36:4e
36:4p
38:4p,38:5e
38:5p,38:6e
40:5p,40:6e
38:6p
40:6p,40:7e
PE
34:2p,34:3e
36:2p,36:3e
36:0e
36:4p
36:3p,36:4e
38:4p,38:5e
38:3p,38:4e
40:4p,40:5e
40:3p,40:4e
PC
1
0.5
0
-0.5
-1
Malay
Indian
Chinese

(c)

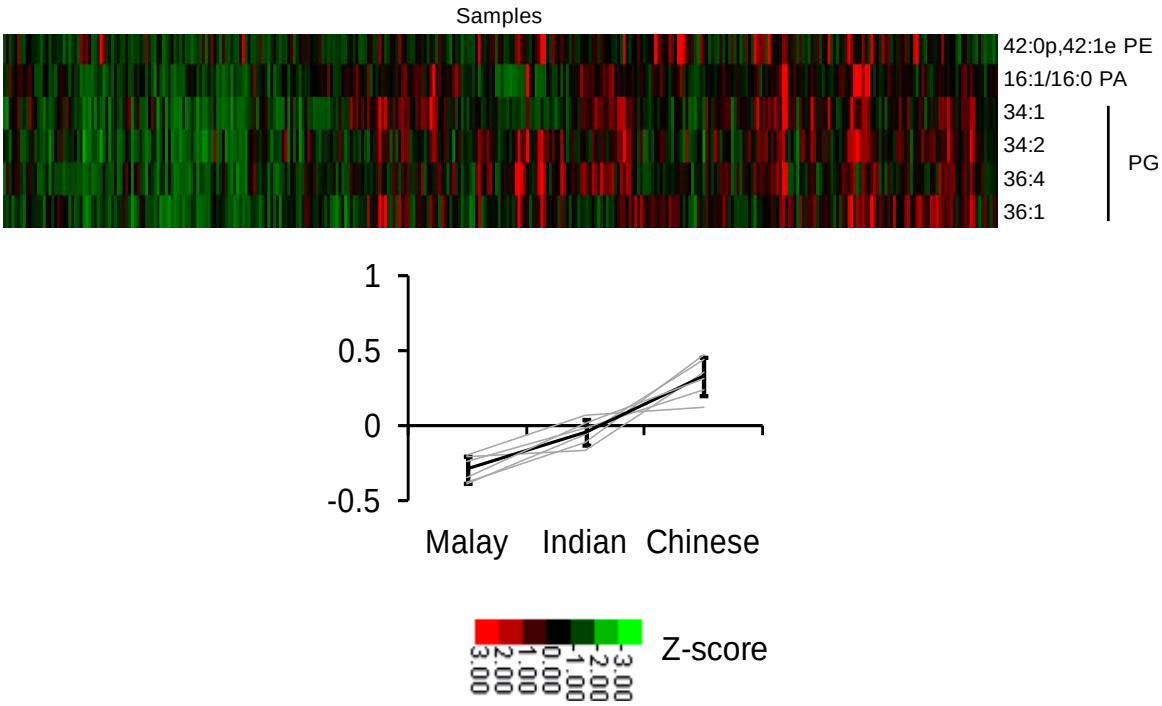


Figure 4-4: Hierarchical clustering of lipid species after removal of lipids significant due to gender, showing heat map (top panel) and cluster line plots (bottom panel) for (a) Cluster 1 (b) Cluster 2 (c) Cluster 3. Heat map colours represent relative values, with green representing values below the mean, red representing values above the mean and black representing the mean of a row (lipid) across all columns (samples) as shown in the colour scale below. Each lipid is represented by a grey line and the mean of each cluster line plot is represented by a bold line. Error bars represent standard deviation.

Most clusters comprised of lipid species from the same class, indicating that lipid species that were most biochemically similar tended to group together. While most of the clusters did not show any particular trend across the three ethnic groups, three clusters (Figure 4-4) showed a difference in abundance of lipids, particularly higher abundance in Chinese when compared to Indian and Malay populations. Specifically, we observed higher levels of PS and PG lipid species in Chinese population when compared to Malay and Indian populations. Not all PE and PC exhibit a

similar trend but only a few species, mainly pPE and pPC species. Lipid classes like PA and GluCer do not show any ethnic related differences (Supplementary Figure 6-2).

4.2.3 Ethnic related trends in carbon chain lengths

In order to see any biological relevant difference across the ethnicities in terms of carbon chain length, each lipid class was analyzed according to chain length composition. In this study, we defined carbon chain composition (C) as the total carbon content found in FA chains attached at sn-1 and sn-2 positions of GPLs. We classified lipids as Lyso (16, 18, 20, 22C), short carbon composition or SC (32,34C), moderate carbon composition or MC (36C) and long carbon composition or LC (38, 40, 42C). The classification is based on knowledge that most mammalian lipid species have FA with18C, thus anything above 36C (i.e. two 18C chains) are likely to be made up of a FA composition of 20C or more; hence LC (Quehenberger et al., 2010). Differences in carbon chain length for lipid species analyzed are reported separately depending on the gender, as previously described (t-test, two-tailed, $p<0.05$). Data is represented as density plot plotted as density (frequency) against normalized intensity for each ethnic group, as previously described. Data is shown for PE and PS (Figure 4-5). Relative intensity for each GPL class is shown in Supplementary Figure 6-3.

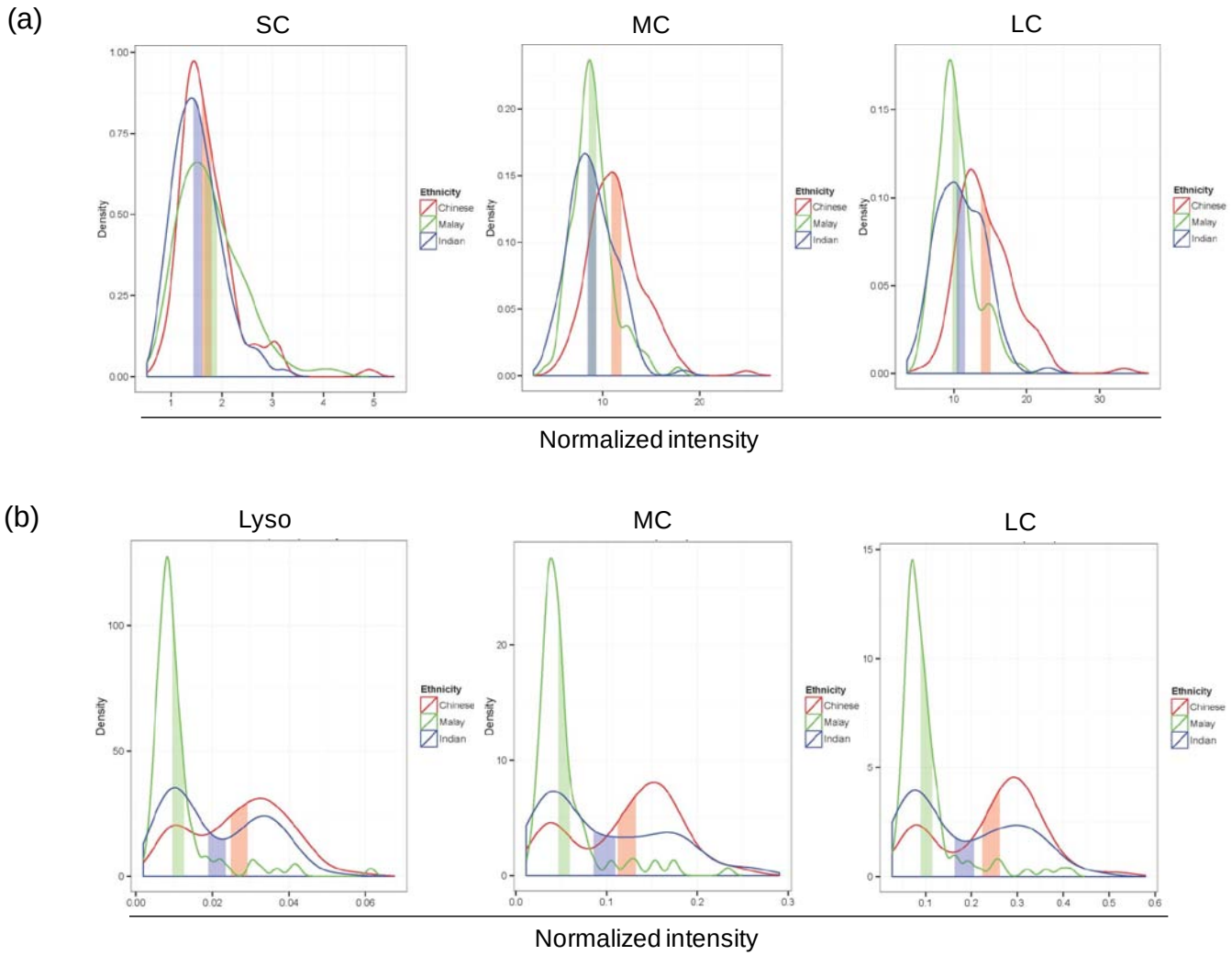


Figure 4-5: Ethnic related trends in carbon chain length represented as density plots, plotted as density (frequency in ethnicity) against normalized intensity for Malay, Indian and Chinese in the Singaporean population for GPLs (a) PE (b) PS. Shaded region indicates the 95% confidence interval based on the mean of each ethnic group.

Chinese population has significantly higher PS and PE when compared to Malay population ($p<0.001$) (Figure 4-5 and Supplementary Figure 6-3). It is interesting to note that the differences between the Malay and Chinese ethnicity are higher for MC and LC lipid species when looking at the 95% confidence intervals for most lipid classes. Furthermore, the distribution of lipid intensities is narrower for Malay population when compared to Indian and Chinese populations, especially for PS where Chinese and Indian populations exhibit two distinct peaks instead of one, indicating a wider range of variation. This two peak distribution is observed across all carbon chain lengths only for PS and not for PE species. Though LC lipid species show a high SEM which is due to difference in relative intensities among the 38, 40 and 42 chain lengths, the ethnic associated differences are still significant when analyzed based on their individual chain length (data not shown). For other GPLs, Chinese have more abundant MC and LC PC lipid species, similar to PE though the inter-ethnic differences are less pronounced. Indian and Malay populations had lower PA and PG species irrespective of carbon chain length; however we could not observe if this trend is there for LC lipids due to the absence of LC lipid species measured in this study. In the case of PI, no significant difference was observed between the three ethnic groups for LC lipid species.

SPs are characterized by their “sphingoid” base backbones, which are long-chain alkanes or alkenes (sometimes with additional unsaturation) of approximately 14–20 carbons in length, with an amino group at position 2 and hydroxyl-substituents at positions 1 and 3 (Merrill et al., 2005). In humans, the

alkyl chain length of the sphingoid base is usually 18 carbons in length, with the exception of brain gangliosides, where substantial amounts of 20-carbon species appear in an age-dependent manner (Merrill et al., 2005). In this study, we exclusively measured SPs containing 18 carbons in their sphingoid base. We further investigated the attached FA composition of the ceramide backbone. SPs were classified as SC (16, 18 C), MC (20, 22 C) and LC (24 C) species, based on their chain length in mammalian plasma (Quehenberger et al., 2010). Data for total SPs is shown in Figure 4-6. Remaining classes are shown in Supplementary Figure 6-4.

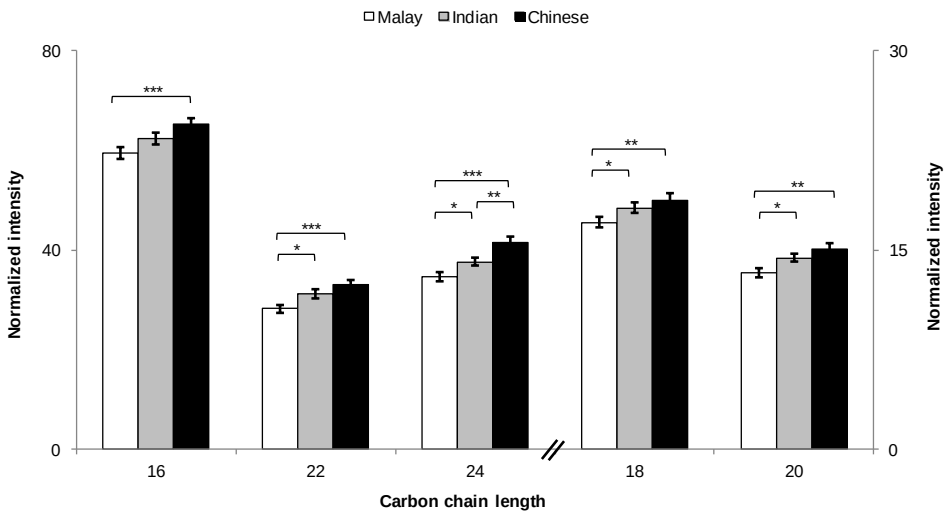


Figure 4-6: Ethnic related trends in carbon chain length (attached FA) across Malay, Indian and Chinese in the Singaporean population for total SPs in terms of their individual chain length (bottom panel). Break in axis indicates normalized lipid intensity on right secondary y-axis. Error bars are plotted as standard error mean (SEM).* is p <0.05, ** is p<0.01 and *** is p<0.001.

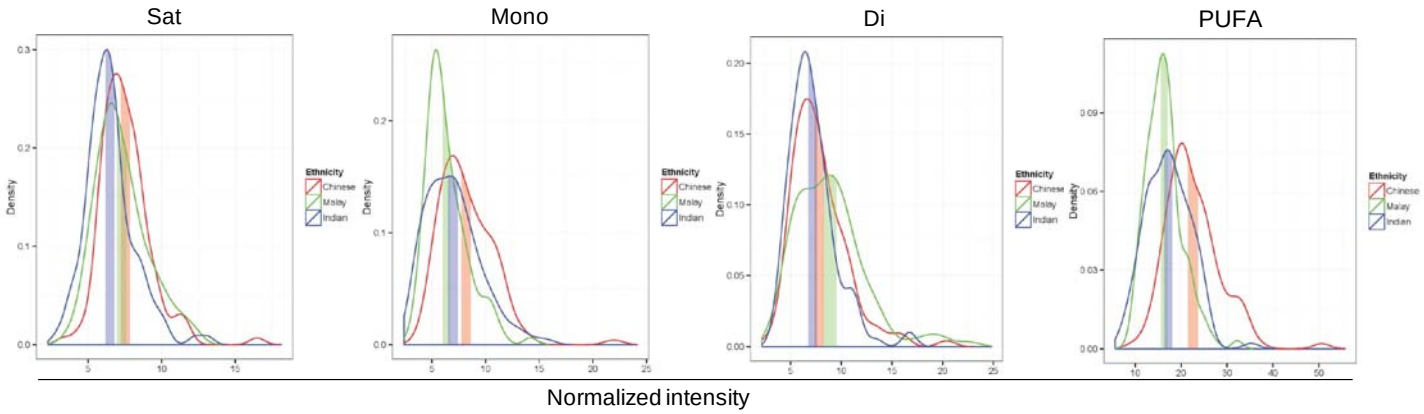
When looking at carbon chain length across all SPs, interestingly, Chinese have a higher relative intensity when compared to Malay and Indian,

populations, especially for MC and LC containing SPs. However, when looking at individual SP classes, the trend of Chinese showing a higher abundance of SPs is more prominent for SM and GM3 than for GluCer. This is consistent with our findings for MC and LC GPLs species (PE, PS and PG) where only selected classes show a more distinctive increase in relative intensity for Chinese when compared to Indian and Malay populations.

4.2.4 Ethnic related trends across degree of unsaturation

Ethnic associated changes were also analyzed across different degrees of unsaturation for all GPLs and SPs. Due to the diversity of lipids measured in this study, various degrees of unsaturation were present and lipids were analysed based on the number of double bonds present across both FA chains. Comparative analyses were done on lipid species classified as saturated (Sat, 0 double bond), monounsaturated (Mono, one double bond), diunsaturated (Di, two double bonds) and polyunsaturated (PUFA, three or more double bonds) lipid species. Student's t-test (two-tailed, $p<0.05$) was done, and data is represented as density plot plotted as density (frequency) against normalized intensity for PE and PS, as previously described (Figure 4-7). Relative intensity for each GPL class is shown in Supplementary Figure 6-5.

(a)



(b)

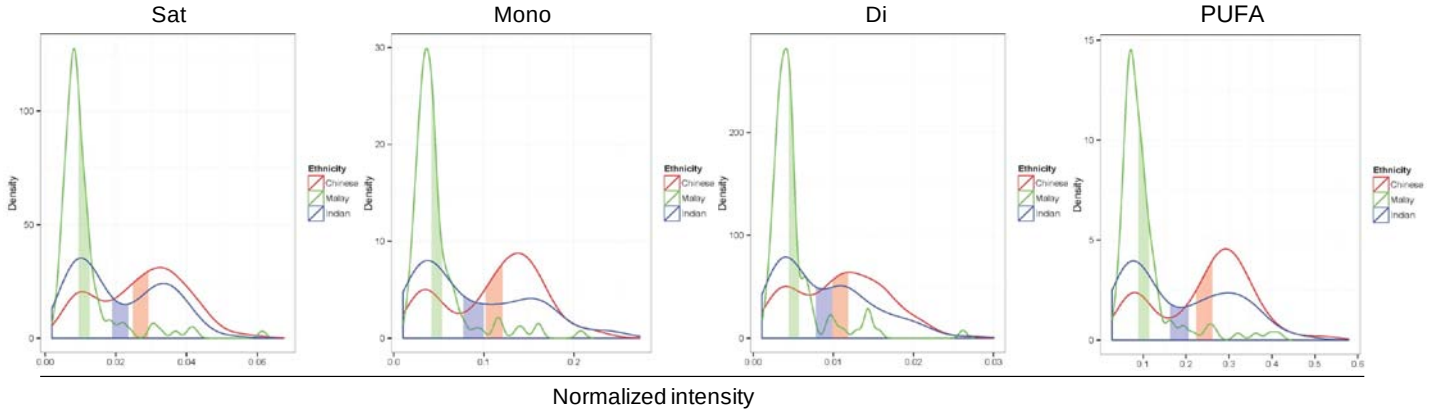


Figure 4-7: Ethnic related trends for degree of unsaturation represented as density plots, plotted as density (frequency in ethnicity) against normalized intensity for Malay, Indian and Chinese in the Singaporean population for GPLs (a) PE (b) PS. Shaded region indicates the 95% confidence interval based on the mean of each ethnic group.

Chinese have a significantly higher intensity when compared to the Malay population, particularly for PUFA PE lipid species when looking at the 95% confidence interval ($p<0.001$) (Figure 4-7). The SEM is high for PUFA lipid species in some lipid classes as the relative intensity between the various numbers of double bonds may be different, and are further represented based on their individual number of double bonds to see if any ethnic related differences exist (Supplementary Figure 6-5). The biggest difference between Malay and Chinese populations is seen for the five double bonds, whereas for PS, the trend is seen across all lipid species irrespective of the degree of unsaturation (Supplementary Figure 6-5). Furthermore, Malay population appears to have a narrower range of variation in lipid intensities when compared to Indian and Chinese populations, particularly for PS and this is also seen across different carbon chain lengths for PS (Figure 4-7). When looking at other GPLs, Chinese have more abundant PUFA PI lipids, particularly for 4 double bonds whereas all PUFA PC lipid species with three or more double bonds show a significantly higher intensity in Chinese as compared to Indian and Malay populations (Supplementary Figure 6-5). In the case of PG lipid species, Indian and Malay populations have less abundant PG species irrespective of the degree of unsaturation when compared to Chinese population (Supplementary Figure 6-5). This observation cannot be confirmed in PA due to the absence of PUFA PA lipid species measured in our list of lipids.

As previously described, the analysis for the degree of unsaturation was done separately for sphingoid backbone and the attached FA chain.

However, some classes like the GM3 and GluCer only have differences in the number of double bonds in either the sphingoid backbone or attached FA, thus the data is reported accordingly as shown in Supplementary Figure 6-6.

For SPs measured in this study, analysis was targeted to lipid species with either zero or one double bond in either the sphingoid backbone or the attached FA. It can be seen that for sphingoid backbone double bonds, gender differences exist like for GM3 where the trend of Chinese being higher than Malay and Indian populations in relative intensity is observed only for the females and not for males. Furthermore this is seen more clearly for saturated species in both the sphingoid backbone and attached FA chain, particularly for SM as SM constitutes the dominant lipid class measured among all SPs (Supplementary Figure 6-6).

4.2.5 Ethnic related trends across ester-linked and plasmalogen PE and PC lipid classes

In our study, PE and PC contain a mixture of a and p lipid species. pPE and pPC measured in this study consist of both ether-linked lipids and plasmalogens which contain a vinyl ether moiety at the sn-1 position of the glycerol backbone (Nagan et al., 1998), as shown in Figure 4-8.

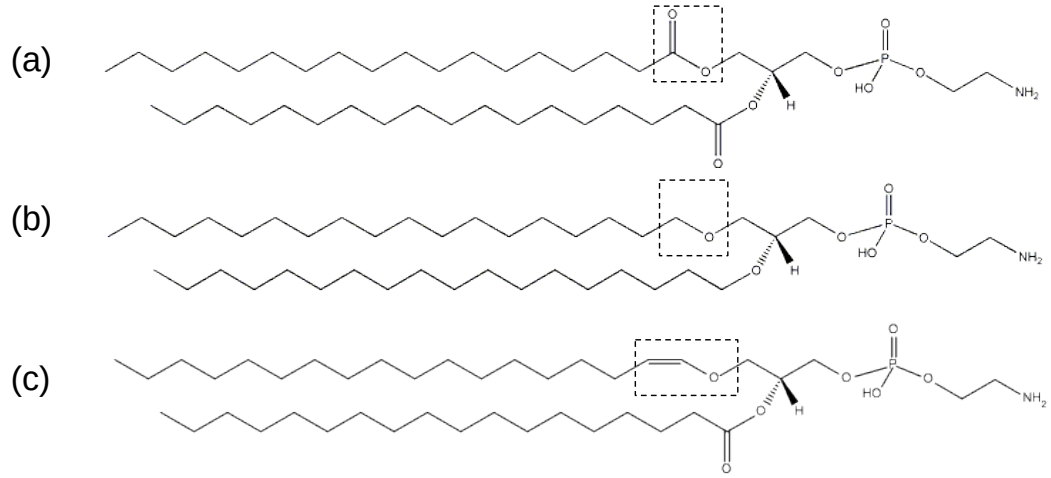


Figure 4-8: Typical structure of an (a) ester PE 18:0/18:0 (b) ether PE 18:0/18:0 (c) plasmalogen PE 18:0/18:0 with a vinyl ether bond. The box indicates the difference in FA linkage. *Adapted from Fahy et al.(Fahy et al., 2009).*

In order to see if any ethnic related differences exist between a and p PE and PC species, a class comparison was done to see if there were any differences that showed up across the 3 ethnicities. Student's t-test (two-tailed, $p<0.05$) was done and gender-dependent differences were analyzed and reported separately (t-test, two-tailed, $p<0.05$). Data is represented as density plot plotted as density (frequency) against normalized intensity for PE

and PC (Figure 4-9) and relative intensity for each GPL class, as previously described (Supplementary Figure 6-7).

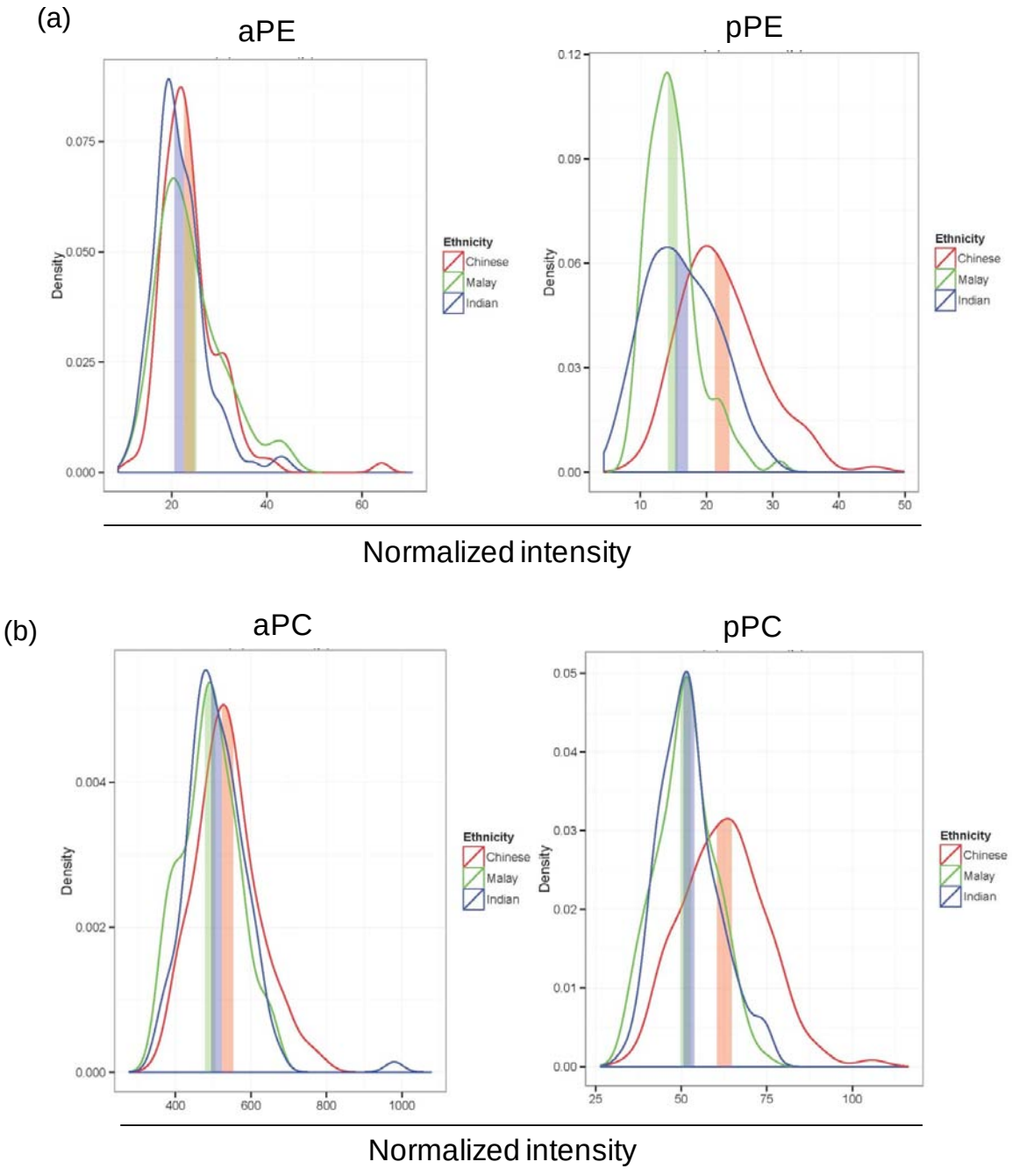


Figure 4-9: Ethnic related trends for a and p lipid species represented as density plots, plotted as density (frequency in ethnicity) against normalized intensity for Malay, Indian and Chinese in the Singaporean population for GPLs (a) PE (b) PC. Shaded region indicates the 95% confidence interval. *Abbreviations*: a denotes ester and p denotes plasmalogen species.

Chinese have a significantly higher intensity when compared to the Indian and Malay populations when looking at the 95% confidence interval, particularly for pPE and pPC lipid species ($p<0.001$) (Figure 4-9). When looking at the population distribution of plasmalogen lipids, Chinese and Indian populations have a broader range of variation in lipid intensities when compared to Malay population, particularly for pPE. However, population distribution for a lipid species were similar across all three ethnic groups. Ethnic related differences between a and p lipid species were more distinct in PE species when compared to PC species (Supplementary Figure 6-7).

.

4.2.6 Statistical validation of ethnic related trends in carbon chain length, degree of unsaturation and plasmalogen/ester species using regression and permutation analyses

To confirm the descriptive findings, linear regression was performed to test any association of age and gender using type II ANOVA (two-tailed, $p<0.05$) as previously described. Four hypotheses were tested:

1. Age, after accounting for gender and ethnic groups, ignoring gender:ethnicity interaction.
2. Gender, after accounting for age and ethnicity, ignoring gender:ethnicity interaction.
3. Ethnicity, after accounting for age and gender, ignoring gender:ethnicity interaction.
4. Gender-ethnicity interaction, after accounting for age, gender, ethnicity.

As most of the p-values obtained are not significant, data for ethnicity (hypothesis 3) only is reported (Supplementary Table 6-2). This is consistent with our findings using student's t-test (two-tailed, $p<0.05$).

In order to validate the observed significant trends across long chain PUFA containing lipid species (PS and PE) as well as p lipid species (PE and PC), permutation test was performed. The main aim of this test was

to ensure differences across the Chinese when compared to Malay and Indian were indeed significant or confounded by other factors such as gender and age. Permutation analysis was performed using R (Team, 2008) as previously described. The following considerations were taken into account while performing the permutation tests:

1. Age, gender and ethnicity was permuted as a whole and the the ratio is maintained to be the same as the original sample set.
2. Permutation was performed 1000 times. The number of permutations analyses was limited as 1000 runs took~1000 seconds to finish running, thus larger number of analyses was not performed.
3. Even if no p-value from permutation was smaller than the true one (obtained from Type II ANOVA), p-values were adjusted to 1, giving a minimum value of 0.001.

A lower permutation p-value would indicate a higher likelihood that the observed significant difference was true and less likely due to random chance. An arbitrary cutoff above 0.005 (i.e. 5 out of 1000 permutations) was used as significant due to random chance. The data for PS, PE, PG and a and p PE/PC lipid species are reported in Supplementary Table 6-2.

It can be seen that most of the lipid classes with ethnic associated significant differences observed from earlier analyses using student's t-test was confirmed by linear regression, and was validated by permutation analysis with good p-value ($p<0.005$), with the exception of Lyso PE. PS showed a statistically significant increase in intensity in Chinese when compared to Malay and Indian populations across all carbon chain lengths and degree of unsaturation. However, type II ANOVA showed an increase in significant difference between ethnicities when progressing from SC to LC PE lipid species. A similar observation was also made for PE at higher degree of unsaturation, i.e. PUFA containing PE lipid species when compared to Sat ones. Lastly, a significant ethnic related difference was observed in pPE and pPC species when compared to their aPE and aPC species , consistent with previous analysis.

4.2.7 Ethnic-related trends across sterols and lipoproteins

Seven different STs were measured in this study using LC-MS (Supplemetary Figure 6-1). Clinical measurements of Chol, HDL and LDL were obtained as previously described (Lee Kian Ng, National University Hospital). The results of these clinical measurements are shown in Figure 4-10.

Clinical Measures	Malay	Indian	Chinese
Cholesterol	5.89	5.54	5.66
HDL	1.34***	1.18***	1.48
LDL	3.93***	3.90***	3.50
HDL/LDL ratio	0.36***	0.32***	0.45

Figure 4-10: Clinical measurements of Chol, HDL and LDL obtained by colorimetry method. *** is $p<0.001$ when compared to Chinese population. *From:* Lee Kian Ng (National University Hospital).

No significant inter-ethnic differences were observed for all STs cpecies measured. These findings, particularly for Chol was similar to clinically obtained measurements of Chol, indicating that all three ethnic groups had similar measurements of Chol. However, clinical measurements of HDL and LDL showed significant inter-ethnic differences. Indian population exhibited the lowest plasma HDL concentrations when compared to Malay and Chinese populations, while Malay population had the highest plasma LDL

concentrations. When looking at HDL/LDL ratio, Chinese had the highest ratio when compared to the other two ethnic groups. All these differences are statistically significant when compared to the Chinese population.

Previous studies have been done on the correlation of lipoprotein fractions and plasmalogen species. In our study, due to the significantly higher abundance of p lipid species and HDL levels in Chinese population when compared to the other two ethnic groups, correlation analysis were done between p lipid species (pPE and pPC) and HDL/ LDL levels. The results of these correlation analyses are shown in Figure 4-11 and Supplementary Figure 6-9.

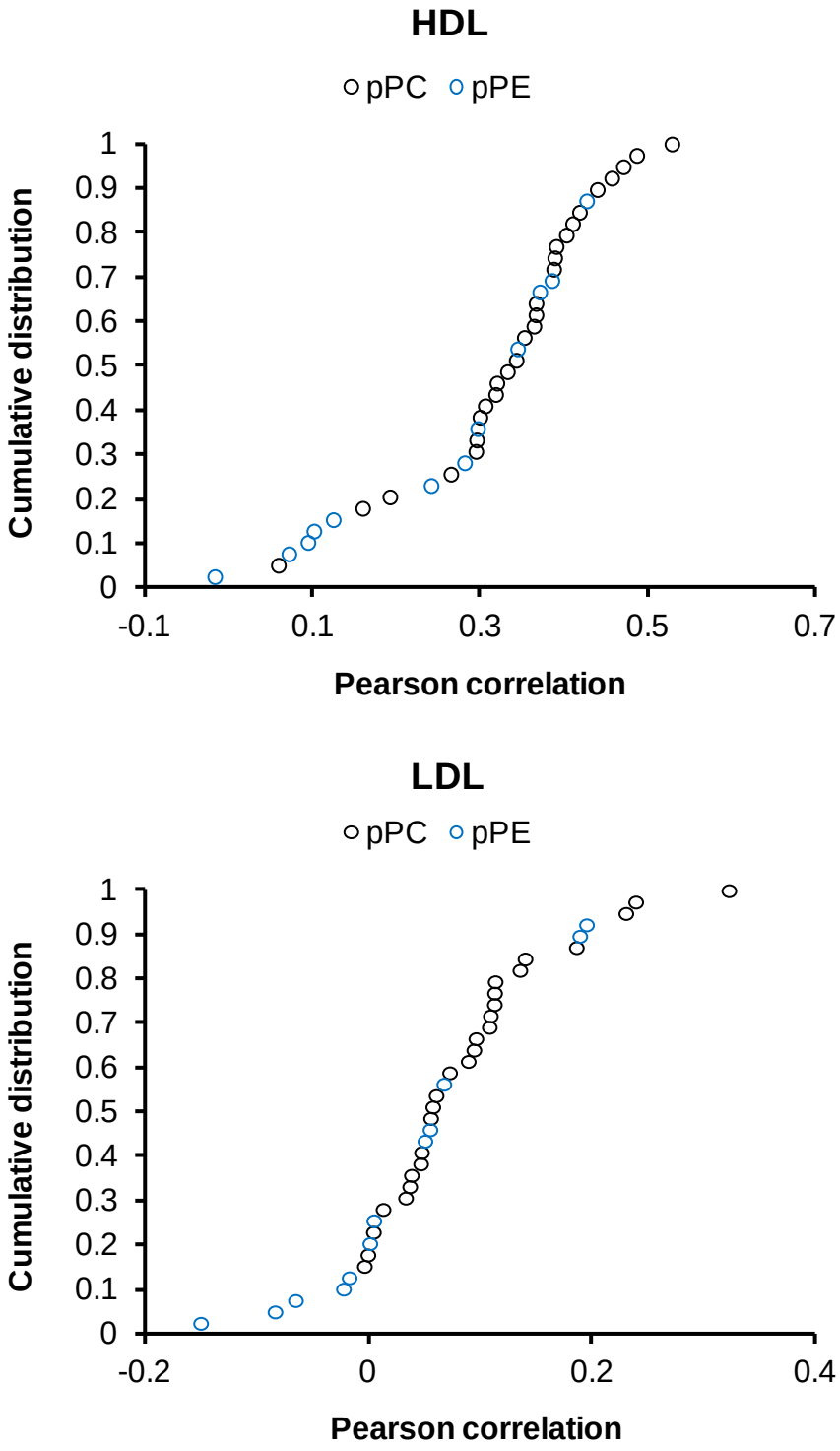


Figure 4-11: Pearson correlation between p lipid species (pPE and pPC) and clinical measurements of HDL (top panel) and LDL (bottom panel).

Most pPE and pPC species are positively correlated with HDL and LDL. When looking at HDL/LDL correlation with p lipid species within each individual ethnic group, Chinese population have a higher positive correlation with HDL compared to Indian and Malay populations (Supplementary Figure 6-9). However, all three ethnic groups exhibit similar correlation between p lipid species and LDL levels (Supplementary Figure 6-9).

4.2.8 Discussion

4.2.8.1 Shotgun lipidomics

Shotgun lipidomics based on intrasource separation only represents one approach for global lipidome analysis with over 90% of lipid mass been accounted for and quantitated by this approach directly from the lipid extracts of biological samples (Han and Gross, 2005), Today, shotgun lipidomics provides a strong foundation for the analysis of thousands of molecular species after data are judiciously analyzed by bioinformatic processes (Han and Gross, 2005).

Shotgun lipidomics has been used extensively for the analysis of lipids, where distinct fragmentation patterns of different lipid species allow for reliable detection of these lipids (mostly abundant lipids) without prior chromatographic separation, with the use of appropriate internal standards for quantification (Han, 2010). The fragmentation patterns of each lipid molecular species depend on both the applied energy for collision-induced dissociation and the structure of individual molecular species (Han and Gross, 2005; Sullards, 2000). Changes in applied collision energy alter the kinetics of individual fragmentation pathways resulting in changes in the distribution of the observed fragment ions (Han and Gross, 2005). In our study, optimization of MS parameters was done for each lipid class based on an existing method to ensure reliable detection of different lipid species in the human plasma biological matrix. Ratiometric comparisons to internal standards are useful

because the relative fragmentation rates of identical molecular species reflect the intrinsic chemical properties of the species in the gas phase (Han and Gross, 2005)

It is important to note that thousands of very low abundance lipid molecular species are not approachable by shotgun analysis and pre-processing methods are required to enrich low abundance lipid molecular species, prior to integration with shotgun analysis (Han and Gross, 2005). In this study, no pre-processing was done to enrich low abundant GPLs and SPs, as such species like Cer and GM3 were excluded from biological variation analysis.

Systematic studies of instrument responses using structurally distinct lipid species have demonstrated that ionization efficiency is predominantly dependent on the lipid polar head group and only weakly dependent on the structure of fatty acid moieties (Han and Gross, 1994, 2005). Previous studies have shown that quantitative shotgun lipidomics approach demonstrated no significant differences in the instrument response of molecular lipids differing in acyl chain lengths and number of double bonds supporting the concept that a single lipid class selective non-endogenous standard would be sufficient for the absolute quantification of the endogenous molecular lipid species (Jung et al., 2011). However, it is noteworthy that precise quantification of molecular species by shotgun lipidomics can be potentially hindered by interferences arising from isotopes, isobars and

isomers (Jung et al., 2011), in this study no isomers were analyzed as targeted MRM approach was used for analysis.

While previous work have demonstrated that shotgun lipidomics yields comparable and even better data than LC-MS, especially in the case of TAG analysis due to more stable background/chemical-noise throughout the analysis and shorter analysis time, caution has to be employed for proper optimization of various MS parameters to ensure reliable detection of lipid species (Jung et al., 2011). Furthermore, shotgun lipidomics has been successfully applied as a measurement of choice across various conditions like hypertension and Alzheimer's disease (Graessler et al., 2009; Han, 2010). As such, this study employs ESI-MS shotgun approach whereby shotgun lipidomics still represents an attractive and promising avenue for fast and reliable analysis of lipids.

4.2.8.2 QC normalization helps to remove batch effects

It can be seen that QC batch normalization works more effectively for some lipid classes than the others (Figure 4-2). This could be due to inherent properties of lipid species like ionization (Quehenberger et al., 2010). Furthermore, normalization works better for lipids detected in the negative ion mode when compared to the positive ion mode. This could be explained by the large variable range of intensity for some lipid classes in the positive ion mode (Liebisch et al., 2004; Quehenberger et al., 2010) such that the

normalization works best for only a subset of lipid species in this range, especially in the case of PC, where both very high and low intensity lipids are detected.

4.2.8.3 Cholesterol as a good biomarker of measurement during fasting state

Targeted lipidomics analysis of GPLs, SPs and STLs showed that ethnic associated significant lipids were mainly detected from GPLs and some SPs (Figure 4-3). However, no major significant lipids were detected for STLs. Abundant lipids like Chol and CholE were found to be similar across all three ethnicities (Supplementary Figure 6-1). This indicates that Chol and CholE are good biomarkers to be measured as clinical parameters across fasting individuals as they do not exhibit any significant change across gender, age and ethnicity. This finding is in line with a recent study by Shui et al. which reports that there is no significant differences in Chol levels measured across healthy monkeys in fed and fasting state (Shui et al., 2011b). Another recent study also reported that there exists no significant biological variations of Chol when looking at diurnal changes and menstrual cycle variation in healthy individuals (Burkard et al., 2007). This reinforces the common current practice of using Chol as a diagnostic form of measurement in clinics. Furthermore, in our study, no significant change across ethnicity was detected for 24-OHChol and is consistent with previous literature where biological variations based on absolute quantifications of 24-OHChol in plasma were observed to be similar

to intra-day technical variation of the analytical method used (Burkard et al., 2007).

4.2.8.4 Ethnic associated biological variation in plasma

Lipids found in plasma are mainly made up of free FA and lipids found in lipoproteins (Dashti et al., 2011). The main GPL and SP constituents in plasma are PC (57–81%) followed by SM (12–26%). Lyso-PC and other GPLs (PE, PS, and Pl) constitute 5–15% with respect to GPLs and SPs (Quehenberger et al., 2010; Skipski, 1972). The lipid levels do not change much across the different classes of lipoproteins during fasting state (E.Vance, 2008). Thus the GPLs and SPs changes that we observe across ethnicity are likely to be metabolically important since all samples were taken under fasting state (Figure 4-3).

The significant lipid species associated with ethnicity are mainly PS, pPE and pPC lipids. PS and aPE/PC are known to be metabolically linked whereas pPE and pPC are linked by headgroup exchange (Vance and Tasseva, 2013). pPE and pPC are biosynthetically different from their a species counterparts as they are synthesized from long chain alcohols. Serine is required for the synthesis of PS as it is the headgroup for PS species (Kanehisa and Goto, 2000). It is well known that PS, PE and PC are metabolically linked due to the similarity of their headgroups, whereby decarboxylation of PS can lead to PE, and headgroup exchange of

ethanolamine to choline leads to conversion of PE to PC (Kanehisa and Goto, 2000). This has been reported experimentally whereby serine can be a precursor of headgroups of both PS and PE, including pPE and aPE in cultured glioma cells due to decarboxylation and headgroup exchange (Xu et al., 1991). Interestingly, previous reports have shown that the incorporation of serine into both pPE and aPE pools and this incorporation is rapid and sustained when compared to using ethanolamine, indicating that serine may be utilized as a preferential pathway for synthesis of pPE species (Cook et al., 1991; Xu et al., 1994; Xu et al., 1991). This could be due to incorporation of serine into the SP pathway which then results in generation of alcohols such as hexadecanal which is subsequently utilized for pPE synthesis pPE can then be converted to pPC by headgroup exchange, similar to their a species counterparts (Nakahara et al., 2012). Furthermore, it has been reported that pPE species are the most abundant mammalian PE and are contained in cellular membranes of many mammalian cells, including neutrophils, erythrocytes, and those of nervous system, heart, vascular smooth muscle and tumors (Nagan and Zoeller, 2001; Synder, 1972). The wide distribution and higher abundance of p lipid species in certain cells and tissues may indicate the requirement for higher precursor pools needed to maintain high levels of p lipid species in these cells and tissues. As such, serine may be utilized as a preferential source of p lipid synthesis, as reported by Xu (1991, 1994). This metabolic relationship could explain why higher levels of both PS and pPE/pPC are observed in Chinese population when compared to the other two ethnicities.

PS is a minor membrane GPL that is produced by eukaryotic cells and is present in moderately low levels in the human plasma (Quehenberger et al., 2010; Vance and Steenbergen, 2005). The exposure of PS on the outer surface of cells is widely believed to play a key role in the removal of apoptotic cells and in initiation of the blood clotting cascade (Vance and Steenbergen, 2005). Recent studies have shown that PS exposure on red blood cells is elevated in patients with metabolic syndrome (Straface et al., 2011). The reason for elevated PS levels in Chinese population across all degrees of unsaturation and carbon chain length is unknown. We hypothesize that higher levels of PS are due to higher levels of serine which are required to maintain higher levels of p lipid biosynthesis in Chinese due to precursor-product relationship between serine and pPE species (Xu et al., 1991). However, high levels of p lipid species in plasma could also be due to dietary influences, if this is the case, then this relationship may not hold true.

In our study, PS, PE and PC species contributing to significant inter-ethnic differences were detected in the plasma with COV less than 20% (Figures 3-4 to 3-6). Previous studies have reported the reliable detection of these lipids in human plasma (Quehenberger et al., 2010; Sato et al., 2010) and the requirement for analytical validation to be within 25%, which is well within the range of COVs measured for PS and pPE and pPC species found to be significantly different across ethnicities in this study (Chau et al., 2008) However, a recent study reports the absence of PS in lipoprotein fraction, while PE were mainly detected in the VLDL fraction (Dashti et al., 2011). This

could be due to the low amounts of PS in plasma under normal conditions which is below the detection limit in their study (Dashti et al., 2011).

4.2.8.4.1 Long chain PUFA containing lipids show inter-ethnic differences

Longer carbon chain lipids with a higher degree of unsaturation were observed to be significantly higher in the Chinese population when compared to Malay and Indian populations, particularly for PE and PS species (Figures 4-5 to 4-7). This effect was more distinct in PUFA containing lipids having 36 carbon atoms and above. LC PUFA lipids are synthesized in the human body by a series of steps involving desaturases like fatty acid desaturase (FADS) and stearoyl coA desaturase (Scd), and elongases (Elovl). The activities of these enzymes have been shown to be affected by genetic variation such as SNPs (Glaser et al., 2010; Lattka et al., 2010; Merino et al., 2010; Schaeffer et al., 2006). Furthermore, SNPs in desaturases have been associated with homeostatic regulation of blood levels of LC PUFA lipids, indicating a direct relationship between genetic variation and variations in lipid levels (Gieger et al., 2008; Schaeffer et al., 2006; Tanaka et al., 2009).

In our study, we observed Chinese having more abundant LC PUFA lipid species when compared to Indian and Malay populations. Genetic differences in the FADS genes (FADS1 and FADS2) could contribute to the

differences in the level of LC-PUFAs observed across the three ethnic groups. Previous reports have shown that FADS cluster represents one of the most variable genes and present a high degree of genetic variation (Gieger et al., 2008; Tanaka et al., 2009). Differences in allele frequencies of different SNPs between the three ethnic groups that affect the activities of FADS genes could explain differences in activities of these enzymes, leading to different homeostatic regulation of blood LC PUFAs.

Some of these LC PUFAs are then incorporated into diacyl lipid species, such as those measured in this study. As such, an increase in lipid species such as PE and PS containing these LC PUFAs could be a reflection of a general increase in plasma LC PUFAs due to an increase in FADS activity, in the case of the Chinese when compared to the Indian and Malay populations, Furthermore, plasma PUFA has been reported to be influenced by diet (Tanaka et al., 2009), Measurement of dietary PUFA intake could be a useful indicator of identifying whether lifestyle factors such as diet could further contribute to inter-ethnic differences observed in this study, in addition to genetic variation of FADS genes. (Toufektsian et al., 2011).

4.2.8.4.2 Plasmalogen lipids are more abundant in the Chinese population

In our study, pPE and pPC are found to be significantly higher in Chinese (Figure 4-10). Plasmalogens (1-O-alk-1'-enyl-2-acyl

glycerophospholipids) belong to a subclass of ether PLs and are characterized by the presence of a *cis* double bond on the alkyl chain adjacent to the ether linkage at the sn-1 position giving rise to the characteristic vinyl-ether bond (Brites et al., 2004; Nagan and Zoeller, 2001). Plasmalogen synthesis is initiated in the peroxisomes, and subsequently completed in the endoplasmic reticulum (Brites et al., 2004). Plasmalogens have a very well characterized and defined structure in which the sn-1 position is usually occupied by a 16:0 (palmitic acid), 18:0 (stearic acid) or 18:1 (oleic acid) carbon chain, the *sn*-2 position is occupied by a polyunsaturated fatty acid (PUFA) and the *sn*-3 is occupied by one of the two polar head groups: ethanolamine or choline which give rise to pPE (plasmenylethanolamine) and pPC (plasmenylcholine) species respectively (Brites et al., 2004). These two types of plasmalogens exist in human tissues in different proportions where pPE is in general more abundant than pPC in most tissues with exception of the heart muscle where pPC dominates in abundance (Panganamala et al., 1971). In mammals, the heart, nervous tissues and inflammatory cells contain the highest concentrations of p lipid species (Diagne et al., 1984).

The abundance of p lipid species varies greatly among various tissues in humans (Nagan and Zoeller, 2001). Liver contains very little p species, but studies have shown that human serum contain significant levels of pPE (Brautigam et al., 1996; Nagan and Zoeller, 2001). . It has been shown that although there was little p lipid species in the liver, hepatocytes actively secrete nascent lipoproteins which contain up to 30% of pPE among total PE

(Vance, 1990). Furthermore, lipoproteins, contain a significant amount of pPE and pPC species (Wiesner et al., 2009).This could mean that liver cells target pPE for export with lipoproteins (Nagan and Zoeller, 2001), though the source of these lipid species for export could come from both endogenous biosynthesis as well as exogenously from the diet.

The physiological functions of p lipid species is yet to be clearly understood (Vance and Tasseva, 2013). Several roles have been proposed. Firstly, previous reports have shown that p species could serve to limit oxidative damage to cells by serving as antioxidants (Reiss et al., 1997; Zoeller et al., 1988) due to the higher susceptibility of the vinyl ether bond to oxidative attack compared to their 1-acyl analogues (Brosche and Platt, 1998; Mangold and Weber, 1987). Secondly, due to their vinyl ether bond, p lipid species have different physical properties and plays a role in membrane dynamics (Brites et al., 2004). p lipid species have been shown to play important roles in membrane fusion (Lohner, 1996) and are abundant in membrane systems that undergo rapid vesicular fusion events, though direct evidence for this is still lacking (Nagan and Zoeller, 2001). Thirdly, the high level of PUFA in p lipid species in the sn-2 position when compared to their diacyl counterparts has prompted the suggestion that they can act as storage depots to maintain high levels of PUFA (docosahexaenoic acid (DHA) and arachidonic acid (AA)) in some tissues (Gaposchkin and Zoeller, 1999; Nagan and Zoeller, 2001). This is consistent with our study where we observed a higher abundance of PUFA containing lipid species in Chinese for PE and PC when compared to the other two ethnic groups.

p lipid species have also been implicated in various pathological conditions, including Zellweger syndrome, Alzheimer's disease, retinitis pigmentosa and Down syndrome (Brites et al., 2004; Igarashi et al., 2011). Most of these conditions reflect a decrease in p lipid levels indicating the importance in maintaining normal levels for healthy physiological functions. Conditions like Alzheimer's have already been reported to have ethnic dependent differences in frequency across blacks and whites (de la Monte et al., 1989). It is possible to speculate that ethnic associated differences observed across the Singapore population may infer differences in underlying susceptibility to p lipid species associated pathological conditions due to significant variations in these lipid levels across the three ethnic groups, though much more biochemical studies on affected populations similar in composition to the Singaporean population need to be done before any valid and conclusive statements can be made.

Statistical validation of the observed significant differences across ethnicity was confirmed by permutation test. A permutation test can be performed in order to test whether differences found between comparison groups are significant (Bijlsma et al., 2006; Carina M. Rubingh, 2006). In the permutation test, lipid classification (carbon chain length, unsaturation, a/p species) were permuted 1000 times. A good permutation test result ($p < 0.005$) was obtained for LC PUFA containing lipid species of PS and PE, consistent with the type II ANOVA and t-test of significance results. Similarly, significant differences between Chinese when compared to Malay and Indian were also observed for pPE and pPC species, indicating that these trends could be

biologically relevant. While the reason for these interethnic differences in p lipid levels remains unknown, more biochemical based studies in similar populations need to be done to confirm these findings.

4.2.8.4.3 HDL/LDL correlate with p lipid species

Significant variations in HDL and LDL levels are observed across Indian and Malay population when compared to Chinese population despite similar levels of cholesterol across all three ethnicities. HDL/LDL ratio is commonly used as a biomarker for cardiovascular disease risk in clinical settings (Fernandez and Webb, 2008; NCEP, 2001). In this study, Indian population had the lowest HDL and highest LDL, giving rise to the lowest HDL/LDL ratio when compared to Malay and Chinese populations. Previous reports have showed that Indian have a high susceptibility towards cardovascular risk (Enas et al., 1992; McKeigue et al., 1989). Furthermore, previous studies done on the Singaporean population show that Indians confer a higher susceptibilty to cardiovascular pathologies when compared to Malay and Chinese populations, particularly Indians have a lower HDL than the other two ethnic groups which is comparable to data obtained in our study (Hughes, 1989; Hughes et al., 1997; Hughes et al., 1990a; Hughes et al., 1990b). Ethnic differences in cardiovascular risk has been reported before indicating Indians as a higher risk group for cardiovascular pathologies (Chaturvedi, 2003). This could mean that despite participants being healthy at

this point in the study, ethnic differences may exist due to their lipoprotein levels which may lead to differences in susceptibility to cardiovascular events, such that Indian population may be at a higher risk of getting cardiovascular events in future compared to the other two ethnic groups.

Previous studies on Chol measurements as well as disease risk associated with aberrant Chol levels have been focused mainly on Caucasian populations (Clausen et al., 1996; Nelson et al., 2002; Santosa et al., 2007; Varbo et al., 2011; Zhang et al., 2003) and there is a lack of published studies in Asian populations (APCSC, 2005; Lyratzopoulos et al., 2005; Shiwaku et al., 2005). As such, it is not surprising that most research on Chol associated therapies like development of statin drugs have also been focused on non-Asian populations (Baigent et al., 2005). The need to address population based studies in Asian populations is important due to reported differences in Chol levels and response to drugs that have been developed mainly in non Asian populations. In this study, lipid measurements were done in a healthy multi-ethnic Singaporean population which is representative of the three major ethnicities in Southeast Asia. We report similar Chol levels were observed across all three ethnic groups and this is in agreement to previous reports where no significant inter-ethnic differences were observed in serum cholesterol levels in the Singaporean population (Hughes et al., 1990b). Despite similar Chol levels, significant inter-ethnic variations were observed in HDL and LDL levels, which contribute to total Chol measurements and are well-known factors involved in affecting the risk and susceptibility to disease such as cardiovascular disease.

In our study, Chinese population had a higher abundance of p lipid species, higher HDL and lower LDL compared to Indian and Malay populations, Recent studies have been done correlating the p lipid species levels to lipoprotein concentrations (Brautigam et al., 1996; Maeba et al., 2007). One of the proposed roles of p lipid species is as an endogenous antioxidant due to the susceptibility of the vinyl ether bond to oxidative damage by peroxy radicals, UV and others (Engelmann, 2004; Maeba and Ueta, 2003; Nagan and Zoeller, 2001; Wallner and Schmitz, 2011). Liver has one of the lowest concentrations of p lipid species, but it has been shown that the liver selectively incorporates p lipid species into lipoproteins (Vance, 1990). This could be due to the ability of p lipid species to protect lipoproteins against oxidative damage. There is strong evidence that p lipid species are preferentially oxidized in conditions of oxidative stress, such as cardiovascular pathologies (Brosche and Platt, 1998; Engelmann et al., 1994; Maeba et al., 2007). Studies have shown that p lipid species act as endogenous antioxidants comparable to or even better than α-tocopherol (Vitamin E, lipidic antioxidant) as they slow down and impede the development of cardiovascular events like atherogenesis (Brautigam et al., 1996).

pPE has been reported to reduce susceptibility of cholesterol in membranes to oxidation by free radicals (Maeba and Ueta, 2003). Furthermore, the increase in p lipid species oxidation products such as 2-hydroxyaldehydes in LDL particles during conditions of oxidative stress indicate the higher susceptibility of LDL particles to oxidation and implications

in protection against cardiovascular disease outcomes (Regnstrom et al., 1992). While p lipid species in HDL have a lower susceptibility to oxidation as compared to LDL, they still play an important role especially in HDL mediated cellular Chol efflux where deficiency in p lipid levels can affect Chol efflux to HDL particles (Mandel et al., 1998; Munn et al., 2003). Correlation analyses between p lipid species and HDL/LDL in our study have shown that p lipid species is positively correlated with HDL and LDL with correlation coefficients that are similar to previously published literature (Jira and Spiteller, 1996; Maeba et al., 2007). As p lipid species have been shown to confer protection and increase oxidative resistance against oxidative stress, the inter-ethnic differences observed in this study in the healthy Singaporean population may be a useful indicator of future risk and susceptibility to metabolic disease caused by oxidative stress (Brosche and Platt, 1998; Jira and Spiteller, 1996).

4.2.8.5 Drawbacks of this study

While ESI-MS is a very promising tool for the analysis of PL, it is hampered by the fact that not all molecular species are detected with equal efficiency (Koivusalo et al., 2001). A previous study showed two major findings: first, the instrument response for both saturated and unsaturated PL species decreased prominently in long chain species with high lipid concentrations. Second, the degree of acyl chain unsaturation had a significant effect on instrument response and was concentration dependent (Koivusalo et al., 2001). It could be that the unsaturated species are more

surface active and thus become more efficiently ionized (Koivusalo et al., 2001). Additionally, the double bonds present in lipid species could weaken the intermolecular interactions in the droplet surface layer, enhancing the evaporation of the molecule to gas phase (Koivusalo et al., 2001). However, if this is the case, then all samples irrespective of ethnicity should exhibit a similar ionization pattern. All samples were randomized prior to MS analysis, thus, the trend of Chinese having higher amounts of PUFA long chain lipid species when compared to Malay and Indian is biologically relevant instead of being an artifact due to the analytical platform used.

It is known that PL can cause ion suppression or enhancement, adversely influencing the mass accuracy of the co-eluting species (particularly in positive ion mode). Crude lipid extract used in this study without solid phase extraction or LC-pre separation could increase matrix effects and reduce the sensitivity of the ions detected. Different biological samples contain different matrices, and this can lead to variability of ion suppression across different samples. This remains an issue for sample analysis, and several measures can be taken to account for this effect. For example, uniformized sample pre-processing and addition of multiple lipid standards per lipid class and subsequent normalization to these standards may reduce this variability as it works under assumption that lipid standards will ionize in different biological matrices similar to members of the same lipid class. Thus, normalization of lipid species to these lipid standards would help to reduce some differences in ion suppression across different biological matrices. However, only one lipid standard per lipid class was used in this study as it is currently not possible to

include internal standards representative of every lipid molecular entity (Stahlman et al., 2009). Furthermore, there is still a limited availability of standards for lipid analysis. In shotgun lipidomics approach, the absence of chromatographic separation result in all lipid species being subjected to the same matrix effect as they are analyzed simultaneously, whereas after chromatographic separation, matrix effect varies based on the different elution times. Though higher overall matrix effect may exist for shotgun approach, this effect is consistent for all samples analyzed (Jung et al., 2011).

Co-eluting species may interact with other LMW metabolites (Urpi-Sarda et al., 2009), often representing noisy variables (Tulipani et al., 2013). Thus, major lipid species like PC may prejudice LMW metabolite coverage and reproducibility, particularly for polar compounds, thus masking biological variation (Tulipani et al., 2013). It is possible that observations made in this study involving low abundant lipid species like GluCer and PA may not hold true, simply due to matrix effects from major lipid species. This can affect the detection of several lipid species whose intensities measured do not reflect their biological levels. For e.g. TAG, due to their high abundance in the human plasma, may act as a matrix for GPLs measured affecting the signals from less abundant classes like PS. However, the high significance observed could indeed be a true difference between the ethnicities but more studies are required to confirm these findings.

Quantification in lipidomics is usually achieved by comparison with internal standards which should ideally have similar chemical and physical

properties to the lipids to which they are being compared in order to minimize any difference in detection between the two, such as the use of stable-isotope labelled lipids or non-physiological lipids belonging to the same class (Peter Meikle, 2009). Usage of multiple lipid standards per lipid class is important as members of the same lipid class may have variations in their ionization pattern. In practice, one or two internal standards are used for a particular class of lipids (Peter Meikle, 2009). Due to limitations in this study, only one lipid standard was used for normalization of each class, which may not be optimal for all lipid species in that class.

4.2.8.6 Previous reports on ethnic associated lipid variation

Lipid levels have been demonstrated to vary between ethnic groups (Ford ES, 2002). Africans and East Asians have higher levels of HDL and lower levels of triglyceride compared to Europeans (Park et al., 2003). PUFA have been reported to be higher in Americans when compared to Italian and Finnish individuals (Dougherty et al., 1987). Not many studies have been done in ethnically diverse Asian populations. This study revealed interesting differences between the Chinese and other two ethnicities. However, in this study, FA analysis was not performed and thus ethnic associated changes observed in LC PUFA containing lipid species need to be interpreted with caution. Though many aspects could contribute to intra and inter-ethnic variability such as genetics and lifestyle (Winnike et al., 2009), the underlying mechanisms of these ethnic differences remain unknown (Elbers

et al., 2012). While evolutionary variations could explain some of our observations, more studies (including genotyping, transcriptomic profiling, GC-MS for FA profiling) need to be done to understand these variations at a genetic and biochemical level.

4.2.8.7 Impact of quantifying biological variation in an ethnically diverse population

Altered concentrations of circulating PL have been implicated in the pathology of type 2 diabetes, dyslipidemia and cardiovascular disease (Hodge et al., 2007a; Malerba et al., 2008; Wang et al., 2003), as well as a wide range of other common diseases including dementia and depression (Farooqui et al., 2000). Ethnic associated differences in metabolic diseases have been reported. A recent study done in the Malaysian population showed that Chinese had a lower prevalence of metabolic syndrome when compared to Indians and Malay (Rampal et al., 2012). In Singapore, previous reports have shown that Malay population has an increased risk of premenopausal breast cancer when compared to Indian and Chinese populations (Verkooijen et al., 2009). Furthermore, Chinese female population with diabetes was at higher risk of acquiring dengue hemorrhagic fever (serotype 2) when compared to the other two ethnicities (Pang et al., 2012). Understanding ethnic associated differences under various pathological conditions is important as it would aid in disease classification, identification and subsequent clinical interventions or

choice of therapy especially in an ethnically diverse population like in Singapore (Nicholson et al., 2012).

This study provides a descriptive analysis of the inter-ethnic differences associated with the natural biological variation across the healthy Singaporean population. The application of quantifying the degree of natural biological variation in healthy individuals is important. Firstly, understanding the natural variation will help to define boundaries for physiologically normal versus diseased individuals. Secondly, understanding differences in these variations across ethnicity will provide a useful tool in future diagnosis and treatment of pathological conditions. These findings will enable us to design and advance future functional studies to test certain hypothesis derived from this data, for example the association between high HDL levels in Chinese and efflux capacity of Chol. But to do this, we need to understand how diverse the HDL lipidome is among normal healthy individuals. Eventually, identifying the natural range of biological variation of various lipid species will be an important step towards understanding metabolic pathways contributing to common human conditions (Demirkan et al., 2012).

CHAPTER 5

CONCLUSIONS AND FUTURE PERSPECTIVES

5.1 Final Conclusions

This is the first comprehensive lipidomic profiling study done across a healthy Singaporean population. From this study, it can be concluded that study design methodology is essential in obtaining good quality samples for analysis. Technical variation is an important component of large population based studies where large sample numbers require standardized processing methods and reliable instrumentation. Technical variation due to analytical platforms like MS should be carefully considered in order to obtain data of good quality. After determining the technical capabilities of the MS used, biological variation can be considered.

Here, we report that two lipid classes, PE and PC show a wide range of technical variation which could be due to a large dynamic range and different abundance in human plasma. Batch effect correction is an important step in normalizing data in large studies where multiple batches need to be analyzed and sample processing and instrumentation are bound to give rise

to variations. After this has been taken into account, ethnic associated differences were observed across LC PUFA containing PE and PS lipid species where Chinese have a higher intensity when compared to Malay and Indians. Furthermore, Chinese have a significantly higher intensity of pPE and pPC lipid species when compared to the other two ethnicities. Why these differences are present is not known but we hypothesize that lifestyle factors like diet and evolutionary lineages could contribute to our observations.

5.2 Future Studies

Future studies include the need to replicate these findings in an independent validation study. Technical variation has to be established in a similar manner before considering biological differences. Lifestyle information obtained from this study should be analyzed to understand if they contribute significantly to the observed ethnic differences. Longitudinal studies looking at time-related differences in lipid profiles within an individual will help to define natural variation more precisely and ultimately lead to an understanding of metabolic space within a healthy ethnic group. FA profiling using GC-MS would be a useful way to determine FA composition to verify significant differences in PUFA content among different ethnic groups.

Cross platform analyses between genomics and lipidomics data obtained from these samples will help us to understand if genomic differences from lipid related genes contribute to the observed lipid variations across the three ethnicities (Wong et al., 2012). Furthermore, differences between ethnic groups could also be investigated at the level of epigenetic control. Availability and subcellular compartmentalization of lipids could contribute to the specificity of epigenetic control which conveys genome flexibility (Sassone-Corsi, 2013). Studies looking at histone modifications and microRNA pathways may provide useful information between ethnic associated regulatory differences.

Lipids are interlinked metabolically in many ways. While we are still beginning to unravel the complexities of these relationships and pathways, a useful way to understand these links is by looking at metabolite ratios (Gieger et al., 2008). Future studies can look at pairs of metabolites which are related to the direct substrates and products of an enzymatic conversion where the ratio between their concentrations can be used as an approximation of the enzymatic activity (Gieger et al., 2008). Furthermore, metabolite ratios often reduce the variability in a dataset while increasing the power of the study. This can help to identify pairs of metabolites that may be physiologically relevant across different ethnicities.

While we try to understand the biological variation across the Singaporean population, it is important to note that profiling these variations will help us to determine healthy natural variation ranges of measured lipids for each ethnic group. This is useful in future detection, diagnosis and treatment of pathological conditions especially when ethnic differences exist. Understanding healthy biological variation can aid in the development of future biomarkers for disease where lipid dysregulation play a role. This study serves as an important step towards our understanding of the biology of lipid variations across healthy individuals.

REFERENCES

A.Å. Karlsson, P.M., Å. Larsen, G. Odham (1996). Normal-phase liquid chromatography class separation and species determination of phospholipids utilizing electrospray mass spectrometry/tandem mass spectrometry. Rapid Commun Mass Spectrom *10*, 775–780.

Adibhatla, R.M., Hatcher, J.F., and Dempsey, R.J. (2006). Lipids and lipidomics in brain injury and diseases. The AAPS journal *8*, E314-321.

Akey, J.M., Biswas, S., Leek, J.T., and Storey, J.D. (2007). On the design and analysis of gene expression studies in human populations. Nature genetics *39*, 807-808; author reply 808-809.

Altmaier, E., Ramsay, S.L., Graber, A., Mewes, H.W., Weinberger, K.M., and Suhre, K. (2008). Bioinformatics analysis of targeted metabolomics--uncovering old and new tales of diabetic mice under medication. Endocrinology *149*, 3478-3489.

APCSC (2005). Joint effects of systolic blood pressure and serum cholesterol on cardiovascular disease in the Asia Pacific region. Circulation *112*, 3384-3390.

Assfalg, M., Bertini, I., Colangiuli, D., Luchinat, C., Schafer, H., Schutz, B., and Spraul, M. (2008). Evidence of different metabolic phenotypes in humans. Proceedings of the National Academy of Sciences of the United States of America *105*, 1420-1424.

Baggerly, K.A., Edmonson, S.R., Morris, J.S., and Coombes, K.R. (2004). High-resolution serum proteomic patterns for ovarian cancer detection. Endocrine-related cancer *11*, 583-584; author reply 585-587.

Baigent, C., Keech, A., Kearney, P.M., Blackwell, L., Buck, G., Pollicino, C., Kirby, A., Sourjina, T., Peto, R., Collins, R.*, et al.* (2005). Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet *366*, 1267-1278.

Bando, K., Kawahara, R., Kunimatsu, T., Sakai, J., Kimura, J., Funabashi, H., Seki, T., Bamba, T., and Fukusaki, E. (2010). Influences of biofluid sample collection and handling procedures on GC-MS based metabolomic studies. Journal of bioscience and bioengineering *110*, 491-499.

Bell, J.D., Sadler, P.J., Morris, V.C., and Levander, O.A. (1991). Effect of aging and diet on proton NMR spectra of rat urine. Magnetic resonance in

medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine *17*, 414-422.

Benton, H.P., Want, E., Keun, H.C., Amberg, A., Plumb, R.S., Goldfain-Blanc, F., Walther, B., Reily, M.D., Lindon, J.C., Holmes, E.*, et al.* (2012). Intra- and interlaboratory reproducibility of ultra performance liquid chromatography-time-of-flight mass spectrometry for urinary metabolic profiling. Analytical chemistry *84*, 2424-2432.

Bernini, P., Bertini, I., Luchinat, C., Nincheri, P., Staderini, S., and Turano, P. (2011). Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. Journal of biomolecular NMR *49*, 231-243.

Bijlsma, S., Bobeldijk, I., Verheij, E.R., Ramaker, R., Kochhar, S., Macdonald, I.A., van Ommen, B., and Smilde, A.K. (2006). Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. Analytical chemistry *78*, 567-574.

Blennow, K., Hampel, H., Weiner, M., and Zetterberg, H. (2010). Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. Nature reviews Neurology *6*, 131-144.

Bligh, E.G., and Dyer, W.J. (1959). A rapid method of total lipid extraction and purification. Canadian journal of biochemistry and physiology *37*, 911-917.

Bollard, M.E., Holmes, E., Lindon, J.C., Mitchell, S.C., Branstetter, D., Zhang, W., and Nicholson, J.K. (2001). Investigations into biochemical changes due to diurnal variation and estrus cycle in female rats using high-resolution (1)H NMR spectroscopy of urine and pattern recognition. Analytical biochemistry *295*, 194-202.

Bollard, M.E., Stanley, E.G., Lindon, J.C., Nicholson, J.K., and Holmes, E. (2005). NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. NMR in biomedicine *18*, 143-162.

Bou Khalil, M., Hou, W., Zhou, H., Elisma, F., Swayne, L.A., Blanchard, A.P., Yao, Z., Bennett, S.A., and Figeys, D. (2010). Lipidomics era: accomplishments and challenges. Mass spectrometry reviews *29*, 877-929.

Brautigam, C., Engelmann, B., Reiss, D., Reinhardt, U., Thiery, J., Richter, W.O., and Brosche, T. (1996). Plasmalogen phospholipids in plasma lipoproteins of normolipidemic donors and patients with hypercholesterolemia treated by LDL apheresis. Atherosclerosis *119*, 77-88.

Breusing, N., Grune, T., Andrisic, L., Atalay, M., Bartosz, G., Biasi, F., Borovic, S., Bravo, L., Casals, I., Casillas, R.*, et al.* (2010). An inter-laboratory validation of methods of lipid peroxidation measurement in UVA-treated human plasma samples. Free radical research *44*, 1203-1215.

Briche, C.S., Carter, D., and Webb, K.S. (2002). Comparison of gas chromatography and liquid chromatography mass spectrometric measurements for high accuracy analysis of cholesterol in human serum by isotope dilution mass spectrometry. Rapid communications in mass spectrometry : RCM *16*, 848-853.

Brinkmann, V. (2007). Sphingosine 1-phosphate receptors in health and disease: mechanistic insights from gene deletion studies and reverse pharmacology. Pharmacology & therapeutics *115*, 84-105.

Brites, P., Waterham, H.R., and Wanders, R.J. (2004). Functions and biosynthesis of plasmalogens in health and disease. Biochimica et biophysica acta *1636*, 219-231.

Brosche, T., and Platt, D. (1998). The biological significance of plasmalogens in defense against oxidative damage. Experimental gerontology *33*, 363-369.

Brouwers, J.F. (2011). Liquid chromatographic-mass spectrometric analysis of phospholipids. Chromatography, ionization and quantification. Biochimica et biophysica acta *1811*, 763-775.

Bruce, S.J., Tavazzi, I., Parisod, V., Rezzi, S., Kochhar, S., and Guy, P.A. (2009). Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. Analytical chemistry *81*, 3285-3296.

Brugger, B., Erben, G., Sandhoff, R., Wieland, F.T., and Lehmann, W.D. (1997). Quantitative analysis of biological membrane lipids at the low picomole level by nano-electrospray ionization tandem mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America *94*, 2339-2344.

Bruins, A.P. (1998). Mechanistic aspects of electrospray ionization. Journal of Chromatography A *794*, 345–357.

Burkard, I., von Eckardstein, A., Waeber, G., Vollenweider, P., and Rentsch, K.M. (2007). Lipoprotein distribution and biological variation of 24S- and 27-hydroxycholesterol in healthy volunteers. Atherosclerosis *194*, 71-78.

Byrdwell, W.C. (1998). Dual parallel mass spectrometers for analysis of sphingolipid, glycerophospholipid and plasmalogen molecular species. Rapid communications in mass spectrometry : RCM *12*, 256-272.

Calzada, C., Colas, R., Guillot, N., Guichardant, M., Laville, M., Vericel, E., and Lagarde, M. (2010). Subgram daily supplementation with docosahexaenoic acid protects low-density lipoproteins from oxidation in healthy men. Atherosclerosis *208*, 467-472.

Carina M. Rubingh, S.B., Eduard P. P. A. Derks, Ivana Bobeldijk, Elwin R. Verheij, Sunil Kochhar and Age K. Smilde (2006). Assessing the performance of statistical validation tools for megavariate metabolomics data. Metabolomics : Official journal of the Metabolomic Society *2*, 53-61.

Chan, R., Uchil, P.D., Jin, J., Shui, G., Ott, D.E., Mothes, W., and Wenk, M.R. (2008). Retroviruses human immunodeficiency virus and murine leukemia virus are enriched in phosphoinositides. J Virol *82*, 11228-11238.

Chaturvedi, N. (2003). Ethnic differences in cardiovascular disease. Heart *89*, 681-686.

Chau, C.H., Rixe, O., McLeod, H., and Figg, W.D. (2008). Validation of analytic methods for biomarkers used in drug development. Clinical cancer research : an official journal of the American Association for Cancer Research *14*, 5967-5976.

Chhabil, D. (2007). Mass analysis and ion detection In: Fundamentals of contemporary mass spectrometry, N.N. Desiderio DM, ed. (Hoboken, New Jersey: John Wiley and Sons Inc).

Clausen, J.O., Borch-Johnsen, K., Ibsen, H., Bergman, R.N., Hougaard, P., Winther, K., and Pedersen, O. (1996). Insulin sensitivity index, acute insulin response, and glucose effectiveness in a population-based sample of 380 young healthy Caucasians. Analysis of the impact of gender, body fat, physical fitness, and life-style factors. The Journal of clinical investigation *98*, 1195-1209.

Cook, H.W., Thomas, S.E., and Xu, Z. (1991). Essential fatty acids and serine as plasmalogen precursors in relation to competing metabolic pathways. Biochemistry and cell biology = Biochimie et biologie cellulaire *69*, 475-484.

Crews, B., Wikoff, W.R., Patti, G.J., Woo, H.K., Kalisiak, E., Heideker, J., and Siuzdak, G. (2009). Variability analysis of human plasma and cerebral spinal fluid reveals statistical significance of changes in mass spectrometry-based metabolomics data. Analytical chemistry *81*, 8538-8544.

Crutchfield, C.A., Lu, W., Melamud, E., and Rabinowitz, J.D. (2010). Mass spectrometry-based metabolomics of yeast. Methods in enzymology *470*, 393-426.

Dallmann, R., Viola, A.U., Tarokh, L., Cajochen, C., and Brown, S.A. (2012). The human circadian metabolome. Proceedings of the National Academy of Sciences of the United States of America *109*, 2625-2629.

Dashti, M., Kulik, W., Hoek, F., Veerman, E.C., Peppelenbosch, M.P., and Rezaee, F. (2011). A phospholipidomic analysis of all defined human plasma lipoproteins. Scientific reports *1*, 139.

de la Monte, S.M., Hutchins, G.M., and Moore, G.W. (1989). Racial differences in the etiology of dementia and frequency of Alzheimer lesions in the brain. Journal of the National Medical Association *81*, 644-652.

Demirkan, A., Amin, N., Isaacs, A., Jarvelin, M.R., Whitfield, J.B., Wichmann, H.E., Kyvik, K.O., Rudan, I., Gieger, C., Hicks, A.A.*, et al.* (2011). Genetic architecture of circulating lipid levels. European journal of human genetics : EJHG *19*, 813-819.

Demirkan, A., van Duijn, C.M., Ugocsai, P., Isaacs, A., Pramstaller, P.P., Liebisch, G., Wilson, J.F., Johansson, A., Rudan, I., Aulchenko, Y.S.*, et al.* (2012). Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. PLoS genetics *8*, e1002490.

Diagne, A., Fauvel, J., Record, M., Chap, H., and Douste-Blazy, L. (1984). Studies on ether phospholipids. II. Comparative composition of various tissues from human, rat and guinea pig. Biochimica et biophysica acta *793*, 221-231.

Dittmer, J.C., and Lester, R.L. (1964). A Simple, Specific Spray for the Detection of Phospholipids on Thin-Layer Chromatograms. Journal of lipid research *5*, 126-127.

Dougherty, R.M., Galli, C., Ferro-Luzzi, A., and Iacono, J.M. (1987). Lipid and phospholipid fatty acid composition of plasma, red blood cells, and platelets and how they are affected by dietary lipids: a study of normal subjects from Italy, Finland, and the USA. The American journal of clinical nutrition *45*, 443-455.

Dunn, W.B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J.D., Halsall, A., Haselden, J.N.*, et al.* (2011). Procedures for large-scale metabolic profiling of serum and plasma

using gas chromatography and liquid chromatography coupled to mass spectrometry. Nature protocols *6*, 1060-1083.

Dunn, W.B., Wilson, I.D., Nicholls, A.W., and Broadhurst, D. (2012). The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. Bioanalysis *4*, 2249-2264.

E.Vance, D.E.V.a.J. (2008). Biochemistry of lipids, lipoproteins and membranes, Fifth Edition edn (Elsevier).

Edmond de Hoffmann, V.S. (2007). Mass Spectrometry, Principles and Applications, Third edn (John Wiley and Sons Ltd).

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America *95*, 14863-14868.

Ekman, S., Westman-Brinkmalm, and Kraj (2009). Mass Spectrometry (John Wiley and Sons Ltd), pp. 1-275.

Ekroos, K., Janis, M., Tarasov, K., Hurme, R., and Laaksonen, R. (2010). Lipidomics: a tool for studies of atherosclerosis. Current atherosclerosis reports *12*, 273-281.

Elbers, C.C., Guo, Y., Tragante, V., van Iperen, E.P., Lanktree, M.B., Castillo, B.A., Chen, F., Yanek, L.R., Wojczynski, M.K., Li, Y.R.*, et al.* (2012). Gene-centric meta-analysis of lipid traits in african, East asian and Hispanic populations. PloS one *7*, e50198.

Enas, E.A., Yusuf, S., and Mehta, J.L. (1992). Prevalence of coronary artery disease in Asian Indians. The American journal of cardiology *70*, 945-949.

Engelmann, B. (2004). Plasmalogens: targets for oxidants and major lipophilic antioxidants. Biochemical Society transactions *32*, 147-150.

Engelmann, B., Brautigam, C., and Thiery, J. (1994). Plasmalogen phospholipids as potential protectors against lipid peroxidation of low density lipoproteins. Biochemical and biophysical research communications *204*, 1235-1242.

Fahy, E., Cotter, D., Sud, M., and Subramaniam, S. (2011). Lipid classification, structures and tools. Biochimica et biophysica acta *1811*, 637-647.

Fahy, E., Subramaniam, S., Murphy, R.C., Nishijima, M., Raetz, C.R., Shimizu, T., Spener, F., van Meer, G., Wakelam, M.J., and Dennis, E.A. (2009). Update of the LIPID MAPS comprehensive classification system for lipids. Journal of lipid research *50 Suppl*, S9-14.

Farooqui, A.A., Horrocks, L.A., and Farooqui, T. (2000). Glycerophospholipids in brain: their metabolism, incorporation into membranes, functions, and involvement in neurological disorders. Chemistry and physics of lipids *106*, 1-29.

Fenn JB, M.M., Meng CK,Wong SF, Whitehouse CM (1989). Electrospray ionization for mass spectrometry of large biomolecules. Science *246*, 64–71.

Fernandez, M.L., and Webb, D. (2008). The LDL to HDL cholesterol ratio as a valuable tool to evaluate coronary heart disease risk. Journal of the American College of Nutrition *27*, 1-5.

Fernandis, A.Z., and Wenk, M.R. (2007). Membrane lipids as signaling molecules. Current opinion in lipidology *18*, 121-128.

Ford ES, G.W., Dietz WH (2002). Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. JAMA *287*, 356–359.

Gao, J., Garulacan, L.A., Storm, S.M., Opiteck, G.J., Dubaquie, Y., Hefta, S.A., Dambach, D.M., and Dongre, A.R. (2005). Biomarker discovery in biological fluids. Methods *35*, 291-302.

Gaposchkin, D.P., and Zoeller, R.A. (1999). Plasmalogen status influences docosahexaenoic acid levels in a macrophage cell line. Insights using ether lipid-deficient variants. Journal of lipid research *40*, 495-503.

Gavaghan, C.L., Wilson, I.D., and Nicholson, J.K. (2002). Physiological variation in metabolic phenotyping and functional genomic studies: use of orthogonal signal correction and PLS-DA. FEBS letters *530*, 191-196.

Giddings, J.C. (1965). Comparison of Theoretical Limit of Separating Speed in Gas and Liquid Chromatography. Analytical chemistry *37*, 60-&.

Gieger, C., Geistlinger, L., Altmaier, E., Hrabe de Angelis, M., Kronenberg, F., Meitinger, T., Mewes, H.W., Wichmann, H.E., Weinberger, K.M., Adamski, J.*, et al.* (2008). Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS genetics *4*, e1000282.

Gika, H.G., Theodoridis, G.A., Earll, M., and Wilson, I.D. (2012). A QC approach to the determination of day-to-day reproducibility and robustness of LC-MS methods for global metabolite profiling in metabonomics/metabolomics. Bioanalysis *4*, 2239-2247.

Glaser, C., Heinrich, J., and Koletzko, B. (2010). Role of FADS1 and FADS2 polymorphisms in polyunsaturated fatty acid metabolism. Metabolism: clinical and experimental *59*, 993-999.

Graessler, J., Schwudke, D., Schwarz, P.E., Herzog, R., Shevchenko, A., and Bornstein, S.R. (2009). Top-down lipidomics reveals ether lipid deficiency in blood plasma of hypertensive patients. PloS one *4*, e6261.

Gross, J.H. (2011). Mass Spectrometry : A textbook, 2nd Edition edn (Springer).

Gross, R.W., and Han, X. (2011). Lipidomics at the interface of structure and function in systems biology. Chemistry & biology *18*, 284-291.

Gross, R.W.a.H., X. (2006). Unlocking the complexity of lipids: Using lipidomics to identify disease mechanisms, biomarkers, and treatment efficacy. Future Lipidology *1*, 539-547.

Guan, X.L., He, X., Ong, W.Y., Yeo, W.K., Shui, G., and Wenk, M.R. (2006). Non-targeted profiling of lipids during kainate-induced neuronal injury. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *20*, 1152-1161.

Hammad, S.M., Pierce, J.S., Soodavar, F., Smith, K.J., Al Gadban, M.M., Rembiesa, B., Klein, R.L., Hannun, Y.A., Bielawski, J., and Bielawska, A. (2010). Blood sphingolipidomics in healthy humans: impact of sample collection methodology. Journal of lipid research *51*, 3074-3087.

Han, X. (2010). Multi-dimensional mass spectrometry-based shotgun lipidomics and the altered lipids at the mild cognitive impairment stage of Alzheimer's disease. Biochimica et biophysica acta *1801*, 774-783.

Han, X., and Gross, R.W. (1994). Electrospray ionization mass spectroscopic analysis of human erythrocyte plasma membrane phospholipids. Proceedings of the National Academy of Sciences of the United States of America *91*, 10635-10639.

Han, X., and Gross, R.W. (2005). Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. Mass spectrometry reviews *24*, 367-412.

Han, X., Yang, K., Cheng, H., Fikes, K.N., and Gross, R.W. (2005). Shotgun lipidomics of phosphoethanolamine-containing lipids in biological samples after one-step in situ derivatization. Journal of lipid research *46*, 1548-1560.

Harkewicz, R., and Dennis, E.A. (2011). Applications of mass spectrometry to lipids and membranes. Annual review of biochemistry *80*, 301-325.

Hicks, A.A., Pramstaller, P.P., Johansson, A., Vitart, V., Rudan, I., Ugocsai, P., Aulchenko, Y., Franklin, C.S., Liebisch, G., Erdmann, J.*, et al.* (2009). Genetic determinants of circulating sphingolipid concentrations in European populations. PLoS genetics *5*, e1000672.

Hodge, A.M., English, D.R., O'Dea, K., Sinclair, A.J., Makrides, M., Gibson, R.A., and Giles, G.G. (2007a). Plasma phospholipid and dietary fatty acids as predictors of type 2 diabetes: interpreting the role of linoleic acid. The American journal of clinical nutrition *86*, 189-197.

Hodge, A.M., Simpson, J.A., Gibson, R.A., Sinclair, A.J., Makrides, M., O'Dea, K., English, D.R., and Giles, G.G. (2007b). Plasma phospholipid fatty acid composition as a biomarker of habitual dietary fat intake in an ethnically diverse cohort. Nutrition, metabolism, and cardiovascular diseases : NMCD *17*, 415-426.

Honda, A., Yamashita, K., Miyazaki, H., Shirai, M., Ikegami, T., Xu, G., Numazawa, M., Hara, T., and Matsuzaki, Y. (2008). Highly sensitive analysis of sterol profiles in human serum by LC-ESI-MS/MS. Journal of lipid research *49*, 2063-2073.

Horwitz, W. (1982). Evaluation of analytical methods used for regulation of foods and drugs. Analytical chemistry *54*, 67-76.

Houshyani, B., Kabouw, P., Muth, D., de Vos, R.C., Bino, R.J., and Bouwmeester, H.J. (2012). Characterization of the natural variation in Arabidopsis thaliana metabolome by the analysis of metabolic distance. Metabolomics : Official journal of the Metabolomic Society *8*, 131-145.

Hsu, F.F., and Turk, J. (2000a). Characterization of phosphatidylinositol, phosphatidylinositol-4-phosphate, and phosphatidylinositol-4,5-bisphosphate by electrospray ionization tandem mass spectrometry: a mechanistic study. Journal of the American Society for Mass Spectrometry *11*, 986-999.

Hsu, F.F., and Turk, J. (2000b). Charge-driven fragmentation processes in diacyl glycerophosphatidic acids upon low-energy collisional activation. A mechanistic proposal. Journal of the American Society for Mass Spectrometry *11*, 797-803.

Hu, C., van der Heijden, R., Wang, M., van der Greef, J., Hankemeier, T., and Xu, G. (2009). Analytical strategies in lipidomics and applications in disease biomarker discovery. Journal of chromatography B, Analytical technologies in the biomedical and life sciences *877*, 2836-2846.

Hughes, K. (1989). Mortality from cardiovascular diseases in Chinese, Malays and Indians in Singapore, in comparison with England and Wales, USA and Japan. Annals of the Academy of Medicine, Singapore *18*, 642-645.

Hughes, K., Aw, T.C., Kuperan, P., and Choo, M. (1997). Central obesity, insulin resistance, syndrome X, lipoprotein(a), and cardiovascular risk in Indians, Malays, and Chinese in Singapore. Journal of epidemiology and community health *51*, 394-399.

Hughes, K., Lun, K.C., and Yeo, P.P. (1990a). Cardiovascular diseases in Chinese, Malays, and Indians in Singapore. I. Differences in mortality. Journal of epidemiology and community health *44*, 24-28.

Hughes, K., Yeo, P.P., Lun, K.C., Thai, A.C., Sothy, S.P., Wang, K.W., Cheah, J.S., Phoon, W.O., and Lim, P. (1990b). Cardiovascular diseases in Chinese, Malays, and Indians in Singapore. II. Differences in risk factor levels. Journal of epidemiology and community health *44*, 29-35.

Igarashi, M., Ma, K., Gao, F., Kim, H.W., Rapoport, S.I., and Rao, J.S. (2011). Disturbed choline plasmalogen and phospholipid fatty acid concentrations in Alzheimer's disease prefrontal cortex. Journal of Alzheimer's disease : JAD *24*, 507-517.

Isaac, G. (2011). Electrospray ionization tandem mass spectrometry (ESI-MS/MS)-based shotgun lipidomics. Methods in Molecular Biology *708*, 259-275.

Isaac, G., Jeannotte, R., Esch, S.W., and Welti, R. (2007). New mass-spectrometry-based strategies for lipids. Genetic engineering *28*, 129-157.

Ismaiel, O.A., Halquist, M.S., Elmamly, M.Y., Shalaby, A., and Thomas Karnes, H. (2008). Monitoring phospholipids for assessment of ion enhancement and ion suppression in ESI and APCI LC/MS/MS for chlorpheniramine in human plasma and the importance of multiple source matrix effect evaluations. Journal of chromatography B, Analytical technologies in the biomedical and life sciences *875*, 333-343.

James, A.T., Martin, A.J.P., and Smith, G.H. (1952). Gas-Liquid Partition Chromatography - the Separation and Micro-Estimation of Ammonia and the Methylamines. Biochemical Journal *52*, 238-242.

Jim Mann, M.S. (2001). Triacylglycerols. In Encyclopedia of Life Sciences (John Wiley and Sons Ltd).

Jira, W., and Spiteller, G. (1996). Plasmalogens and their oxidative degradation products in low and high density lipoprotein. Chemistry and physics of lipids *79*, 95-100.

Jung, H.R., Sylvanne, T., Koistinen, K.M., Tarasov, K., Kauhanen, D., and Ekroos, K. (2011). High throughput quantitative molecular lipidomics. Biochimica et biophysica acta *1811*, 925-934.

Kamleh, M.A., Ebbels, T.M., Spagou, K., Masson, P., and Want, E.J. (2012). Optimizing the use of quality control samples for signal drift correction in large-scale urine metabolic profiling studies. Analytical chemistry *84*, 2670-2677.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research *28*, 27-30.

Kathiresan, S., Melander, O., Guiducci, C., Surti, A., Burtt, N.P., Rieder, M.J., Cooper, G.M., Roos, C., Voight, B.F., Havulinna, A.S.*, et al.* (2008). Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nature genetics *40*, 189-197.

Koivusalo, M., Haimi, P., Heikinheimo, L., Kostiainen, R., and Somerharju, P. (2001). Quantitative determination of phospholipid compositions by ESI-MS: effects of acyl chain length, unsaturation, and lipid concentration on instrument response. Journal of lipid research *42*, 663-672.

Kontush, A., Lhomme, M., and Chapman, M.J. (2013). Unravelling the Complexities of the HDL Lipidome. Journal of lipid research.

Krug, S., Kastenmuller, G., Stuckler, F., Rist, M.J., Skurk, T., Sailer, M., Raffler, J., Romisch-Margl, W., Adamski, J., Prehn, C.*, et al.* (2012). The dynamic range of the human metabolome revealed by challenges. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *26*, 2607-2619.

Kwiterovich, P.O., Jr. (2000). The metabolic pathways of high-density lipoprotein, low-density lipoprotein, and triglycerides: a current review. The American journal of cardiology *86*, 5L-10L.

Lattka, E., Illig, T., Koletzko, B., and Heinrich, J. (2010). Genetic variants of the FADS1 FADS2 gene cluster as related to essential fatty acid metabolism. Current opinion in lipidology *21*, 64-69.

Lawton, K.A., Berger, A., Mitchell, M., Milgram, K.E., Evans, A.M., Guo, L., Hanson, R.W., Kalhan, S.C., Ryals, J.A., and Milburn, M.V. (2008). Analysis of the adult human plasma metabolome. Pharmacogenomics *9*, 383-397.

Leek, J.T., Scharpf, R.B., Bravo, H.C., Simcha, D., Langmead, B., Johnson, W.E., Geman, D., Baggerly, K., and Irizarry, R.A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. Nature reviews Genetics *11*, 733-739.

Lei, Z., Huhman, D.V., and Sumner, L.W. (2011). Mass spectrometry strategies in metabolomics. The Journal of biological chemistry *286*, 25435-25442.

Li, Z., Otvos, J.D., Lamon-Fava, S., Carrasco, W.V., Lichtenstein, A.H., McNamara, J.R., Ordovas, J.M., and Schaefer, E.J. (2003). Men and women differ in lipoprotein response to dietary saturated fat and cholesterol restriction. The Journal of nutrition *133*, 3428-3433.

Liebisch, G., Lieser, B., Rathenberg, J., Drobnik, W., and Schmitz, G. (2004). High-throughput quantification of phosphatidylcholine and sphingomyelin by electrospray ionization tandem mass spectrometry coupled with isotope correction algorithm. Biochimica et biophysica acta *1686*, 108-117.

Lin, J., Hu, F.B., Rimm, E.B., Rifai, N., and Curhan, G.C. (2006). The association of serum lipids and inflammatory biomarkers with renal function in men with type II diabetes mellitus. Kidney international *69*, 336-342.

Lindon, J.C., Keun, H.C., Ebbels, T.M., Pearce, J.M., Holmes, E., and Nicholson, J.K. (2005). The Consortium for Metabonomic Toxicology (COMET): aims, activities and achievements. Pharmacogenomics *6*, 691-699.

Lohner, K. (1996). Is the high propensity of ethanolamine plasmalogens to form non-lamellar lipid structures manifested in the properties of biomembranes? Chemistry and physics of lipids *81*, 167-184.

Luo, J., Schumacher, M., Scherer, A., Sanoudou, D., Megherbi, D., Davison, T., Shi, T., Tong, W., Shi, L., Hong, H.*, et al.* (2010). A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data. The pharmacogenomics journal *10*, 278-291.

Lyratzopoulos, G., McElduff, P., Heller, R.F., Hanily, M., and Lewis, P.S. (2005). Comparative levels and time trends in blood pressure, total cholesterol, body mass index and smoking among Caucasian and South-Asian participants of a UK primary-care based cardiovascular risk factor screening programme. BMC public health *5*, 125.

Lytle, C.A., Gan, Y.D., and White, D.C. (2000). Electrospray ionization/mass spectrometry compatible reversed-phase separation of phospholipids: piperidine as a post column modifier for negative ion detection. Journal of microbiological methods *41*, 227-234.

Maeba, R., Maeda, T., Kinoshita, M., Takao, K., Takenaka, H., Kusano, J., Yoshimura, N., Takeoka, Y., Yasuda, D., Okazaki, T.*, et al.* (2007). Plasmalogens in human serum positively correlate with high- density lipoprotein and decrease with aging. Journal of atherosclerosis and thrombosis *14*, 12-18.

Maeba, R., and Ueta, N. (2003). Ethanolamine plasmalogen and cholesterol reduce the total membrane oxidizability measured by the oxygen uptake method. Biochemical and biophysical research communications *302*, 265-270.

Magnusson, C.D., and Haraldsson, G.G. (2011). Ether lipids. Chemistry and physics of lipids *164*, 315-340.

Malerba, G., Schaeffer, L., Xumerle, L., Klopp, N., Trabetti, E., Biscuola, M., Cavallari, U., Galavotti, R., Martinelli, N., Guarini, P.*, et al.* (2008). SNPs of the FADS gene cluster are associated with polyunsaturated fatty acids in a cohort of patients with cardiovascular disease. Lipids *43*, 289-299.

Mamas, M., Dunn, W.B., Neyses, L., and Goodacre, R. (2011). The role of metabolites and metabolomics in clinically applicable biomarkers of disease. Archives of toxicology *85*, 5-17.

Mandel, H., Sharf, R., Berant, M., Wanders, R.J., Vreken, P., and Aviram, M. (1998). Plasmalogen phospholipids are involved in HDL-mediated cholesterol efflux: insights from investigations with plasmalogen-deficient cells. Biochemical and biophysical research communications *250*, 369-373.

Mangold, H.K., and Weber, N. (1987). Biosynthesis and biotransformation of ether lipids. Lipids *22*, 789-799.

Masood, M.A., Rao, R.P., Acharya, J.K., Blonder, J., and Veenstra, T.D. (2012). Quantitation of multiple sphingolipid classes using normal and reversed-phase LC-ESI-MS/MS: comparative profiling of two cell lines. Lipids *47*, 209-226.

Masson, P., Spagou, K., Nicholson, J.K., and Want, E.J. (2011). Technical and biological variation in UPLC-MS-based untargeted metabolic profiling of liver extracts: application in an experimental toxicity study on galactosamine. Analytical chemistry *83*, 1116-1123.

McKeigue, P.M., Miller, G.J., and Marmot, M.G. (1989). Coronary heart disease in south Asians overseas: a review. Journal of clinical epidemiology *42*, 597-609.

Merino, D.M., Ma, D.W., and Mutch, D.M. (2010). Genetic variation in lipid desaturases and its impact on the development of human disease. Lipids in health and disease *9*, 63.

Merrill, A.H., Jr. (2011). Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. Chemical reviews *111*, 6387-6422.

Merrill, A.H., Jr., Sullards, M.C., Allegood, J.C., Kelly, S., and Wang, E. (2005). Sphingolipidomics: high-throughput, structure-specific, and quantitative analysis of sphingolipids by liquid chromatography tandem mass spectrometry. Methods *36*, 207-224.

Minami, Y., Kasukawa, T., Kakazu, Y., Iigo, M., Sugimoto, M., Ikeda, S., Yasui, A., van der Horst, G.T., Soga, T., and Ueda, H.R. (2009). Measurement of internal body time by blood metabolomics. Proceedings of the National Academy of Sciences of the United States of America *106*, 9890-9895.

Molloy, M.P., Brzezinski, E.E., Hang, J., McDowell, M.T., and VanBogelen, R.A. (2003). Overcoming technical variation and biological variation in quantitative proteomics. Proteomics *3*, 1912-1919.

Munn, N.J., Arnio, E., Liu, D., Zoeller, R.A., and Liscum, L. (2003). Deficiency in ethanolamine plasmalogen leads to altered cholesterol transport. Journal of lipid research *44*, 182-192.

Nagan, N., Hajra, A.K., Larkins, L.K., Lazarow, P., Purdue, P.E., Rizzo, W.B., and Zoeller, R.A. (1998). Isolation of a Chinese hamster fibroblast variant defective in dihydroxyacetonephosphate acyltransferase activity and plasmalogen biosynthesis: use of a novel two-step selection protocol. The Biochemical journal *332 (Pt 1)*, 273-279.

Nagan, N., and Zoeller, R.A. (2001). Plasmalogens: biosynthesis and functions. Progress in lipid research *40*, 199-229.

Nakahara, K., Ohkuni, A., Kitamura, T., Abe, K., Naganuma, T., Ohno, Y., Zoeller, R.A., and Kihara, A. (2012). The Sjogren-Larsson syndrome gene encodes a hexadecenal dehydrogenase of the sphingosine 1-phosphate degradation pathway. Molecular cell *46*, 461-471.

NCEP, A.T.P.I.E.s.o.t.t.r.o.t.n.c.e.p.N.e.p.o.d.e., and treatment of high blood cholesterol in adults (2001). Executive Summary of The Third Report of The

National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). JAMA *285*, 2486-2497.

Nelson, K., Norris, K., and Mangione, C.M. (2002). Disparities in the diagnosis and pharmacologic treatment of high serum cholesterol by race and ethnicity: data from the Third National Health and Nutrition Examination Survey. Archives of internal medicine *162*, 929-935.

Nicholson, J.K., Holmes, E., Kinross, J.M., Darzi, A.W., Takats, Z., and Lindon, J.C. (2012). Metabolic phenotyping in clinical and surgical environments. Nature *491*, 384-392.

Nygren, H., Seppanen-Laakso, T., Castillo, S., Hyotylainen, T., and Oresic, M. (2011). Liquid chromatography-mass spectrometry (LC-MS)-based lipidomics for studies of body fluids and tissues. Methods Mol Biol *708*, 247-257.

Pang, J., Salim, A., Lee, V.J., Hibberd, M.L., Chia, K.S., Leo, Y.S., and Lye, D.C. (2012). Diabetes with hypertension as risk factors for adult dengue hemorrhagic fever in a predominantly dengue serotype 2 epidemic: a case control study. PLoS neglected tropical diseases *6*, e1641.

Panganamala, R.V., Horrocks, L.A., Geer, J.C., and Cornwell, D.G. (1971). Positions of double bonds in the monounsaturated alk-1-enyl groups from the plasmalogens of human heart and brain. Chemistry and physics of lipids *6*, 97-102.

Park, Y.W., Zhu, S., Palaniappan, L., Heshka, S., Carnethon, M.R., and Heymsfield, S.B. (2003). The metabolic syndrome: prevalence and associated risk factor findings in the US population from the Third National Health and Nutrition Examination Survey, 1988-1994. Archives of internal medicine *163*, 427-436.

Peter Meikle, C.B.a.J.W. (2009). Lipidomics and lipid biomarker discovery. Australian Biochemist *40*, 12-16.

Petricoin, E.F., Ardekani, A.M., Hitt, B.A., Levine, P.J., Fusaro, V.A., Steinberg, S.M., Mills, G.B., Simone, C., Fishman, D.A., Kohn, E.C.*, et al.* (2002). Use of proteomic patterns in serum to identify ovarian cancer. Lancet *359*, 572-577.

Pulfer, M., and Murphy, R.C. (2003). Electrospray mass spectrometry of phospholipids. Mass spectrometry reviews *22*, 332-364.

Quehenberger, O., Armando, A.M., Brown, A.H., Milne, S.B., Myers, D.S., Merrill, A.H., Bandyopadhyay, S., Jones, K.N., Kelly, S., Shaner, R.L.*, et al.*

(2010). Lipidomics reveals a remarkable diversity of lipids in human plasma. Journal of lipid research *51*, 3299-3305.

Radich, J.P., Mao, M., Stepaniants, S., Biery, M., Castle, J., Ward, T., Schimmack, G., Kobayashi, S., Carleton, M., Lampe, J.*, et al.* (2004). Individual-specific variation of gene expression in peripheral blood leukocytes. Genomics *83*, 980-988.

Raith, K., Brenner, C., Farwanah, H., Muller, G., Eder, K., and Neubert, R.H. (2005). A new LC/APCI-MS method for the determination of cholesterol oxidation products in food. Journal of chromatography A *1067*, 207-211.

Rampal, S., Mahadeva, S., Guallar, E., Bulgiba, A., Mohamed, R., Rahmat, R., Arif, M.T., and Rampal, L. (2012). Ethnic differences in the prevalence of metabolic syndrome: results from a multi-ethnic population-based survey in Malaysia. PloS one *7*, e46365.

Regnstrom, J., Nilsson, J., Tornvall, P., Landou, C., and Hamsten, A. (1992). Susceptibility to low-density lipoprotein oxidation and coronary atherosclerosis in man. Lancet *339*, 1183-1186.

Reiss, D., Beyer, K., and Engelmann, B. (1997). Delayed oxidative degradation of polyunsaturated diacyl phospholipids in the presence of plasmalogen phospholipids in vitro. The Biochemical journal *323 (Pt 3)*, 807-814.

Rivera, J., and Olivera, A. (2007). Src family kinases and lipid mediators in control of allergic inflammation. Immunological reviews *217*, 255-268.

Roberts, L.D., Souza, A.L., Gerszten, R.E., and Clish, C.B. (2012). Targeted metabolomics. Current protocols in molecular biology / edited by Frederick M Ausubel [et al] *Chapter 30*, Unit 30.32.31-24.

Rudkowska, I., Paradis, A.M., Thifault, E., Julien, P., Tchernof, A., Couture, P., Lemieux, S., Barbier, O., and Vohl, M.C. (2013). Transcriptomic and metabolomic signatures of an n-3 polyunsaturated fatty acids supplementation in a normolipidemic/normocholesterolemic Caucasian population. The Journal of nutritional biochemistry *24*, 54-61.

Rutherford, J.N., McDade, T.W., Feranil, A.B., Adair, L.S., and Kuzawa, C.W. (2010). High prevalence of low HDL-c in the Philippines compared to the US: population differences in associations with diet and BMI. Asia Pacific journal of clinical nutrition *19*, 57-67.

Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics *20*, 3246-3248.

Sandra, K., Pereira Ados, S., Vanhoenacker, G., David, F., and Sandra, P. (2010). Comprehensive blood plasma lipidomics by liquid chromatography/quadrupole time-of-flight mass spectrometry. Journal of chromatography A *1217*, 4087-4099.

Santosa, S., Varady, K.A., AbuMweis, S., and Jones, P.J. (2007). Physiological and therapeutic factors affecting cholesterol metabolism: does a reciprocal relationship between cholesterol absorption and synthesis really exist? Life sciences *80*, 505-514.

Sassone-Corsi, P. (2013). Physiology. When metabolism and epigenetics converge. Science *339*, 148-150.

Sato, Y., Nakamura, T., Aoshima, K., and Oda, Y. (2010). Quantitative and wide-ranging profiling of phospholipids in human plasma by two-dimensional liquid chromatography/mass spectrometry. Analytical chemistry *82*, 9858-9864.

Schaeffer, L., Gohlke, H., Muller, M., Heid, I.M., Palmer, L.J., Kompauer, I., Demmelmair, H., Illig, T., Koletzko, B., and Heinrich, J. (2006). Common genetic variants of the FADS1 FADS2 gene cluster and their reconstructed haplotypes are associated with the fatty acid composition in phospholipids. Human molecular genetics *15*, 1745-1756.

Schreiner, P.J., Heiss, G., Tyroler, H.A., Morrisett, J.D., Davis, C.E., and Smith, R. (1996). Race and gender differences in the association of Lp(a) with carotid artery wall thickness. The Atherosclerosis Risk in Communities (ARIC) Study. Arteriosclerosis, thrombosis, and vascular biology *16*, 471-478.

Schwarz, E., Prabakaran, S., Whitfield, P., Major, H., Leweke, F.M., Koethe, D., McKenna, P., and Bahn, S. (2008). High throughput lipidomic profiling of schizophrenia and bipolar disorder brain tissue reveals alterations of free fatty acids, phosphatidylcholines, and ceramides. Journal of proteome research *7*, 4266-4277.

Sevanian, A., Seraglia, R., Traldi, P., Rossato, P., Ursini, F., and Hodis, H. (1994). Analysis of plasma cholesterol oxidation products using gas- and high-performance liquid chromatography/mass spectrometry. Free radical biology & medicine *17*, 397-409.

Shevchenko, A., and Simons, K. (2010). Lipidomics: coming to grips with lipid diversity. Nature reviews Molecular cell biology *11*, 593-598.

Shiwaku, K., Anuurad, E., Enkhmaa, B., Nogi, A., Kitajima, K., Yamasaki, M., Yoneyama, T., Oyunsuren, T., and Yamane, Y. (2005). Predictive values of

anthropometric measurements for multiple metabolic disorders in Asian populations. Diabetes research and clinical practice *69*, 52-62.

Shui, G., Bendt, A.K., Pethe, K., Dick, T., and Wenk, M.R. (2007). Sensitive profiling of chemically diverse bioactive lipids. Journal of lipid research *48*, 1976-1984.

Shui, G., Cheong, W.F., Jappar, I.A., Hoi, A., Xue, Y., Fernandis, A.Z., Tan, B.K., and Wenk, M.R. (2011a). Derivatization-independent cholesterol analysis in crude lipid extracts by liquid chromatography/mass spectrometry: applications to a rabbit model for atherosclerosis. Journal of chromatography A *1218*, 4357-4365.

Shui, G., Stebbins, J.W., Lam, B.D., Cheong, W.F., Lam, S.M., Gregoire, F., Kusonoki, J., and Wenk, M.R. (2011b). Comparative plasma lipidome between human and cynomolgus monkey: are plasma polar lipids good biomarkers for diabetic monkeys? PloS one *6*, e19731.

Skipski, V.P. (1972). Blood Lipids and Lipoproteins: Quantitalion, Composition and Metabolism (New York.: Wiley-lnterscience).

Spiegel, S., and Milstien, S. (2003). Sphingosine-1-phosphate: an enigmatic signalling lipid. Nature reviews Molecular cell biology *4*, 397-407.

Spielman, R.S., Bastone, L.A., Burdick, J.T., Morley, M., Ewens, W.J., and Cheung, V.G. (2007). Common genetic variants account for differences in gene expression among ethnic groups. Nature genetics *39*, 226-231.

Stahlman, M., Ejsing, C.S., Tarasov, K., Perman, J., Boren, J., and Ekroos, K. (2009). High-throughput shotgun lipidomics by quadrupole time-of-flight mass spectrometry. Journal of chromatography B, Analytical technologies in the biomedical and life sciences *877*, 2664-2672.

Stojiljkovic, M.P., Lopes, H.F., Zhang, D., Morrow, J.D., Goodfriend, T.L., and Egan, B.M. (2002). Increasing plasma fatty acids elevates F2-isoprostanes in humans: implications for the cardiovascular risk factor cluster. Journal of hypertension *20*, 1215-1221.

Stoop, M.P., Coulier, L., Rosenling, T., Shi, S., Smolinska, A.M., Buydens, L., Ampt, K., Stingl, C., Dane, A., Muilwijk, B.*, et al.* (2010). Quantitative proteomics and metabolomics analysis of normal human cerebrospinal fluid samples. Molecular & cellular proteomics : MCP *9*, 2063-2075.

Straface, E., Gambardella, L., Mattatelli, A., Canali, E., Boccalini, F., Agati, L., and Malorni, W. (2011). The red blood cell as a gender-associated biomarker

in metabolic syndrome: a pilot study. International journal of cell biology *2011*, 204157.

Suhre, K., and Gieger, C. (2012). Genetic variation in metabolic phenotypes: study designs and applications. Nature reviews Genetics *13*, 759-769.

Sullards, M.C. (2000). Analysis of sphingomyelin, glucosylceramide, ceramide, sphingosine, and sphingosine 1-phosphate by tandem mass spectrometry. Methods in enzymology *312*, 32-45.

Suo, C., Xu, H., Khor, C.C., Ong, R.T., Sim, X., Chen, J., Tay, W.T., Sim, K.S., Zeng, Y.X., Zhang, X.*, et al.* (2012). Natural positive selection and north-south genetic diversity in East Asia. European journal of human genetics : EJHG *20*, 102-110.

Sweeney, G., Nazir, D., Clarke, C., and Goettsche, G. (1996). Ethanolamine and choline phospholipids in nascent very-low-density lipoprotein particles. Clinical and investigative medicine Medecine clinique et experimentale *19*, 243-250.

Synder, F. (1972). Ether Lipids: Chemistry and Biology (New York: Academic Press).

Tanaka, T., Shen, J., Abecasis, G.R., Kisialiou, A., Ordovas, J.M., Guralnik, J.M., Singleton, A., Bandinelli, S., Cherubini, A., Arnett, D.*, et al.* (2009). Genome-wide association study of plasma polyunsaturated fatty acids in the InCHIANTI Study. PLoS genetics *5*, e1000338.

Teahan, O., Gamble, S., Holmes, E., Waxman, J., Nicholson, J.K., Bevan, C., and Keun, H.C. (2006). Impact of analytical bias in metabonomic studies of human blood serum and plasma. Analytical chemistry *78*, 4307-4318.

Team, R.D.C. (2008). R: A Language and Environment for Statistical Computing.

Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J.*, et al.* (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707-713.

Toufektsian, M.C., Salen, P., Laporte, F., Tonelli, C., and de Lorgeril, M. (2011). Dietary flavonoids increase plasma very long-chain (n-3) fatty acids in rats. The Journal of nutrition *141*, 37-41.

Trotzmuller, M., Guo, X., Fauland, A., Kofeler, H., and Lankmayr, E. (2011). Characteristics and origins of common chemical noise ions in negative ESI LC-MS. Journal of mass spectrometry : JMS *46*, 553-560.

Tsai, M.J., and O'Malley, B.W. (1994). Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annual review of biochemistry *63*, 451-486.

Tulipani, S., Llorach, R., Urpi-Sarda, M., and Andres-Lacueva, C. (2013). Comparative analysis of sample preparation methods to handle the complexity of the blood fluid metabolome: when less is more. Analytical chemistry *85*, 341-348.

U.S. Department of Health and Human Services, F.a.D.A., Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM) (2001). Guidance for Industry : Bioanalytical Method Validation (U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM)), pp. 1-22.

Urpi-Sarda, M., Monagas, M., Khan, N., Llorach, R., Lamuela-Raventos, R.M., Jauregui, O., Estruch, R., Izquierdo-Pulido, M., and Andres-Lacueva, C. (2009). Targeted metabolic profiling of phenolics in urine and plasma after regular consumption of cocoa by liquid chromatography-tandem mass spectrometry. Journal of chromatography A *1216*, 7258-7267.

Vance, J.E. (1990). Lipoproteins secreted by cultured rat hepatocytes contain the antioxidant 1-alk-1-enyl-2-acylglycerophosphoethanolamine. Biochimica et biophysica acta *1045*, 128-134.

Vance, J.E., and Steenbergen, R. (2005). Metabolism and functions of phosphatidylserine. Progress in lipid research *44*, 207-234.

Vance, J.E., and Tasseva, G. (2013). Formation and function of phosphatidylserine and phosphatidylethanolamine in mammalian cells. Biochimica et biophysica acta *1831*, 543-554.

Varbo, A., Nordestgaard, B.G., Tybjaerg-Hansen, A., Schnohr, P., Jensen, G.B., and Benn, M. (2011). Nonfasting triglycerides, cholesterol, and ischemic stroke in the general population. Annals of neurology *69*, 628-634.

Verkooijen, H.M., Yap, K.P., Bhalla, V., Chow, K.Y., and Chia, K.S. (2009). Multiparity and the risk of premenopausal breast cancer: different effects across ethnic groups in Singapore. Breast cancer research and treatment *113*, 553-558.

Vuckovic, D. (2012). Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. Analytical and bioanalytical chemistry *403*, 1523-1548.

Wallner, S., and Schmitz, G. (2011). Plasmalogens the neglected regulatory and scavenging lipid species. Chemistry and physics of lipids *164*, 573-589.

Wang, L., Folsom, A.R., Zheng, Z.J., Pankow, J.S., and Eckfeldt, J.H. (2003). Plasma fatty acid composition and incidence of diabetes in middle-aged adults: the Atherosclerosis Risk in Communities (ARIC) Study. The American journal of clinical nutrition *78*, 91-98.

Want, E.J., Wilson, I.D., Gika, H., Theodoridis, G., Plumb, R.S., Shockcor, J., Holmes, E., and Nicholson, J.K. (2010). Global metabolic profiling procedures for urine using UPLC-MS. Nature protocols *5*, 1005-1018.

Watson, A.D. (2006). Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Lipidomics: a global approach to lipid analysis in biological systems. Journal of lipid research *47*, 2101-2111.

Weckwerth, W. (2003). Metabolomics in systems biology. Annual review of plant biology *54*, 669-689.

Wenk, M.R. (2005). The emerging field of lipidomics. Nature reviews Drug discovery *4*, 594-610.

Wenk, M.R. (2010). Lipidomics: new tools and applications. Cell *143*, 888-895.

Wenk, M.R., Lucast, L., Di Paolo, G., Romanelli, A.J., Suchy, S.F., Nussbaum, R.L., Cline, G.W., Shulman, G.I., McMurray, W., and De Camilli, P. (2003). Phosphoinositide profiling in complex lipid mixtures using electrospray ionization mass spectrometry. Nat Biotechnol *21*, 813-817.

Wheeler, D., Bandaru, V.V., Calabresi, P.A., Nath, A., and Haughey, N.J. (2008). A defect of sphingolipid metabolism modifies the properties of normal appearing white matter in multiple sclerosis. Brain : a journal of neurology *131*, 3092-3102.

Whitney, A.R., Diehn, M., Popper, S.J., Alizadeh, A.A., Boldrick, J.C., Relman, D.A., and Brown, P.O. (2003). Individuality and variation in gene expression patterns in human blood. Proceedings of the National Academy of Sciences of the United States of America *100*, 1896-1901.

Wiesner, P., Leidl, K., Boettcher, A., Schmitz, G., and Liebisch, G. (2009). Lipid profiling of FPLC-separated lipoprotein fractions by electrospray ionization tandem mass spectrometry. Journal of lipid research *50*, 574-585.

Wiest, M.M., and Watkins, S.M. (2007). Biomarker discovery using high-dimensional lipid analysis. Current opinion in lipidology *18*, 181-186.

Winnike, J.H., Busby, M.G., Watkins, P.B., and O'Connell, T.M. (2009). Effects of a prolonged standardized diet on normalizing the human metabolome. The American journal of clinical nutrition *90*, 1496-1501.

Wong, L.P., Ong, R.T., Poh, W.T., Liu, X., Chen, P., Li, R., Lam, K.K., Pillai, N.E., Sim, K.S., Xu, H.*, et al.* (2012). Deep Whole-Genome Sequencing of 100 Southeast Asian Malays. American journal of human genetics.

Wymann, M.P., and Schneiter, R. (2008). Lipid signalling in disease. Nature reviews Molecular cell biology *9*, 162-176.

Xu, Z., Byers, D.M., Palmer, F.B., and Cook, H.W. (1994). Serine and ethanolamine incorporation into different plasmalogen pools: subcellular analyses of phosphoglyceride synthesis in cultured glioma cells. Neurochemical research *19*, 769-775.

Xu, Z.L., Byers, D.M., Palmer, F.B., Spence, M.W., and Cook, H.W. (1991). Serine utilization as a precursor of phosphatidylserine and alkenyl-(plasmenyl)-, alkyl-, and acylethanolamine phosphoglycerides in cultured glioma cells. The Journal of biological chemistry *266*, 2143-2150.

Yorek, M.A. (1993). Phospholipid Handbook (Marcel Dekker, Inc).

Yost, R.A., and Enke, C.G. (1978). Selected Ion Fragmentation with a Tandem Quadrupole Mass-Spectrometer. Journal of the American Chemical Society *100*, 2274-2275.

Young, S.P., and Wallace, G.R. (2009). Metabolomic analysis of human disease and its application to the eye. Journal of ocular biology, diseases, and informatics *2*, 235-242.

Yu, Z., Zhai, G., Singmann, P., He, Y., Xu, T., Prehn, C., Romisch-Margl, W., Lattka, E., Gieger, C., Soranzo, N.*, et al.* (2012). Human serum metabolic profiles are age dependent. Aging cell *11*, 960-967.

Zelena, E., Dunn, W.B., Broadhurst, D., Francis-McIntyre, S., Carroll, K.M., Begley, P., O'Hagan, S., Knowles, J.D., Halsall, A., Wilson, I.D.*, et al.* (2009).

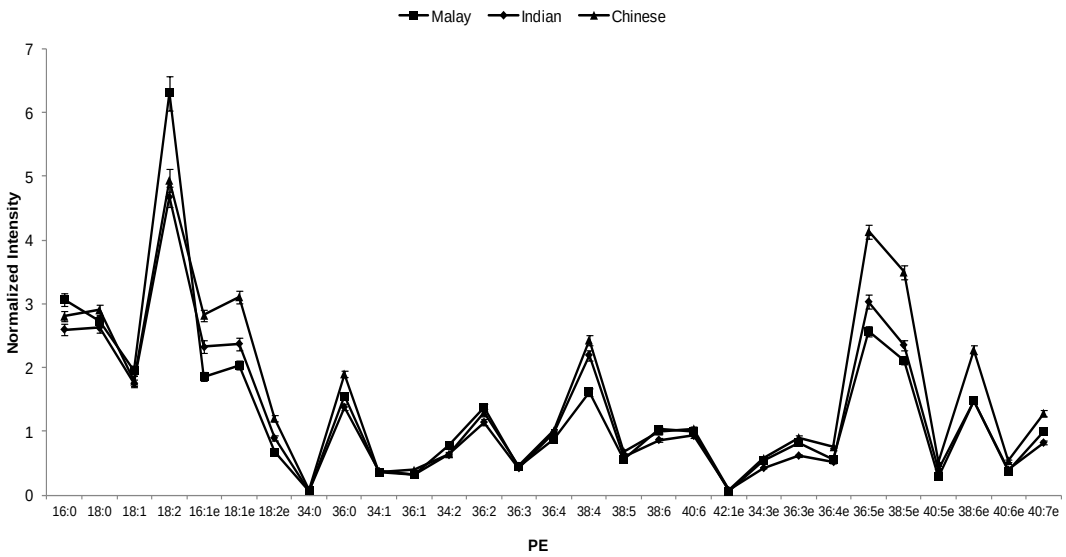
Development of a robust and repeatable UPLC-MS method for the long-term metabolomic study of human serum. Analytical chemistry *81*, 1357-1364.

Zhang, X., Patel, A., Horibe, H., Wu, Z., Barzi, F., Rodgers, A., MacMahon, S., and Woodward, M. (2003). Cholesterol, coronary heart disease, and stroke in the Asia Pacific region. International journal of epidemiology *32*, 563-572.

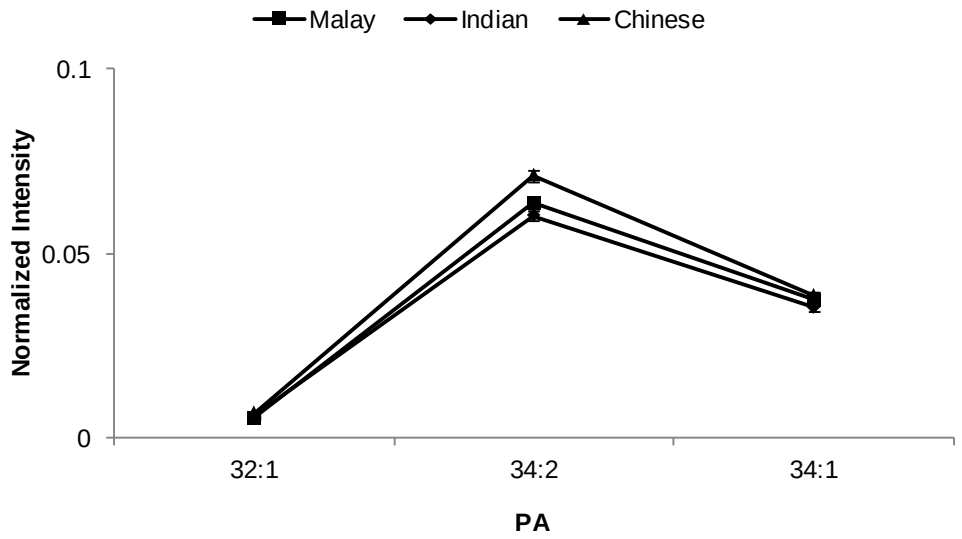
Zoeller, R.A., Morand, O.H., and Raetz, C.R. (1988). A possible role for plasmalogens in protecting animal cells against photosensitized killing. The Journal of biological chemistry *263*, 11590-11596.

APPENDIX

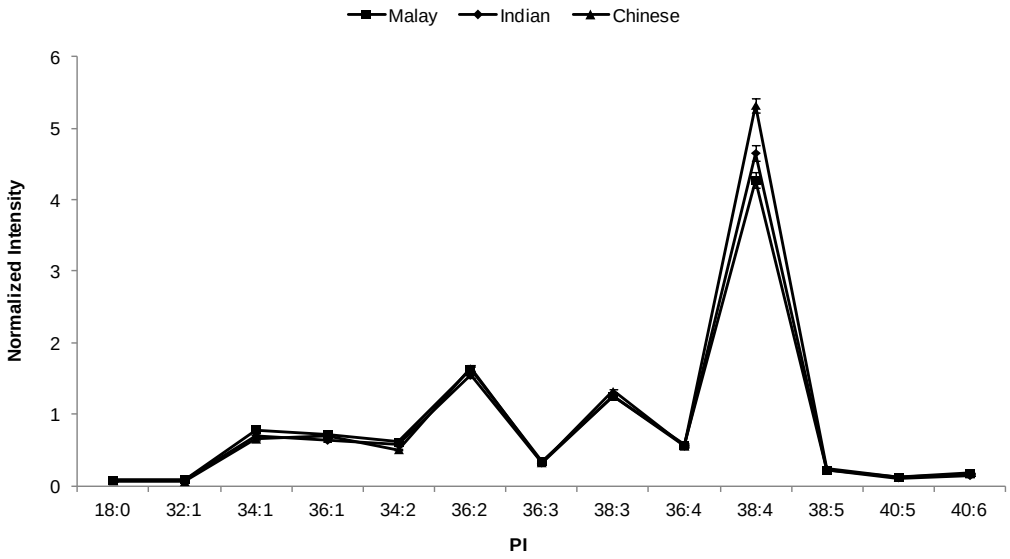
(a)


Malay
Indian
Chinese
Normalized Intensity
0
1
2
3
4
5
6
7
16:0 18:0 18:1 18:2 16:1e 18:1e 18:2e 34:0 36:0 34:1 36:1 34:2 36:2 36:3 36:4 38:4 38:5 38:6 40:6 42:1e 34:3e 36:3e 36:4e 36:5e 38:5e 40:5e 38:6e 40:6e 40:7e
PE

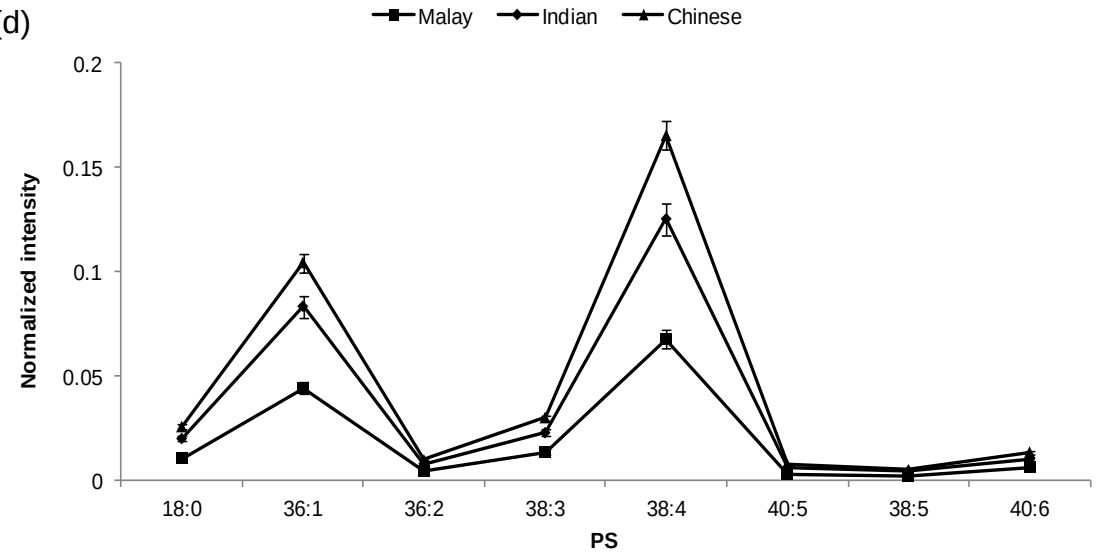
(b)


Malay
Indian
Chinese
Normalized Intensity
0
0.05
0.1
32:1
34:2
34:1
PA

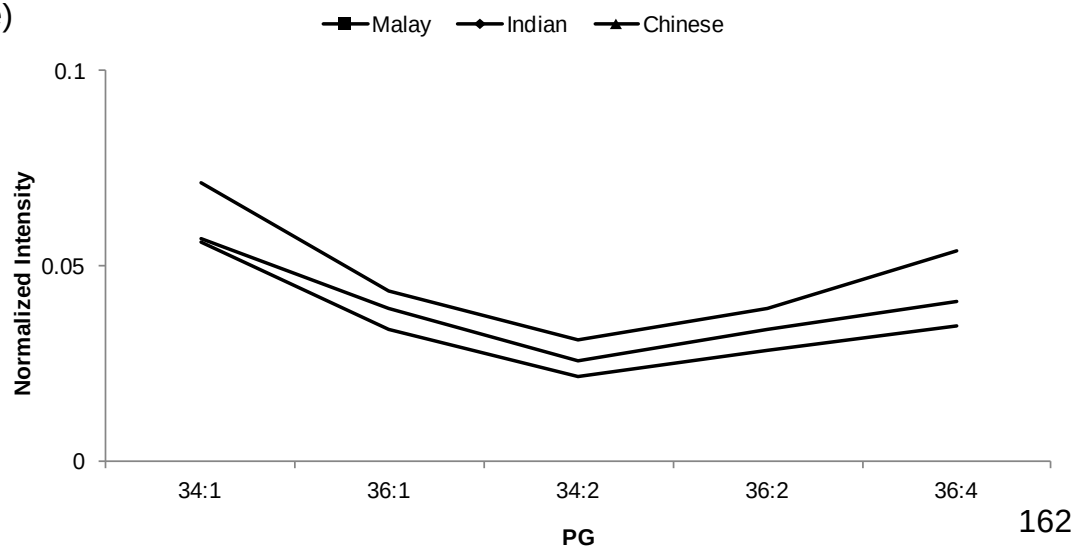
(c)


Malay
Indian
Chinese
Normalized Intensity
0
1
2
3
4
5
6
18:0
32:1
34:1
36:1
34:2
36:2
36:3
38:3
36:4
38:4
38:5
40:5
40:6
PI

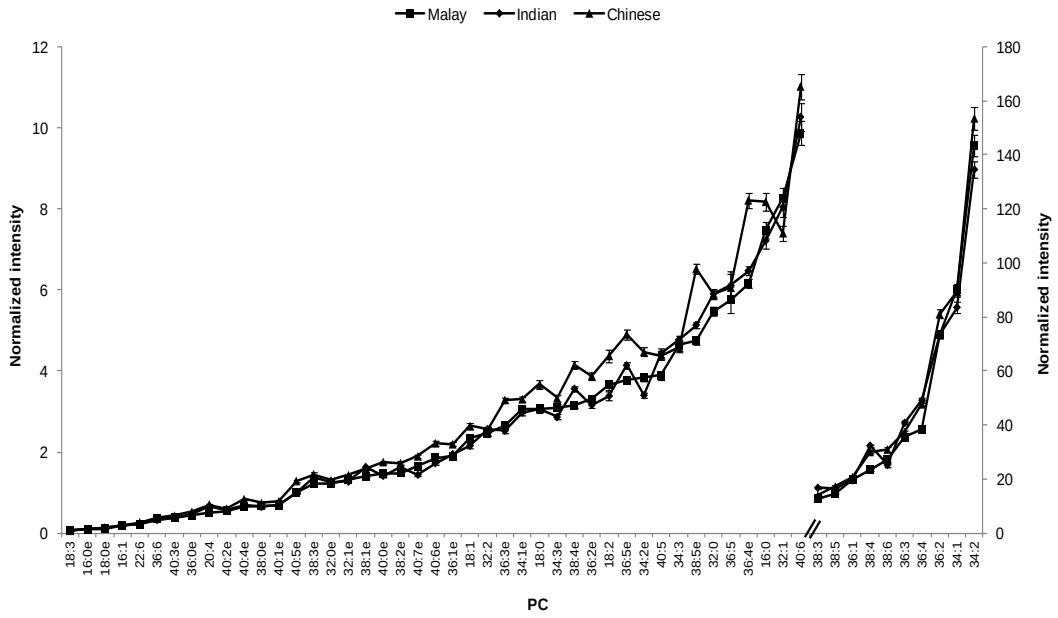
(d)


Malay
Indian
Chinese
Normalized intensity
0
0.05
0.1
0.15
0.2
18:0
36:1
36:2
38:3
38:4
40:5
38:5
40:6
PS

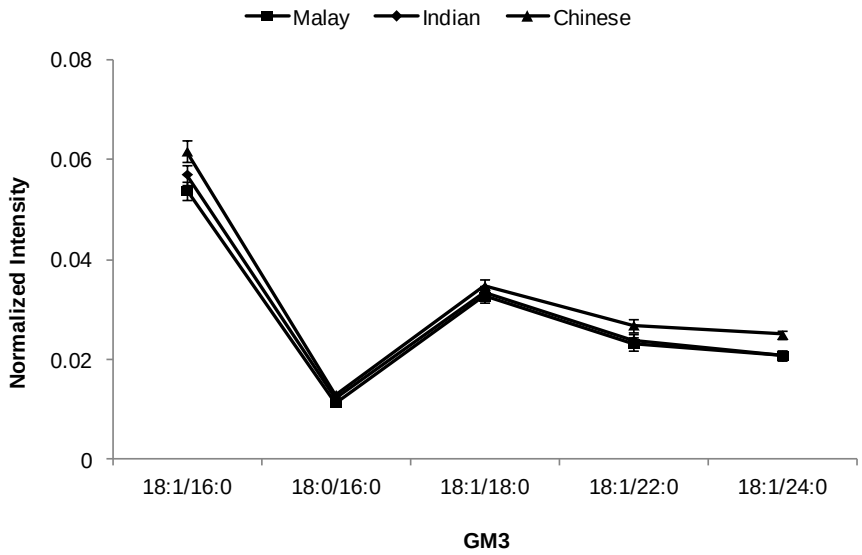
(e)


Malay
Indian
Chinese
Normalized Intensity
0
0.05
0.1
34:1
36:1
34:2
36:2
36:4
PG

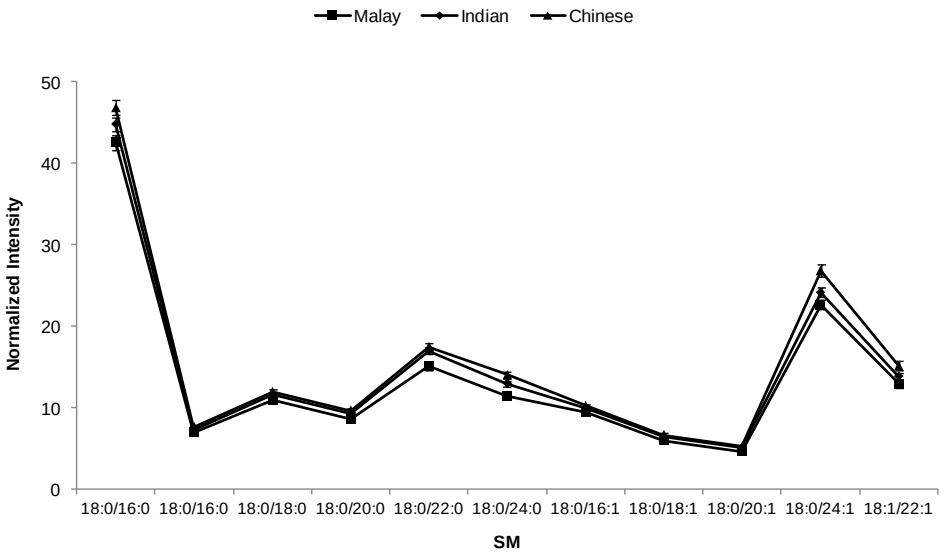
(f)


Malay
Indian
Chinese
Normalized intensity
Normalized intensity
PC
18:3 16:0e 18:0e 16:1 22:6 36:6 40:3e 36:0e 20:4 40:2e 40:4e 38:0e 40:1e 40:5e 38:3e 32:0e 32:1e 38:1e 40:0e 38:2e 40:7e 40:6e 36:1e 18:1 32:2 36:3e 34:1e 18:0 34:3e 38:4e 36:2e 18:2 36:5e 34:2e 40:5 34:3 38:5e 32:0 36:5 36:4e 16:0 32:1 40:6 38:3 38:5 36:1 38:4 38:6 36:3 36:4 36:2 34:1 34:2

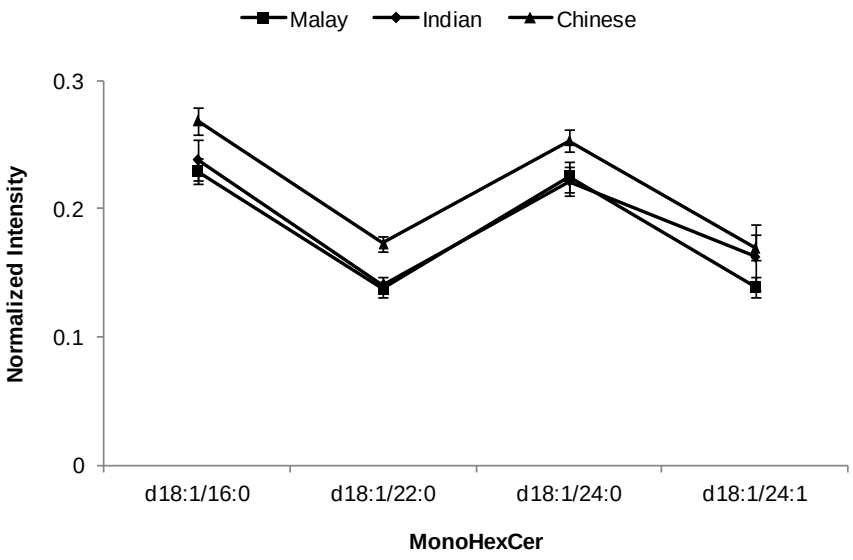
(g)


Malay
Indian
Chinese
Normalized Intensity
18:1/16:0
18:0/16:0
18:1/18:0
18:1/22:0
18:1/24:0
GM3

(h)


Malay
Indian
Chinese
Normalized Intensity
18:0/16:0 18:0/16:0 18:0/18:0 18:0/20:0 18:0/22:0 18:0/24:0 18:0/16:1 18:0/18:1 18:0/20:1 18:0/24:1 18:1/22:1
SM

(i)



(j)

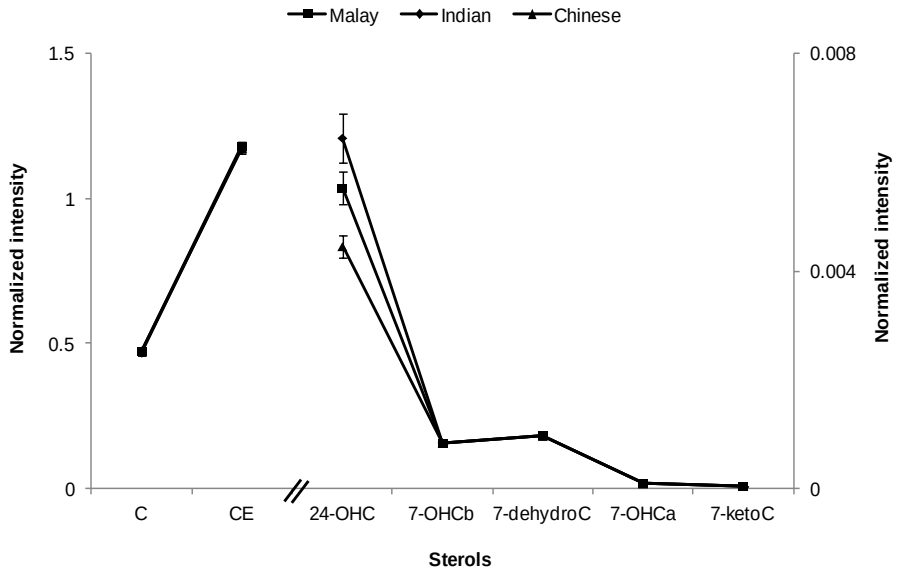
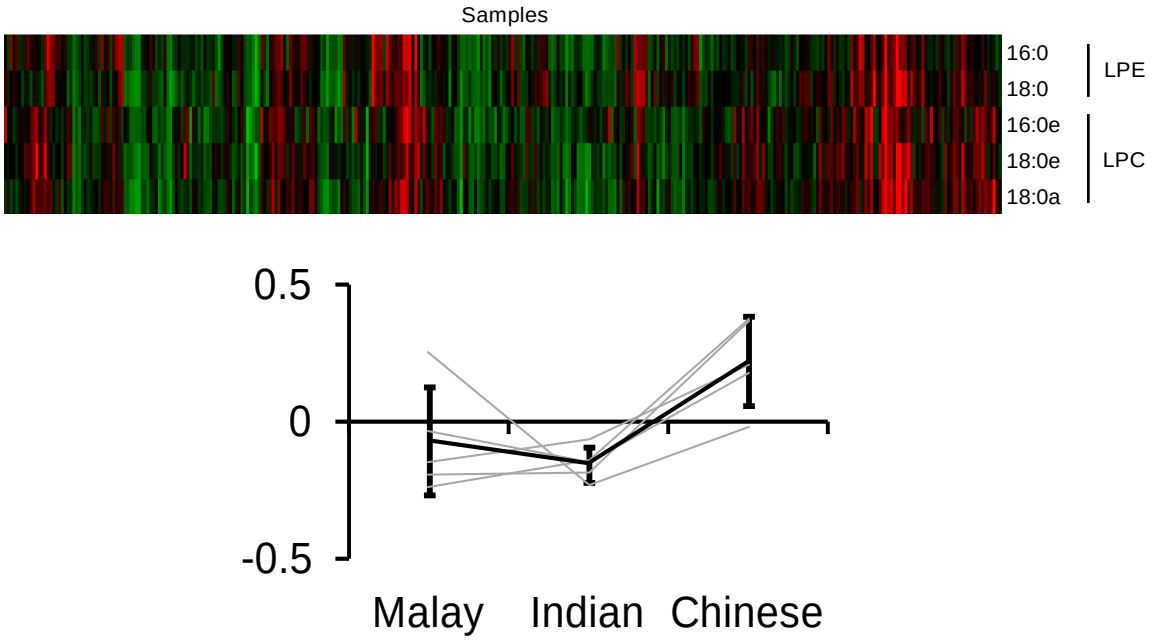
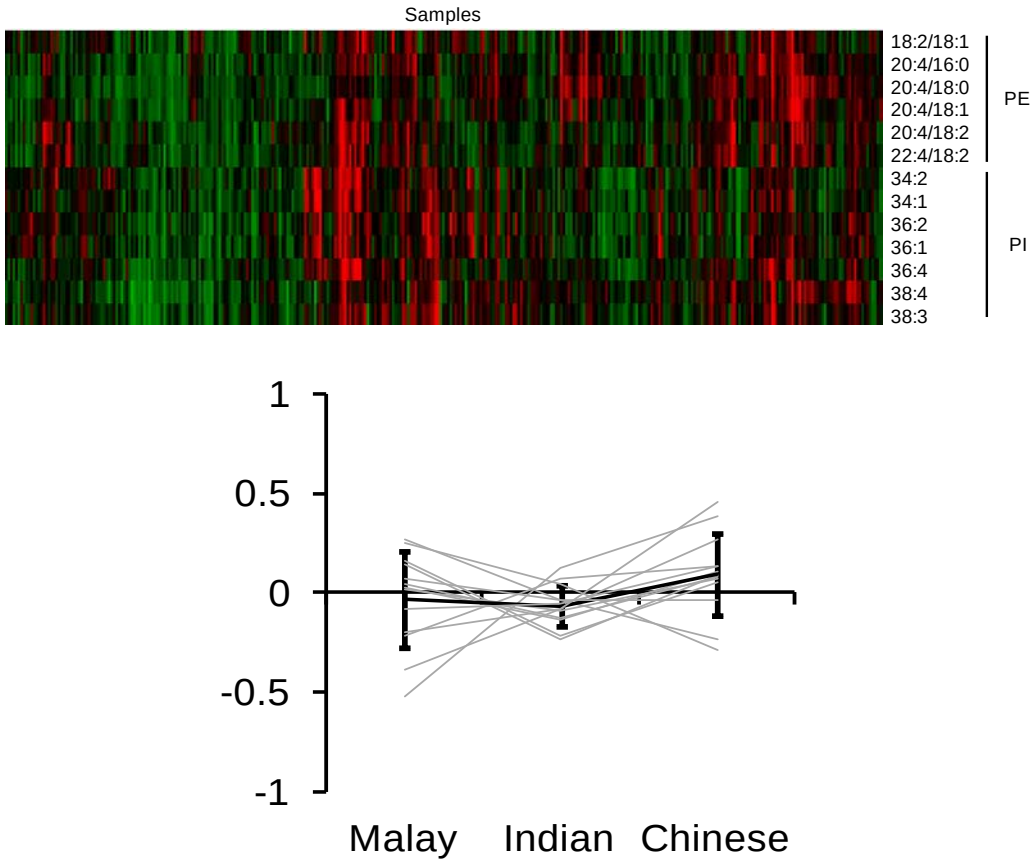


Figure 6-1: Line plots showing lipid profile trends across Malay, Indian and Chinese in the Singaporean population for GPLs (a) PE(b) PA (c) PI (d) PS (e) PG (f) PC (g) GM3 (h) SM (i) GluCer (j) STLs. Break in axis indicates secondary y axis. Error bars are plotted as standard error mean (SEM).

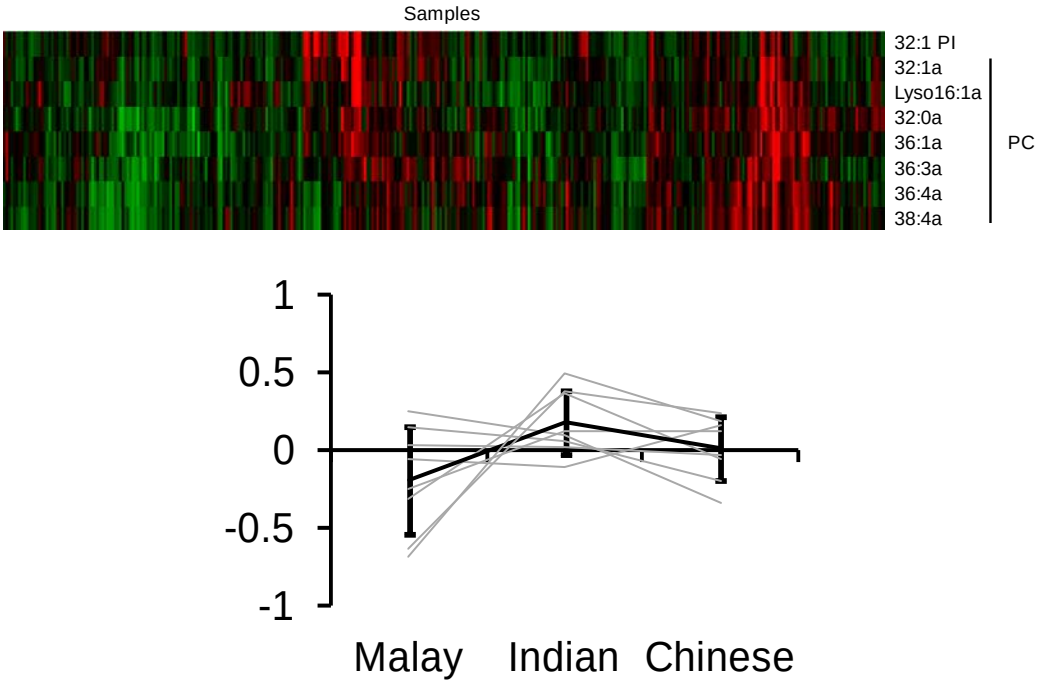
(a)


Samples
16:0
18:0
16:0e
18:0e
18:0a
LPE
LPC
0.5
0
-0.5
Malay
Indian
Chinese

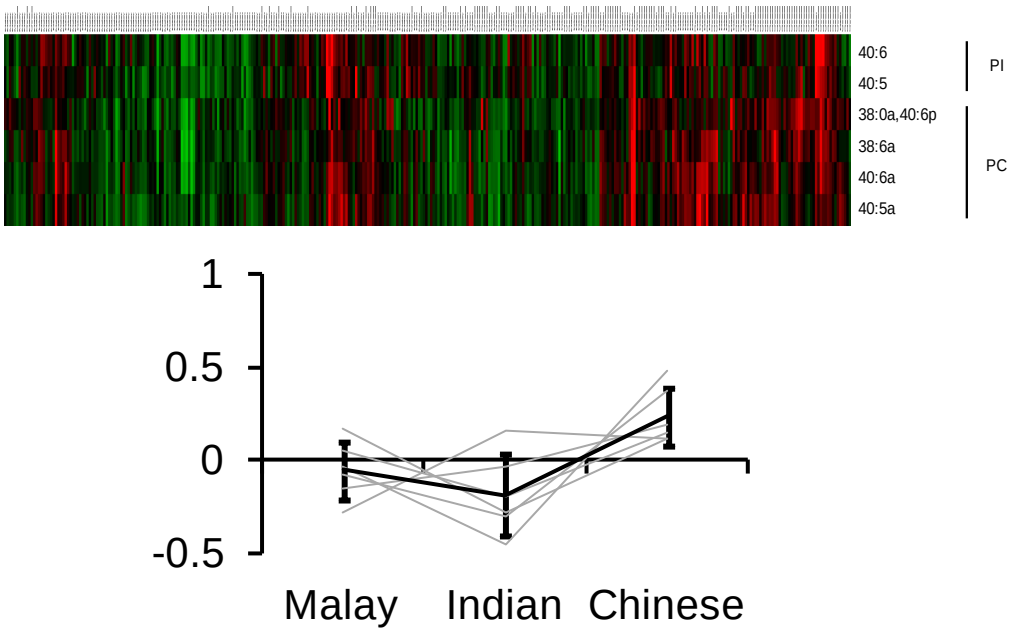
(b)


Samples
18:2/18:1
20:4/16:0
20:4/18:0
20:4/18:1
20:4/18:2
22:4/18:2
34:2
34:1
36:2
36:1
36:4
38:4
38:3
PE
PI
1
0.5
0
-0.5
-1
Malay
Indian
Chinese

(c)


Samples
32:1 PI
32:1a
Lyso16:1a
32:0a
36:1a
36:3a
36:4a
38:4a
PC
1
0.5
0
-0.5
-1
Malay
Indian
Chinese

(d)


40:6
40:5
38:0a,40:6p
38:6a
40:6a
40:5a
PI
PC
1
0.5
0
-0.5
Malay
Indian
Chinese

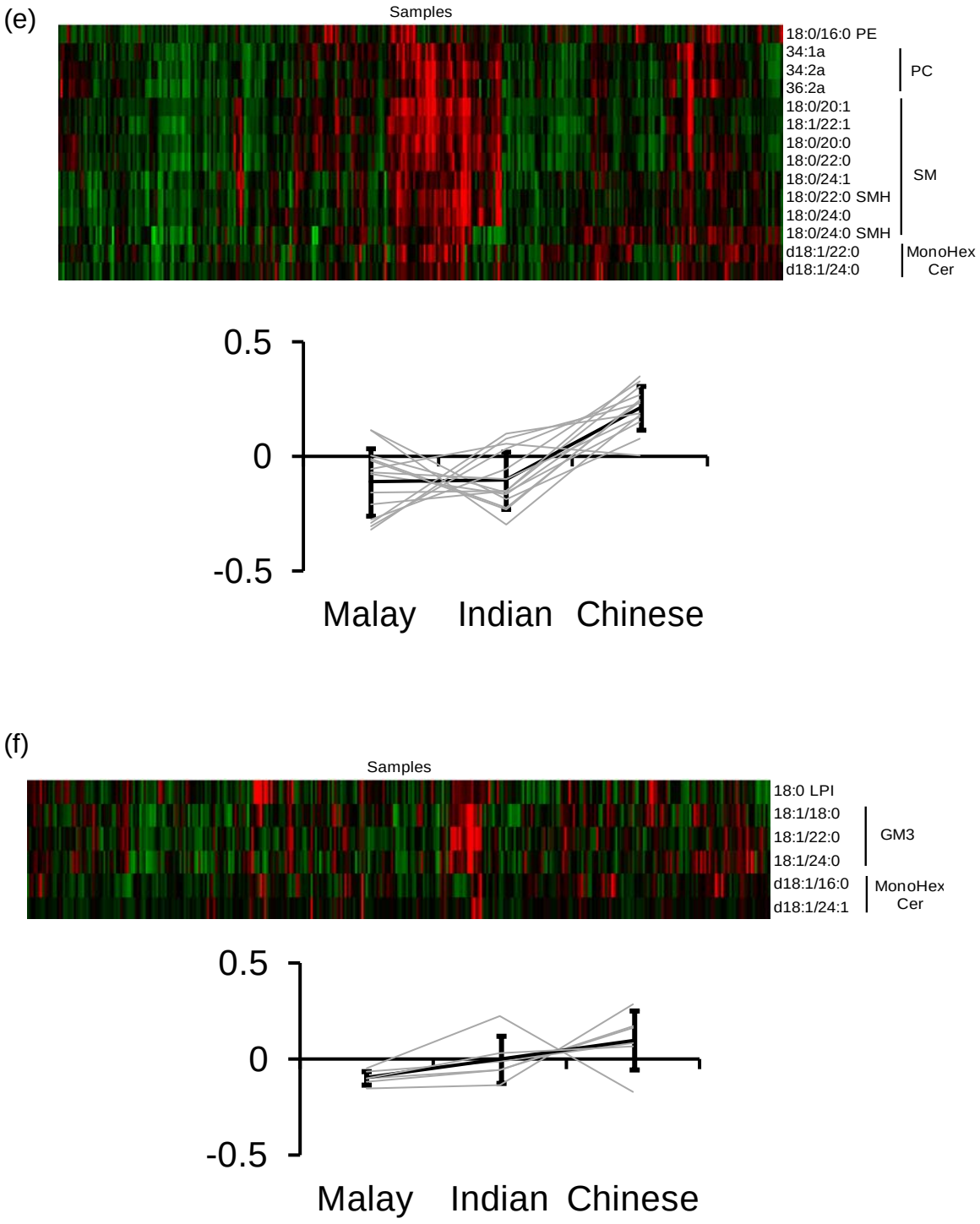
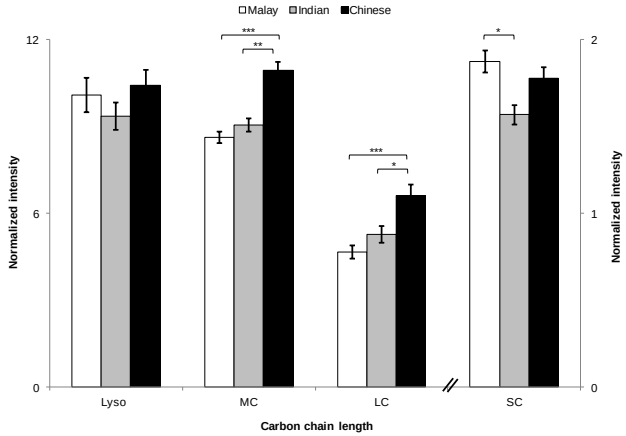
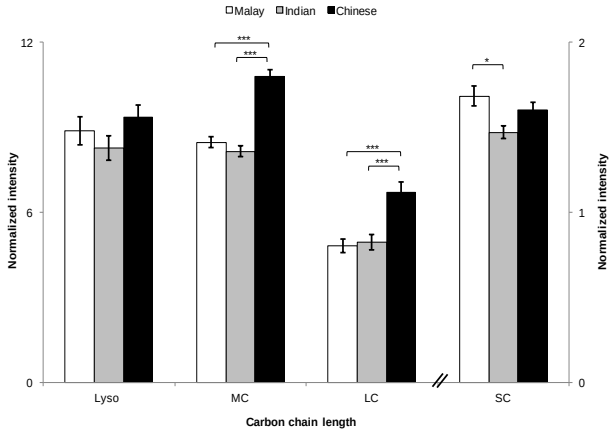


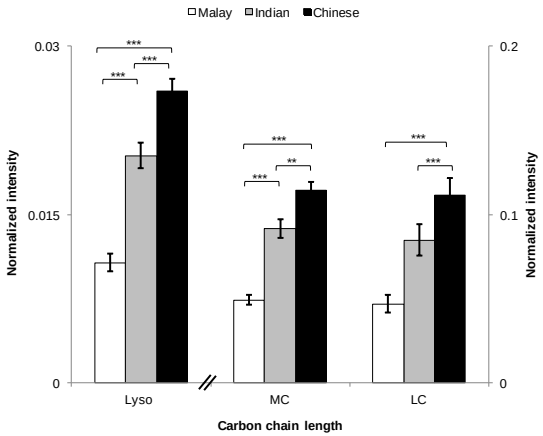
Figure 6-2: Hierarchical clustering of lipid species showing heat map and cluster line plots for (a) Cluster 4 (b) Cluster 5 (c) Cluster 6 (d) Cluster 7 (e) Cluster 8 (f) Cluster 9. Heat map colours represent relative values, with green representing values below the mean, red representing values above the mean and black representing the mean of a row (lipid) across all columns (samples). The mean of each cluster line plot is represented by a bold line. Error bars represent standard deviation.

(a)

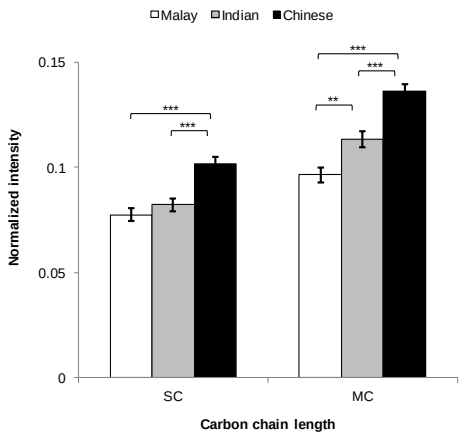

Malay Indian Chinese
Normalized intensity
Normalized intensity
Lyso
MC
LC
SC
Carbon chain length


Malay Indian Chinese
Normalized intensity
Normalized intensity
Lyso
MC
LC
SC
Carbon chain length

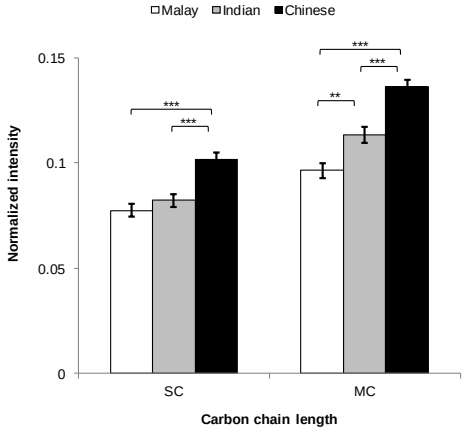
(b)


Malay Indian Chinese
Normalized intensity
Normalized intensity
Lyso
MC
LC
Carbon chain length

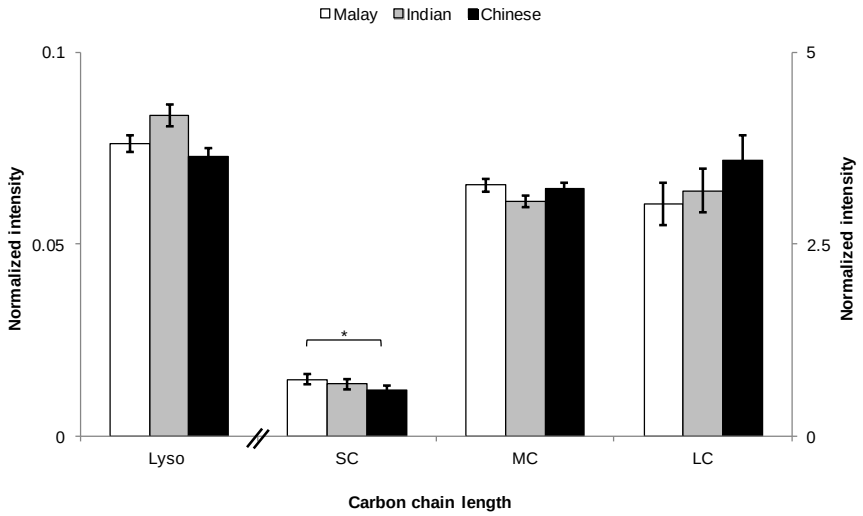
(c)


Malay Indian Chinese
Normalized intensity
SC
MC
Carbon chain length

(d)


Malay Indian Chinese
Normalized intensity
SC
MC
Carbon chain length

(e)



(f)

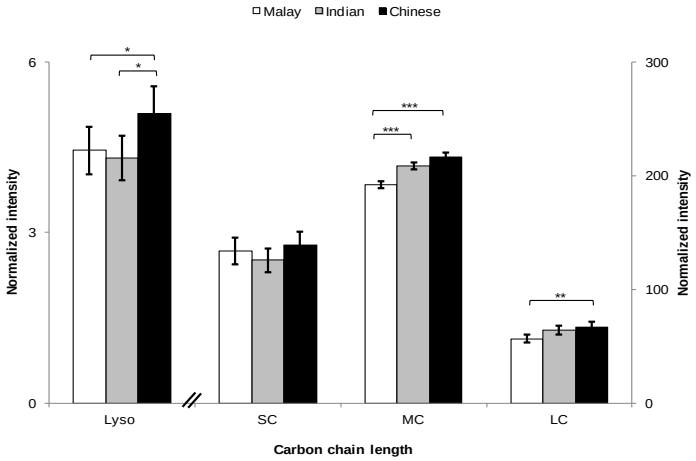


Figure 6-3: Ethnic related trends in carbon chain length across Malay, Indian Singaporean population for GPLs (a) PE, left panel for male and right panel for female, (b) PS (c) PG (d) PA (e) PI (f) PC. Break in axis indicates normalized lipid intensity on right secondary y-axis. Error bars are plotted as standard error mean (SEM).).* is p <0.05, ** is p<0.01 and *** is p<0.001.

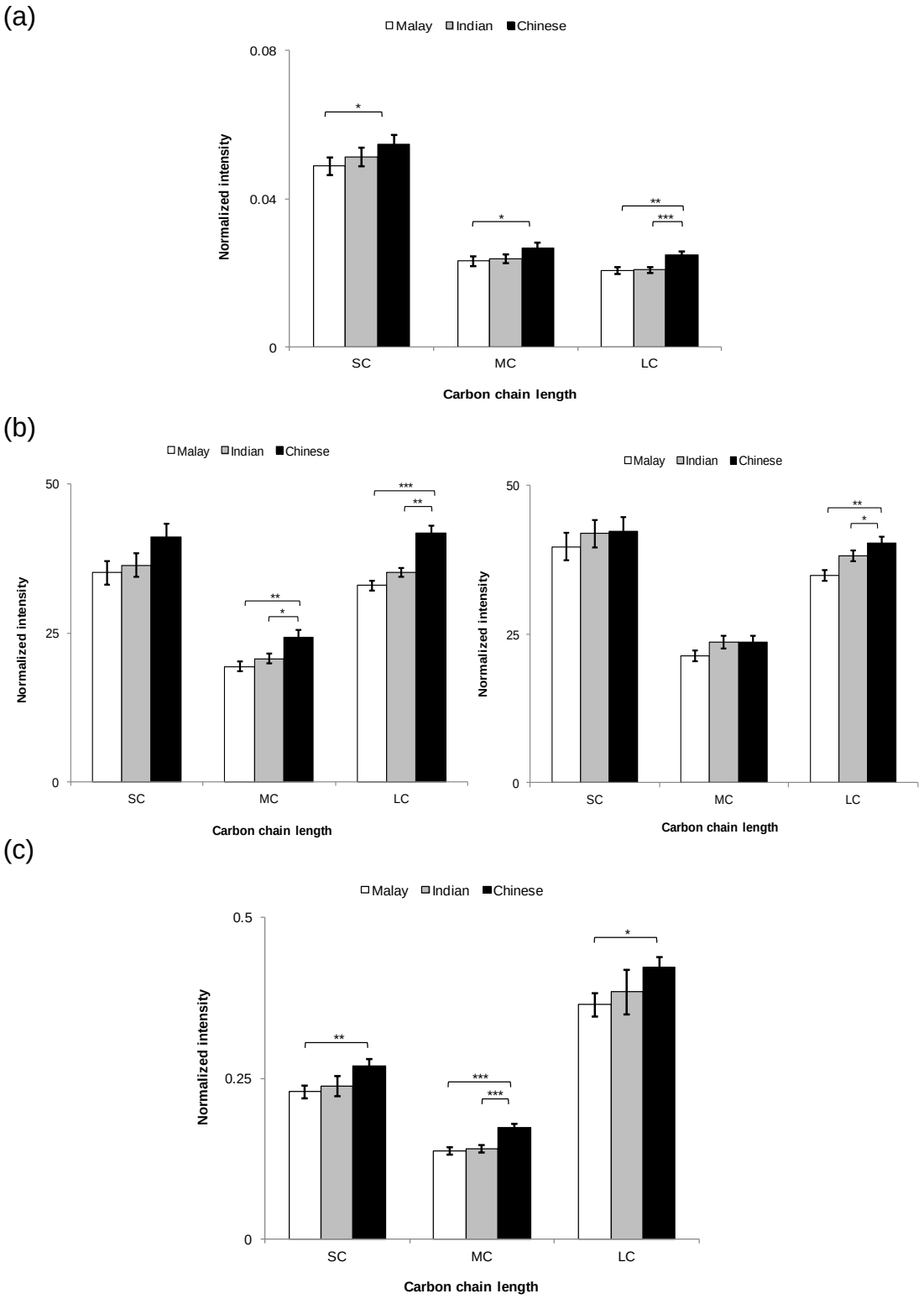
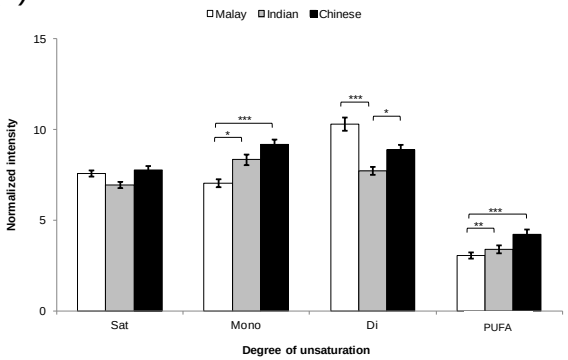
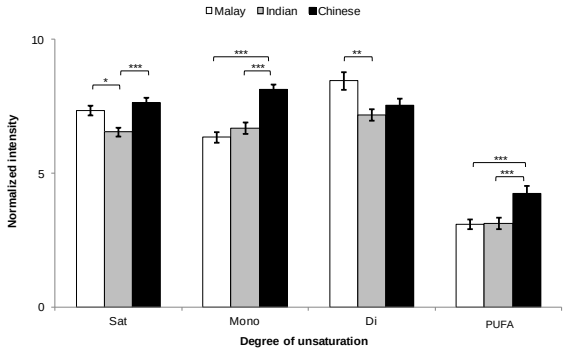
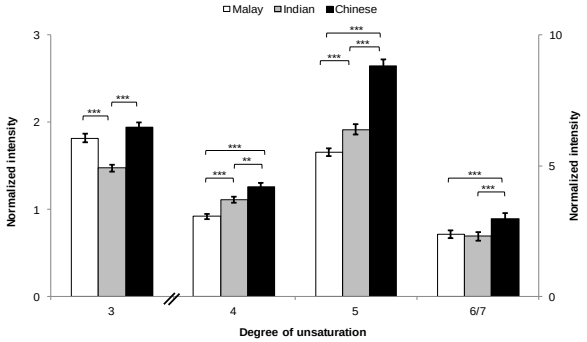


Figure 6-4: Ethnic related trends in carbon chain length (attached FA) across Malay, Indian and Chinese in the Singaporean population for SPs (a) GM3 (b) SM, left panel for male and right panel for female (c) GluCer. Error bars are plotted as standard error mean (SEM).* is p <0.05, ** is p<0.01 and *** is p<0.001.

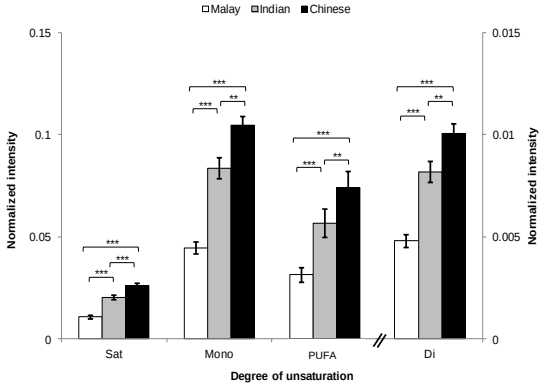
(a)

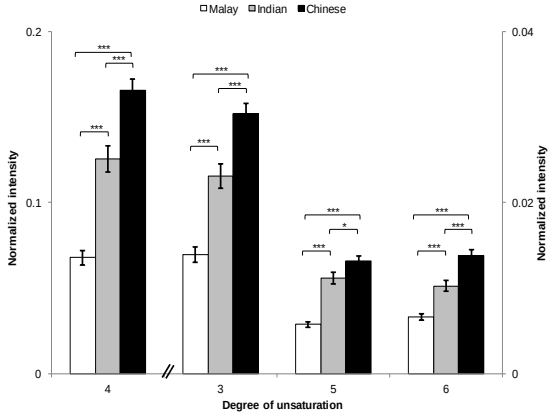

Malay Indian Chinese
Normalized intensity
Sat
Mono
Di
PUFA
Degree of unsaturation


Malay Indian Chinese
Normalized intensity
Sat
Mono
Di
PUFA
Degree of unsaturation

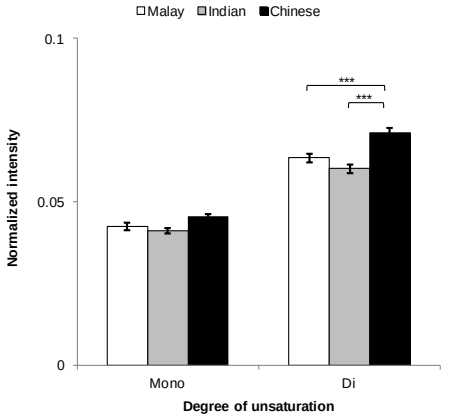

Malay Indian Chinese
Normalized intensity
3
4
5
6/7
Degree of unsaturation

(b)

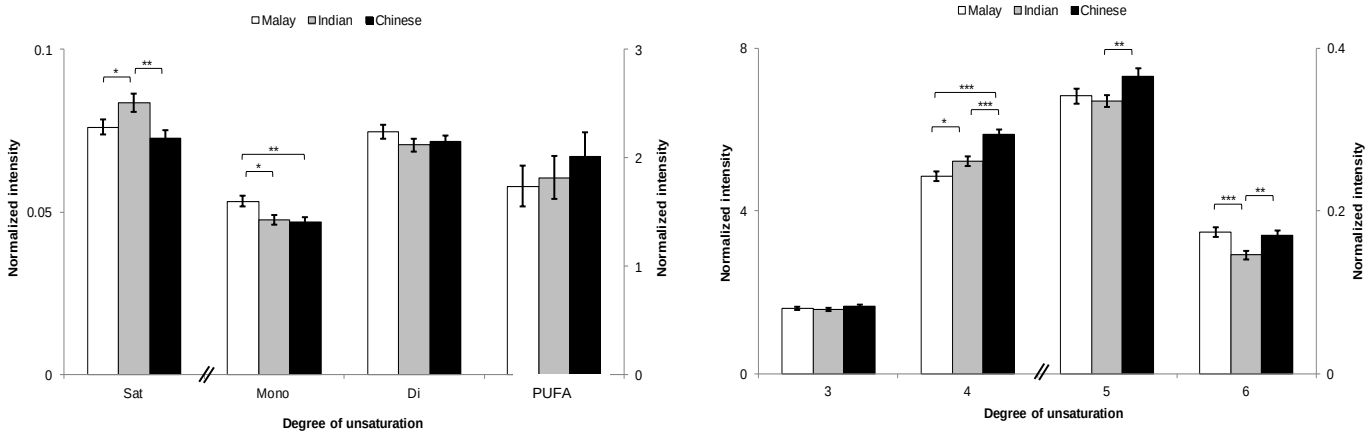

Malay Indian Chinese
Normalized intensity
Sat
Mono
PUFA
Di
Degree of unsaturation


Malay Indian Chinese
Normalized intensity
4
3
5
6
Degree of unsaturation

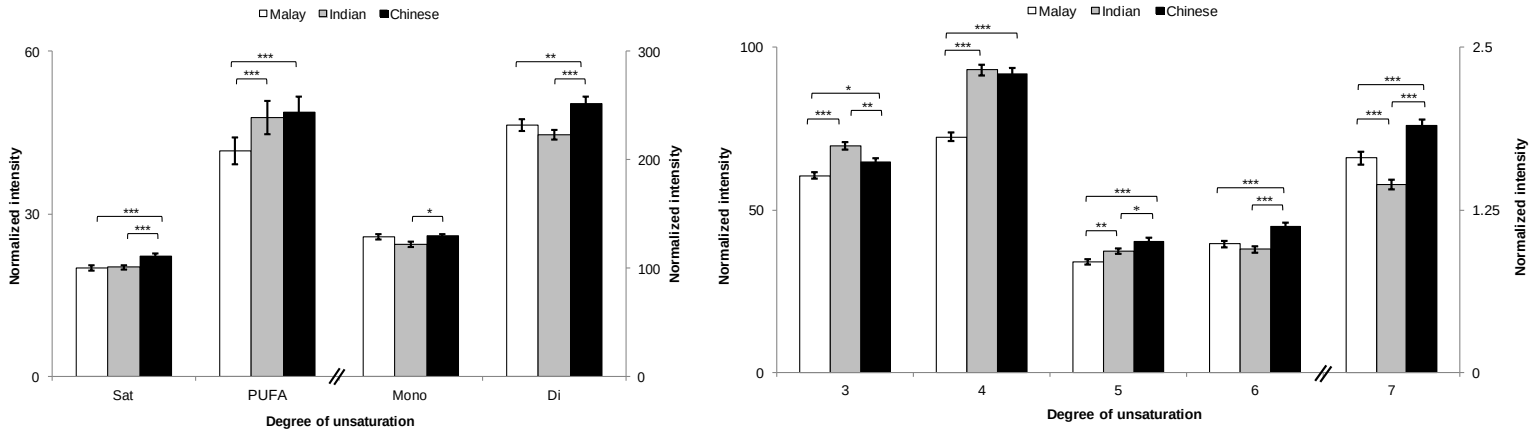
(c)


Malay Indian Chinese
Normalized intensity
Mono
Di
Degree of unsaturation

(d)


Malay Indian Chinese
Normalized intensity
Sat
Mono
Di
PUFA
Degree of unsaturation
3
4
5
6

(e)


Malay Indian Chinese
Normalized intensity
Sat
PUFA
Mono
Di
Degree of unsaturation
3
4
5
6
7

(f)

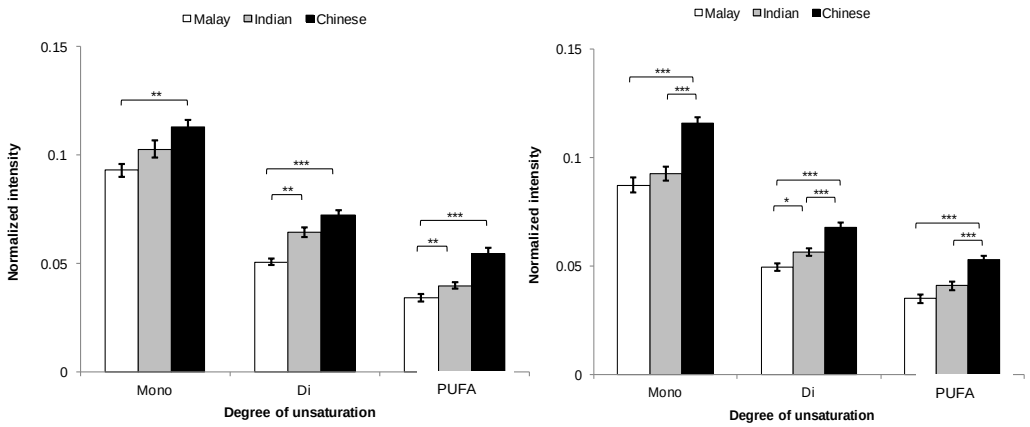
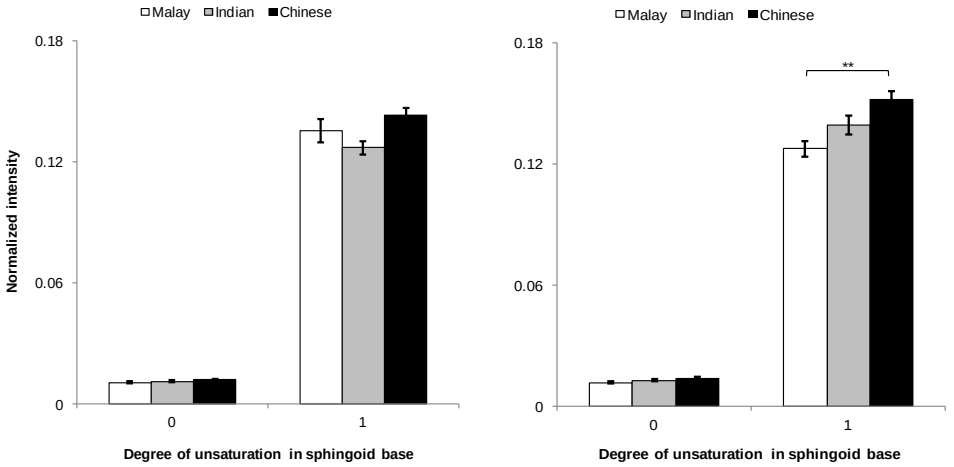
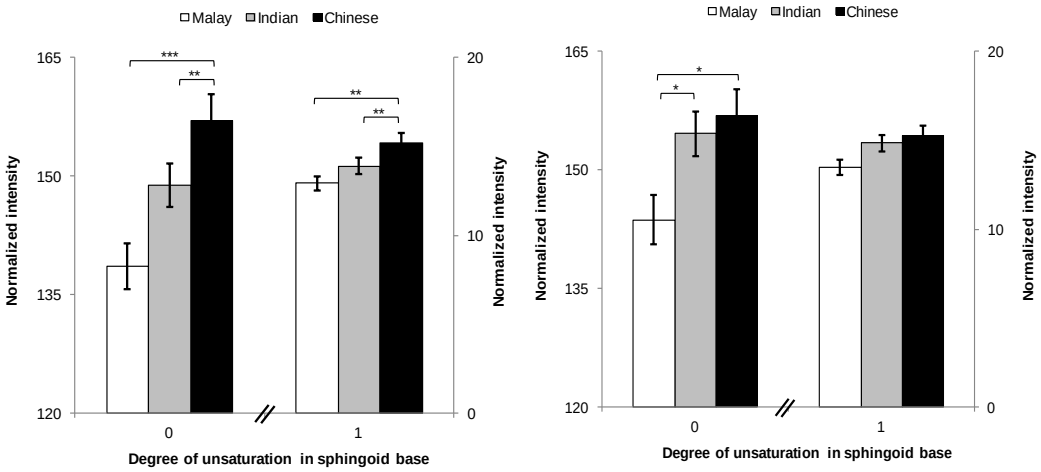


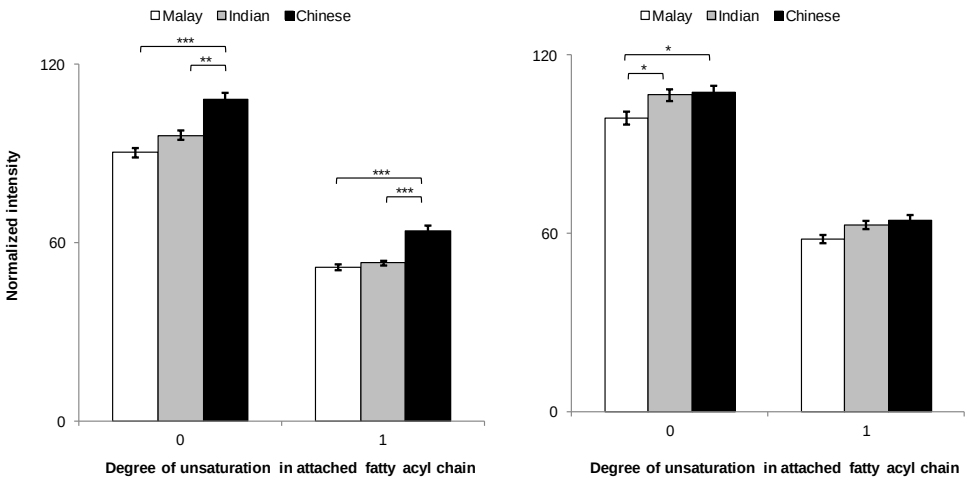
Figure 6-5: Ethnic related trends for degree of unsaturation across Malay, Indian and Chinese in the Singaporean population for GPLs (a) PE,left panel for male and center panel for female (b) PS (c) PA (d) PI (e) PC (f) PG, left panel for male and right panel for female. PUFA lipids are further represented based on their individual number of double bonds (right panel). Break in axis indicates normalized lipid intensity on right secondary y-axis Error bars are plotted as standard error mean (SEM). * is p <0.05, ** is p<0.01 and *** is p<0.001.

(a)


Malay Indian Chinese
0.18
0.12
0.06
0
Normalized intensity
0
1
Degree of unsaturation in sphingoid base
**

(b)


Malay Indian Chinese
165
150
135
120
20
10
0
Normalized intensity
***
**
*
0
1
Degree of unsaturation in sphingoid base


Malay Indian Chinese
120
60
0
Normalized intensity
***
**
*
0
1
Degree of unsaturation in attached fatty acyl chain

(c)

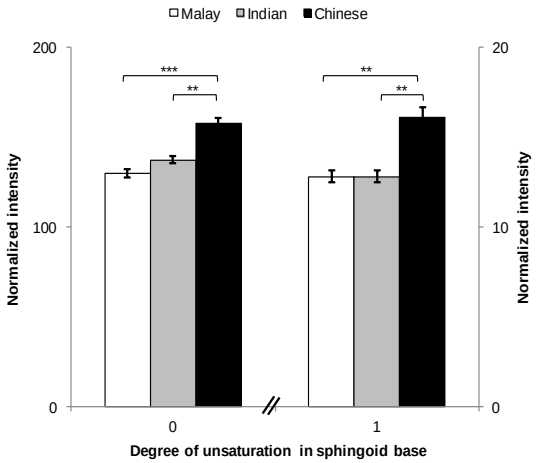
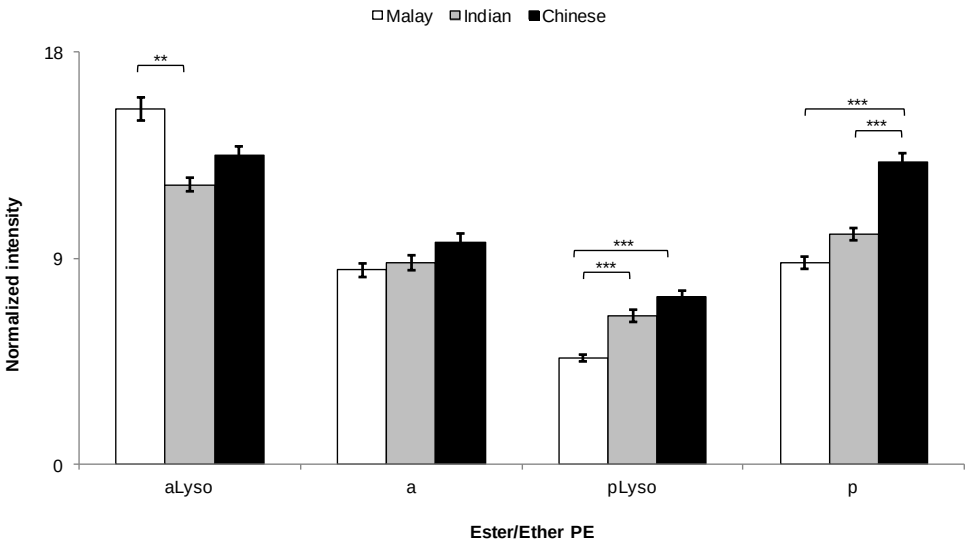
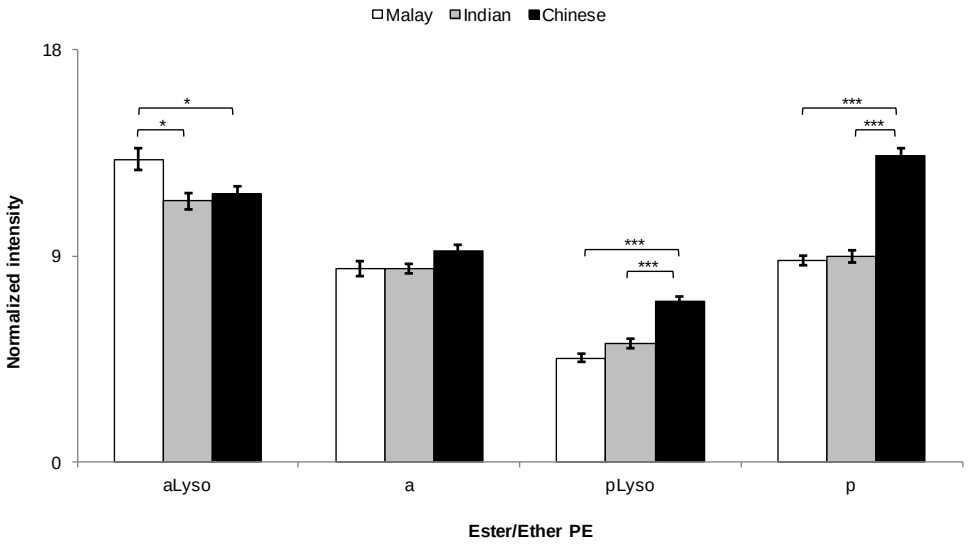


Figure 6-6: Ethnic related trends for degree of unsaturation across Malay, Indian and Chinese in the Singaporean population for SPs (a) GM3 for sphingoid base (b) SM for sphingoid base and attached FA chain (c) GluCer for attached FA chain. Left panel for male and right panel for female. Error bars are plotted as standard error mean (SEM). * is p <0.05, ** is p<0.01 and *** is p<0.001.

(a)


Malay
Indian
Chinese
18
9
0
Normalized intensity
**
***
***
***
***
aLyso
a
pLyso
p
Ester/Ether PE


Malay
Indian
Chinese
18
9
0
Normalized intensity
*
*
***
***
***
***
aLyso
a
pLyso
p
Ester/Ether PE

(b)

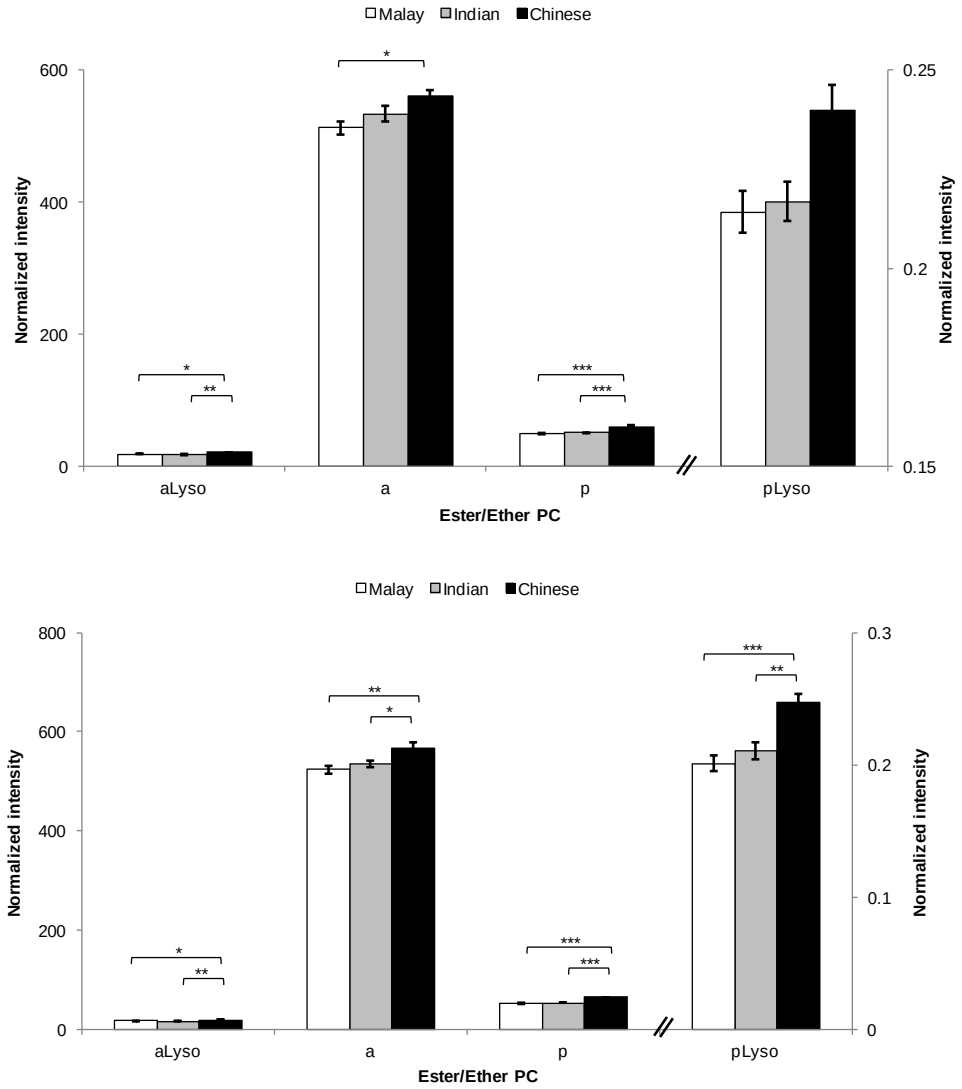


Figure 6-7: Ethnic related trends for a and p lipid species across Malay, Indian and Chinese in the Singaporean population for GPLs (a) PE (b) PC. Top panel is for male and bottom panel is for female. *Abbreviations:* a denotes ester and p denotes plasmalogenspecies. Error bars are plotted as standard error mean (SEM). * is p <0.05, ** is p<0.01 and *** is p<0.001.

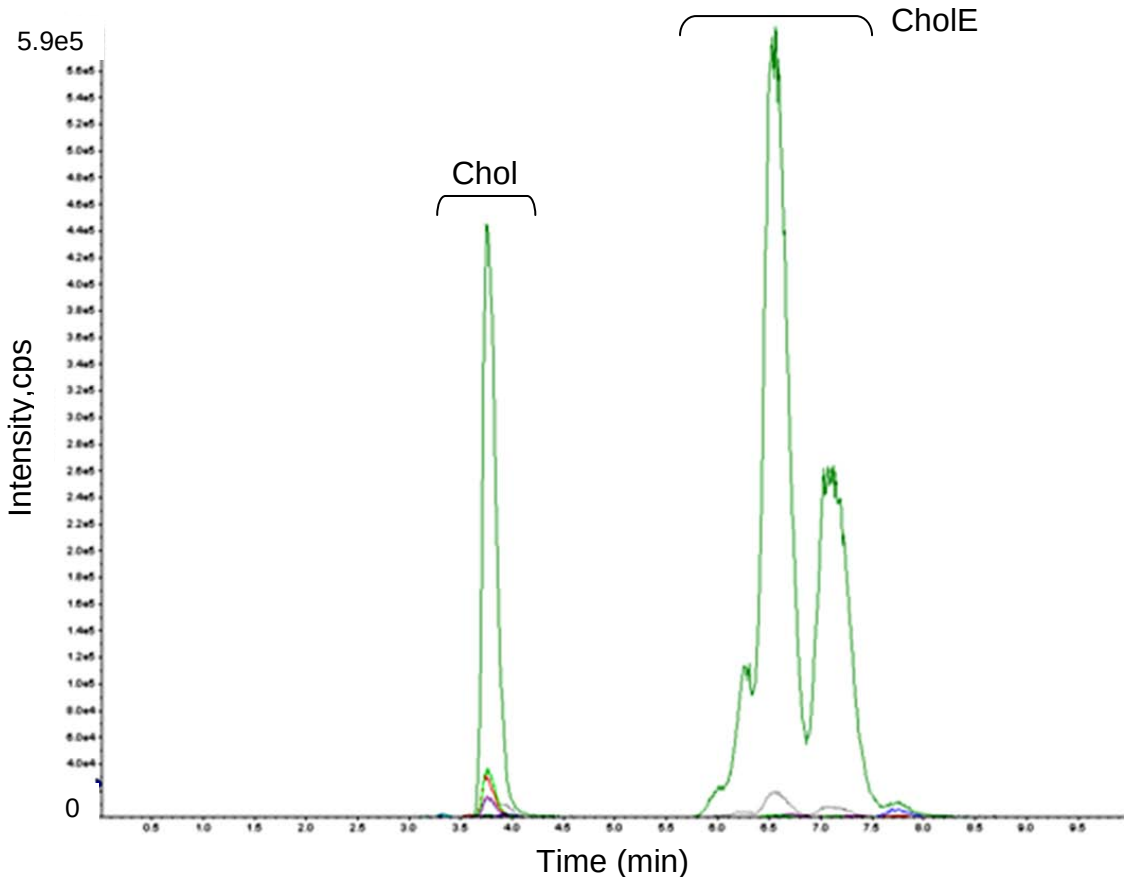
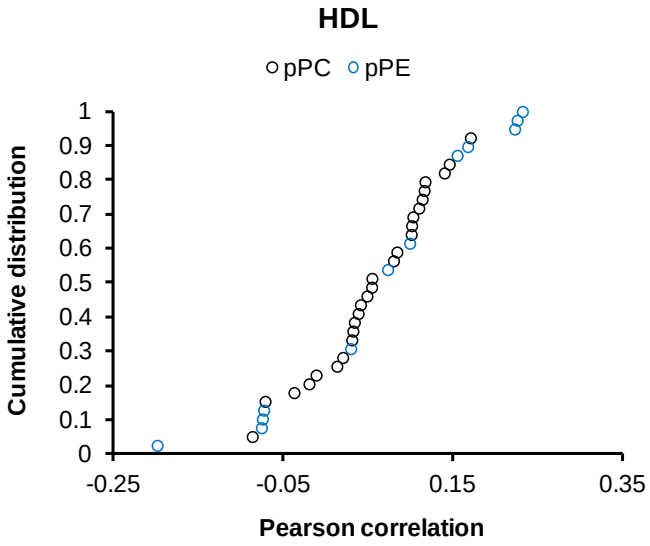
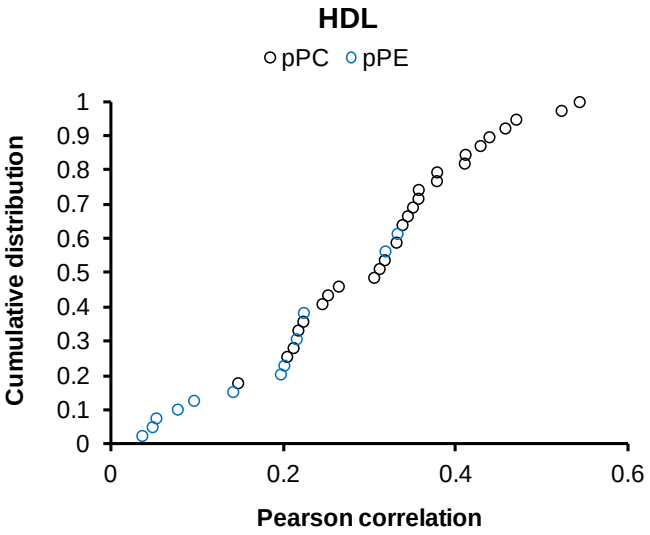


Figure 6-8: Sample reconstituted chromatogram from sterol analysis showing chromatographic separation between the different sterols. Cholesterol (Chol) and all other sterol derivatives elute between 2.7 to 4.1 min, cholesterol esters (CholE) elute between 5.5 to 9 min.

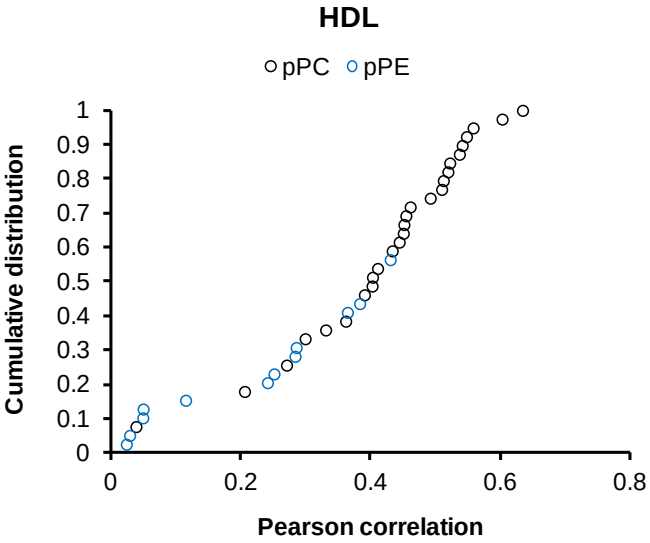
(a)



Malay

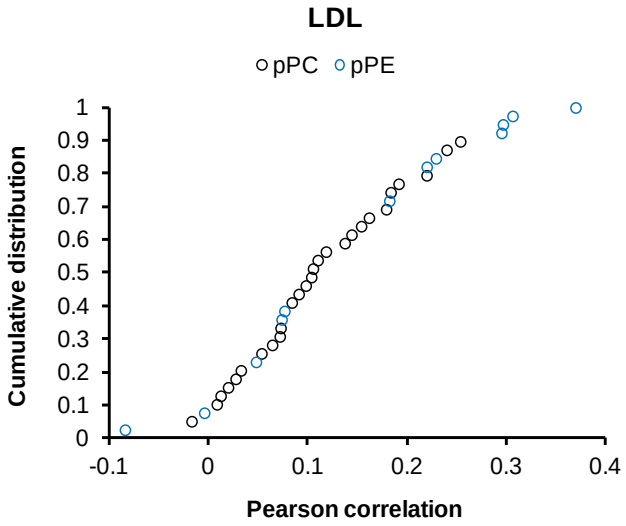


Indian

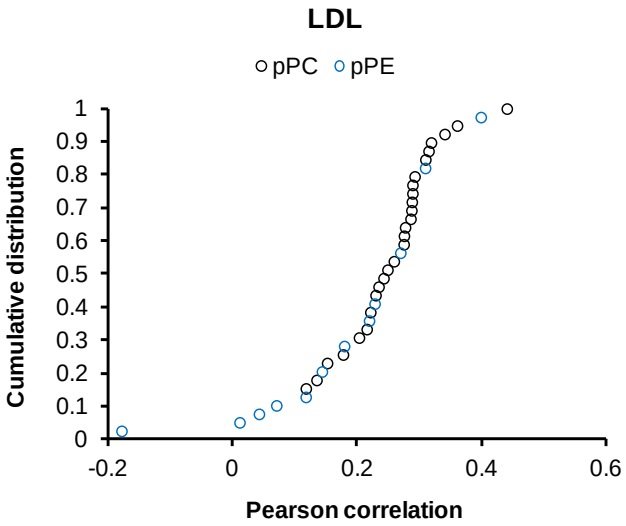


Chinese

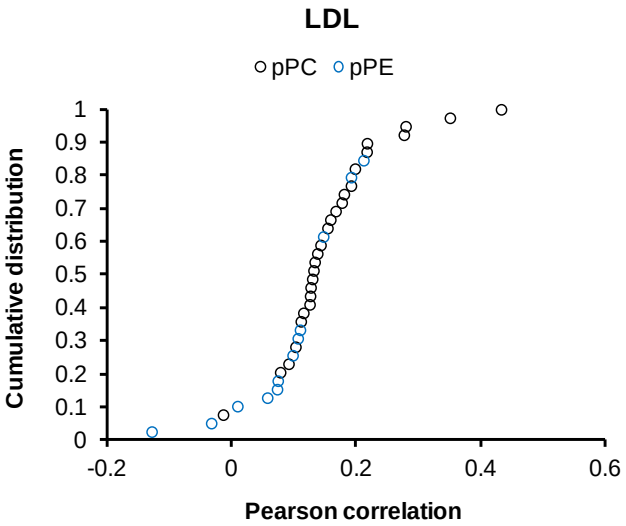
(b)



Malay



Indian



Chinese

Figure 6-9: Pearson correlation between p lipid species (pPE and pPC) and clinical measurements of (a) HDL (top panel) and (b)LDL for Malay (top panel), Indian (middle panel) and Chinese (bottom panel).

(a)

Negative ion mode			
Lipid species	**CE (volts)**	**Lipid species**	**CE (volts)**
PE_436.5/196.1	-30	PI Std_585.6/241.0	-48
PE_452.5/196.1	-40	PI_599.6/241.0	-45
PE_462.5/196.1	-40	PI_807.8/241.0	-55
PE_464.5/196.1	-40	PI_833.8/241.0	-60
PE_476.5/196.1	-40	PI_835.8/241.0	-60
PE_478.5/196.1	-45	PI_857.8/241.0	-60
PE_480.5/196.1	-45	PI_859.8/241.0	-60
PE Std_634.6/196.1	-50	PI_861.8/241.0	-60
PE_698.7/196.1	-55	PI_863.8/241.0	-60
PE_714.7/196.1	-55	PI_883.8/241.0	-60
PE_716.7/196.1	-55	PI_885.8/241.0	-65
PE_718.7/196.1	-55	PI_887.8/241.0	-65
PE_722.7/196.1	-55	PI_909.9/241.0	-65
PE_724.7/196.1	-55	PI_911.9/241.0	-65
PE_726.7/196.1	-55	PS Std_678.6/591.5	-25
PE_738.7/196.1	-55	PS_524.5/437.4	-23
PE_740.7/196.1	-55	PS_786.8/699.6	-25
PE_742.7/196.1	-55	PS_788.8/701.6	-25
PE_744.8/196.1	-55	PS_808.8/721.7	-26
PE_746.8/196.1	-55	PS_810.8/723.7	-26
PE_748.8/196.1	-55	PS_812.8/725.7	-26
PE_750.8/196.1	-55	PS_834.8/747.7	-26
PE_762.8/196.1	-55	PS_836.8/749.7	-26
PE_764.8/196.1	-55	PG Std_665.6/153.0	-50
PE_766.8/196.1	-55	PG_745.8/153.0	-53
PE_774.8/196.1	-55	PG_747.8/153.0	-55
PE_776.8/196.1	-55	PG_769.8/153.0	-55
PE_778.8/196.1	-55	PG_773.8/153.0	-55
PE_790.8/196.1	-55	PG_775.8/153.0	-55
PE_814.8/196.1	-55	GM3_1151.6/290.1	-65
PA Std_591.6/153.0	-50	GM3_1153.6/290.1	-65
PA_645.7/153.0	-55	GM3_1179.6/290.1	-65
PA_671.7/153.0	-55	GM3_1235.6/290.1	-65
PA_673.7/153.0	-55	GM3_1263.6/290.1	-65

(b)

Positive ion mode			
Lipid species	**CE (volts)**	**Lipid species**	**CE (volts)**
PC_482.5/184.1	50	PC_820.8/184.1	60
PC_494.5/184.1	55	PC_822.8/184.1	60
PC_496.5/184.1	55	PC_824.8/184.1	60
PC_510.5/184.1	50	PC_826.8/184.1	60
PC_518.5/184.1	50	PC_828.8/184.1	60
PC_520.5/184.1	50	PC_830.8/184.1	60
PC_522.5/184.1	50	PC_832.8/184.1	60
PC_524.5/184.1	50	PC_834.8/184.1	60
PC_544.5/184.1	50	PC_836.8/184.1	60
PC_568.5/184.1	50	S184M Std_647.7/184.1	60
PC Std_678.7/184.1	60	S184M_701.7/184.1	60
PC_718.7/184.1	60	S184M_703.7/184.1	60
PC_720.7/184.1	60	S184M_705.7/184.1	60
PC_730.7/184.1	60	S184M_729.8/184.1	60
PC_732.7/184.1	60	S184M_731.8/184.1	60
PC_734.7/184.1	60	S184M_757.7/184.1	60
PC_742.7/184.1	60	S184M_759.7/184.1	60
PC_744.7/184.1	60	S184M_785.7/184.1	60
PC_746.7/184.1	60	S184M_787.7/184.1	60
PC_756.7/184.1	60	S184M_813.9/184.1	60
PC_758.7/184.1	60	S184M_815.9/184.1	60
PC_760.7/184.1	60	SM Std_647.7/629.7	30
PC_766.7/184.1	60	SM_701.7/683.7	30
PC_768.7/184.1	60	SM_703.7/685.7	30
PC_770.7/184.1	60	SM_705.7/687.7	30
PC_772.7/184.1	60	SM_729.7/711.7	30
PC_774.7/184.1	60	SM_731.7/713.7	30
PC_776.7/184.1	60	SM_757.7/739.7	30
PC_778.7/184.1	60	SM_759.7/741.7	30
PC_780.8/184.1	60	SM_785.7/767.7	30
PC_782.8/184.1	60	SM_787.7/769.7	30
PC_784.8/184.1	60	SM_813.8/795.8	30
PC_786.8/184.1	60	SM_815.8/797.8	30
PC_788.8/184.1	60	Cer_566.7/264.3	50
PC_794.8/184.1	60	Cer_580.7/264.3	50
PC_796.8/184.1	60	Cer_622.7/264.3	50
PC_798.8/184.1	60	Cer_648.7/264.3	50
PC_800.8/184.1	60	Cer_650.7/264.3	50
PC_802.8/184.1	60	Cer Std_552.7/264.3	50
PC_804.8/184.1	60	GluCer Std_588.7/264.3	50
PC_806.8/184.1	60	GluCer_700.7/264.3	50
PC_808.8/184.1	60	GluCer_784.7/264.3	50
PC_810.8/184.1	60	GluCer_810.7/264.3	50
PC_812.8/184.1	60	GluCer_812.7/264.3	50
PC_818.8/184.1	60		

(c)

Sterols	
Lipid species	**CE (volts)**
Chol Std_375.3/161.0	45
Chol_369.3/161.0	45
CholE Std_375.3/167.0	45
CholE_369.3/161.0	45
24-OHC Std_374.3/159.0	50
24-OHC_367.3/159.0	45
7-OHC Std_392.3/159.0	50
7-OHC_374.3/159.0	50
7-OHC_385.3/109.0	50
7-ketone Std_408.3/175.0	50
7-ketone_401.3/175.0	50
7-dehydroC Std_372.3/220.2	50
7-dehydroC_367.3/215.2	50

Table 6-1: List of MRMs measured in this study and their respective collision energies (CE) for (a) negative ion mode (b) positive ion mode (c) sterol analysis.

Classification		Lipid species	Type II ANOVA p-value	Permutation p-value
Carbon chain composition	32, 34 C	PE	1.00×10^{-03}	0.001
	36 C		1.98×10^{-15}	
	38, 40, 42 C		8.45×10^{-20}	
	32, 34 C	PG	2.15×10^{-6}	
	36 C		2.27×10^{-13}	
	36 C	PS	2.40×10^{-19}	
	38, 40, 42 C		4.41×10^{-22}	
Degree of unsaturation	0	PE	4.15×10^{-05}	
	1		1.78×10^{-07}	
	2		8.66×10^{-06}	
	>3		2.39×10^{-20}	
	0	PS	1.78×10^{-22}	
	1		2.59×10^{-19}	
	2		3.33×10^{-16}	
	>3		4.41×10^{-22}	
Ester/ether species	Ether	PE	7.00×10^{-03}	
	Ester		8.34×10^{-25}	0.003
	Ether	PC	1.20×10^{-18}	0.001
	Ester		3.14×10^{-04}	

Table 6-2: Regression Analysis from type II ANOVA and permutation analysis showing p-value for ethnicity (after accounting for age, gender and interaction between gender and ethnicity), for carbon chain composition, degree of unsaturation and a/p species.