UNIVERSITY OF CENTRAL OKLAHOMA

GRADUATE COLLEGE

Identification of Geographic and Cultural Ancestry through Chemical Analysis of Latent Fingerprint Residues

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

BY

BRYNN MYERS
Edmond, Oklahoma
[2023]

IDENTIFICATION OF GEOGRAPHIC AND CULTURAL ANCESTRY THROUGH CHEMICAL ANALYSIS OF LATENT FINGERPRINTS

By: Brynn Myers

Jackson School of Graduate Studies at the University of Central Oklahoma

A THESIS APPROVED FOR
THE W. ROGER WEBB FORENSIC SCIENCE INSTITUTE

May 2024

By: _____
Dr. Mark McCoy

_____
Dr. Thomas Jourdan

_____
Ms. Cait Porterfield

_____
Dr. Zhibo Yang

Acknowledgments

This journey was long, but well worth it. I would like to begin by first and foremost thanking my family: my mom, my dad, and my wonderful husband who have been with me and encouraged me along the way. I also want to thank the amazing faculty and staff at the University of Central Oklahoma, as well as those at the University of Oklahoma. Without them, this project would not have been possible. Finally, but not least important, I want to give thanks to God. My faith has gotten me through so many hard times, and through all of it, He has been there to guide me.

Contents

Table of Figures

Abstract

The permanence and uniqueness of fingerprints and the analysis of physical characteristics associated with fingerprints have allowed for the identification of individuals in forensic investigations. However, there is additional information that may be obtained from latent fingerprints. Fingerprint residues are comprised of chemical components – those which are created and secreted from the body naturally (endogenous) and those from products outside the body that are ingested or applied on the skin (exogeneous) – that have previously been used by researchers to identify an individual's age, gender, and even lifestyle. This information may be useful for the inclusion and exclusion of suspects when physical fingerprint characteristics may be unidentifiable or may be a means of adding additional information to preexisting physical characteristic analyses. However, there is currently no scientific literature regarding the ability to identify an individual's geographic and cultural ancestry through analysis of latent fingerprint residues. The purpose of this study was to identify a chemical pattern which might allow investigators to identify an individual's geographic and cultural ancestry through the analysis of latent fingerprint residues using Mass Spectrometry – specifically Desorption Electrospray Ionization - Mass Spectrometry – as a medium for analysis. Individuals from varying racial backgrounds (Hispanic, Asian, European, African, and Middle Eastern) were asked to deposit a single right index print onto a piece of Mylar film. Following collection of prints, samples were analyzed using the DESI Synapt G2 instrument. Results of the study indicated that identification of geographic and cultural ancestry based on latent fingerprint residues may be possible, but further study is needed. In most cases, it was seen that an individual has significant agreement of chemical residues within their fingerprints. However, for individuals of the same or even different geographic and cultural ancestries, there are less definitive results as to if the chemistry of their latent print residues is

similar or significantly different. In total, no distinct observations can be made from the data obtained, and further investigation is required to establish any relationships between geographic and cultural ancestries and the chemistry of fingerprint residues.

Introduction

Latent fingerprints are a type of identifiable information which are often left by criminals when they are committing crimes. Latent fingerprints are typically used for the identification or exclusion of a known print to an unknown print. However, with proper application of other technologies, it may be possible to obtain additional information about a suspect.

Some research has been done in the application of chemical instrumentation as a method of characteristic identification, such as age, gender, or lifestyle traits (Hinners et al., 2018, 2020; O'Neill et al., 2020; Zhou & Zare, 2017). However, no research has been conducted on the application of these techniques for the identification of an individual's geographic and cultural ancestry.

The use of mass spectrometry has been proven to be useful in identifying other characteristics of an individual (Huynh et al., 2015; O'Neill et al., 2020; Popa et al., 2010). As a result, the application of this type of instrumentation may prove useful in the identification of geographic and cultural ancestry from latent fingerprint residues. Specifically, the application of imaging mass spectrometry (such as MALDI-MS or DESI-MS) or traditional chromatographic mass spectrometry (such as GC-MS) has proven to provide the most information in regard to analysis of latent fingerprint residues (Hinners et al., 2018, 2020; Huynh et al., 2015; O'Neill et al., 2020; Popa et al., 2010; Zhou & Zare, 2017).

While methodologies for determining geographic and cultural ancestry have been scrutinized in the past, this type of identification may provide information that is useful for inclusions and exclusions - to aid investigators when trying to narrow down a list of suspects. The identification of geographic and cultural ancestry through latent fingerprint residues has not been researched previously and thus has the potential to be of benefit to the field of forensic science.

Literature Review

Fingerprint History

Fingerprints have long been used in forensic investigations as a way of individualization. Friction ridge skin is a type of skin that is made up of a series of ridges and furrows that can leave behind a unique impression. It appears on different areas of the body, including fingers, toes, palms of the hands, and soles of the feet (Houck & Siegel, 2010). The use of friction ridge skin, specifically fingerprints, as a means of identification dates to around 221 B.C. when patterns were used for identification by the Chinese community (Holder & Laub, 2012). Based on artifacts that have been found from this time period, it is known that the Chinese recognized the uniqueness of fingerprints to an individual as they would often use them as a way to show authorship or sign contracts (Holder & Laub, 2012).

Some of the first recorded observations about friction ridge skin were made by a European scientist named Dr. Nehemiah Grew (Holder & Laub, 2012). In 1788, J. C. A. Mayer would be the first individual to record the uniqueness of friction ridge skin, quoting that "arrangement…is never duplicated in two persons" (Holder & Laub, 2012). Although the study of friction ridge skin continued after 1788, only a few observations led to major developments in forensic science. Dr. Johannes Purkinje classified a series of nine categories that fingerprints could fall into and later, Sir William James Herschel was cited as being the first person to note the permanence of latent fingerprints (Holder & Laub, 2012). One of the most well-known names in friction ridge analysis and latent fingerprint research is Sir Francis Galton. Galton is known as having written the first book on fingerprints, published in 1892 (Holder & Laub, 2012). He further established the case for friction ridge skin's uniqueness and persistence – the eventual basis for use of friction ridge skin for forensic identification (Holder & Laub, 2012). Additionally, Galton established a series

of markings to individualize a fingerprint, called Galton details. Galton details include specific fingerprint characteristics such as bifurcations, ending ridges, short ridges, and enclosures (Holder & Laub, 2012). Galton was the first individual to conduct statistical studies on fingerprints and was therefore able to establish their uniqueness. The first use of fingerprints in any type of law enforcement setting was in 1891 by statistician named Juan Vucetich who was employed by police in Argentina (Holder & Laub, 2012). He studied what Galton had previously researched and created his own classification system and used it as a way to individualize prisoners (Holder & Laub, 2012). In 1898, Bengali would become the first country to rely on fingerprints as a method of individualization during a criminal trial and to secure a conviction using fingerprints (Holder & Laub, 2012).

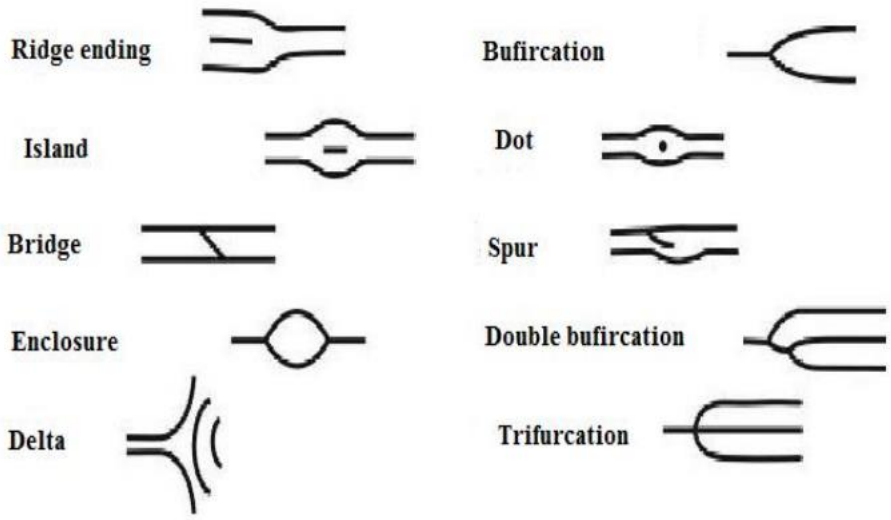


Figure 1 Galton Details (Dwairi et al., 2020).

In addition to Galton details, there have been various systems of fingerprint classification which have been developed over time. One of the most popular systems is that developed by Henry Battley and Fredrick Cherrill (Holder & Laub, 2012). Their classification allowed individual fingerprints to be classified into categories. These categories fall into three broad groups: loops, whorls, and arches (Houck & Siegel, 2010). Within these categories, there are further

classifications: ulnar loop (left and right), radial loop (left and right), plain whorl, central pocket loop whorl, double loop whorl, accidental whorl, plain arch, and tented arch (Houck & Siegel, 2010). The classifications of ulnar and radial loop are less commonly used today and are often designated as just left or right loops. Figure 2 shows examples of these pattern classifications.

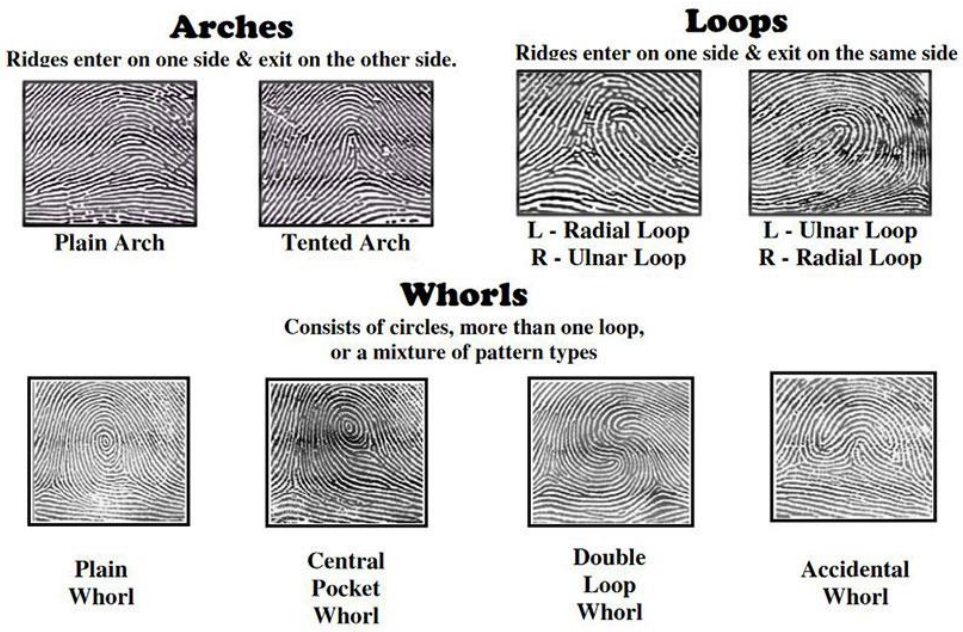


Figure 2 Fingerprint Classifications and Sub-Types (Thakkar, 2007)

Characteristics and Processing of Friction Ridge Skin

Impressions of friction ridge skin patterns at crime scenes can be found in multiple forms, including patent, plastic, or latent. Patent prints are fingerprints that are immediately visible and are left behind in some type of substrate, such as paint or blood (Houck & Siegel, 2010). Plastic prints are true impressions which are 3-dimensional and left behind in a soft, malleable material such as clay or putty (Croxton et al., 2010). Finally, there are latent prints. Latent prints are ridge patterns that are deposited in sweat and are not visible to the naked eye, requiring additional processing to be visible (Holder & Laub, 2012) Latent fingerprints are the most common type of fingermark found at a crime scene (Croxton et al., 2010).

Many forms of development exist that can be used to visualize ridge patterns of latent fingerprints (Holder & Laub, 2012). Some of these detection methods include chemical methods, physical methods, optical techniques, photography, or the use of chemical instrumentation (Croxton et al., 2010). These development techniques are often based on the composition of latent fingerprint residues and the substrate the latent print is deposited on (Croxton et al., 2010).

Latent fingerprint residues are made up of a series of exogenous and endogenous compounds (Hinners et al., 2018), as well as both water-soluble and water-insoluble compounds. Endogenous compounds are chemical compounds that come directly from the human body (Hinners et al., 2018). They are biological and include common compounds such as triacylglycerols, proteins, and amino acids (Hinners et al., 2018). Exogenous compounds, however, are chemical compounds that do not occur naturally in the human body (Hinners et al., 2018). These compounds can come from many places, including things such as drug compounds and gunshot residue, as well as more common products such as soaps, lotions, and other cosmetics (Hinners et al., 2018). Many endogenous and exogenous compounds, as those described above, can be found in latent fingerprint residues due to the production of sweat and oils, or from the use of products (Holder & Laub, 2012). Components of the water-soluble portion of fingerprints include salts and amino acids, while the water-insoluble portion is often made of molecules (such as proteins) and nonpolar acids (such as fatty acids) (Hinners et al., 2018).

In recent years, latent fingerprint research has shifted away from processing latent fingerprints and analyzing Galton features. Instead, many researchers have begun to look at the types of chemical information that can be obtained about an individual from the residues found in latent fingerprint secretions. Much of this research has utilized various forms of mass spectrometry to identify a variety of characteristics (Hinners et al., 2018, 2020; Huynh et al., 2015; O'Neill et

al., 2020; Zhou & Zare, 2017). Researchers so far have had some success in distinguishing gender, age, and even lifestyle (Hinners et al., 2018, 2020; Huynh et al., 2015; O'Neill et al., 2020; Zhou & Zare, 2017).

Mass Spectrometry

Mass spectrometry (MS) is a form of chemical instrumentation commonly used in analytical chemistry to determine the chemical composition of a substance by distinguishing the mass of molecules that make up the substance (Gross, 2017). The first mass spectrometer was developed by Joseph John Thomas in the early 1900s and helped lead to the discovery of atoms and isotopes (Gross, 2017). Mass spectrometry is founded on the basic principle that ions (protons and electrons) can be formed from organic compounds, allowing for the generation of a mass-to-charge (m/z) ratio (Gross, 2017). This m/z ratio can then be detected and allows for the formation of a graph that can be interpreted to determine what molecules are contained within the analyzed substance, as well as the individual abundance of each ion (Gross, 2017).

The types and numbers of mass spectrometry systems have grown over the years. In addition to standard mass spectrometry, various forms of mass spectrometry analysis have been introduced such as Secondary Ion Mass Spectrometry (SIMS), Matrix-Assisted Laser Desorption/Ionization (MALDI), Electron Ionization (EI), Inductively Coupled Plasma (ICP), and many others (Gross, 2017). Additionally, mass spectrometry has been partnered with other forms of chemical analysis, such as Gas or Liquid Chromatography, allowing for the GC-MS and LC-MS analyses that are commonly used in the forensic field for drug analysis (Gross, 2017). While mass spectrometry technology gives the ability to identify components within a chemical mixture, it does not have the ability to separate the components (Jones, 2019). Addition of gas or liquid chromatography allows for the separation of components in chemical mixtures (Jones, 2019).

Together, the two instruments allow an analyst to separate and identify components of a chemical mixture in a single step (Jones, 2019). Some of these mass spectrometry systems have been used to determine important forensic information from the chemical residues of latent fingerprints such as allowing for the identification of age, lifestyle, and even gender (Hinners et al., 2018, 2020; Huynh et al., 2015; O'Neill et al., 2020; Popa et al., 2010; Zhou & Zare, 2017).

Mass spectrometers are made up of a series of components that each serve a different purpose. The first is an ionization source that produces ions from the sample (de Hoffmann & Stroobant, 2007). An ion is an atom that has a different number of protons and electrons, giving the atom an overall net charge (Flowers et al., n.d.). The mass analyzer is then used to separate the ions by size before the ions move into the detector where they are counted (de Hoffmann & Stroobant, 2007). The final component is the processing system that produces a set of data that can be interpreted. Each of these general components can be further broken down into categories based on how they work – such as chemical ionization, field ionization, or desorption (de Hoffmann & Stroobant, 2007). Due to variations in the way that the different ion sources work, it is important to make sure that a mass spectrometer is equipped with the proper components prior to beginning any experimentation (de Hoffmann & Stroobant, 2007).

Mass spectrometry can be used on latent fingerprints due to their chemical composition, specifically the components of sweat residues, which are primarily composed of proteins, amino acids, salts, and water-insoluble fatty acids (Hinners et al., 2018; Holder & Laub, 2012). Matrix-Assisted Laser Desorption-Ionization Mass Spectrometry (MALDI-MS) or Desorption Electrospray Ionization (DESI) are two types of mass spectrometry that are often used to study proteins (de Hoffmann & Stroobant, 2007). One benefit of MALDI-MS is that these types of mass spectrometers have a limit of detection that is typically reported somewhere between femtomoles

(10^{-15}) and picomoles (10^{-12}) (de Hoffmann & Stroobant, 2007). This exceedingly small limit of detection means that MALDI-MS has the potential to find micro molecules that other types of mass spectrometry may be unable to detect.

Lipids can also be identified using mass spectrometry. These methods of detection are more often based on simple mass spectrometry or may be coupled with Gas Chromatography (GC) or High-Performance Liquid Chromatography (HPLC) (de Hoffmann & Stroobant, 2007). One downfall of the use of this form of mass spectrometry detection is that while general information can be obtained about the elemental composition, no other structural information can be obtained, thus leaving many gaps in assessing the chemical make-up (de Hoffmann & Stroobant, 2007).

Desorption Electron Spray Ionization-Mass Spectrometry

Desorption electron spray ionization-mass spectrometry (DESI-MS) has recently had increased use in the world of Imaging Mass Spectrometry (IMS). IMS is a type of mass spectrometry in which a surface area of a sample is scanned and translated into a distribution pattern on the surface (Gross, 2017). The distribution pattern then correlates to an optical image which can be viewed as a picture (Gross, 2017). DESI-MS was first introduced in 2004, making it one of the newer forms of mass spectrometry (Gross, 2017). This type of mass spectrometry allows an examiner to produce not only a mass spectrum, but also an image which can be analyzed for additional information. DESI offers advantages over other types of mass spectrometry as it does not require any treatment of the sample prior to insertion into the instrument base and it additionally analyzes samples at atmospheric pressure, leading to a type of mass spectrometry known as ambient mass spectrometry (Gross, 2017). The general process of ambient mass spectrometry can be seen in Figure 3.

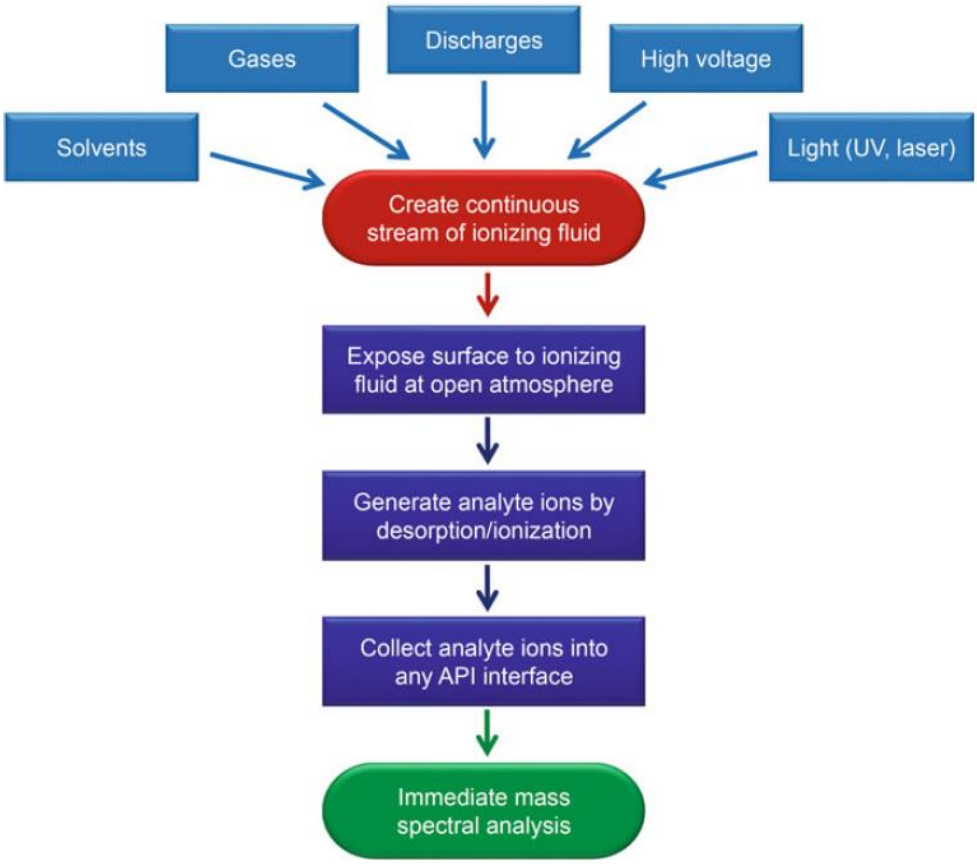


Figure 3 General Process for Ambient Mass Spectrometry (Gross, 2017)

DESI is a relatively non-destructive technique which uses a stream of ions to desorb a sample. This process is based off the concept of electrospray ionization (ESI). In ESI and also DESI, a capillary set at a high voltage (in comparison to a counter electrode) releases a spray of sample ions contained in the sprayer solution (Gross, 2017). The main difference between ESI and DESI comes from the angle of the spray. While ESI is traditionally sprayed at a straight angle, parallel to the instrument stage, DESI shifts this to an angle of 45° toward the sample stage (Gross, 2017). The 45° angle of the sprayer in relation to the sample and the mass spectrometer entrance creates an alignment which is shaped like a V (Gross, 2017). As the DESI instrument analyzes the sample, a carrier gas at a high velocity moves the charged aerosol droplets towards the sample surface, creating a buildup of charge on the sample (Gross, 2017). Ultimately, this creates a thin

layer of liquid film on the sample, which will dissolve and lift analyte ions from the sample (Gross, 2017). The lifted ions then move toward the mass spectrometer inlet where they can be analyzed (Gross, 2017). This process is shown in Figure 4.

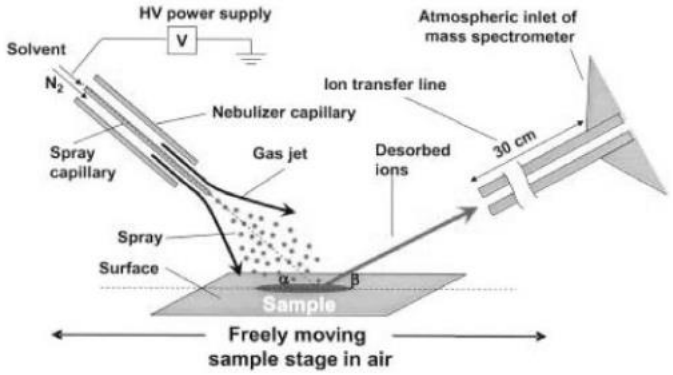


Figure 4 Process of ionization by DESI-MS (de Hoffmann & Stroobant, 2007).

Mass Spectrometry for Latent Fingerprint Development

Authors Hinners, Thomas, and Lee used mass spectrometry as a way to determine the age of latent fingerprints (2020). In their work, they stated that this type of analysis could potentially be beneficial in determining the time since deposition occurred (Hinners et al., 2020). They stated additionally that there have been various prior investigations into this same type of analysis using fatty acid diffusion with time-of-flight secondary ion mass spectrometry as well as with MALDI-MSI (Matrix Assisted Laser Desorption Ionization – Mass Spectrometry Imaging) (Hinners et al., 2020). GC/MS has also been shown to be useful in determining fingerprint age into two broad categories – less than or greater than 8 days - however, this depended on the substrate on which the latent was deposited and the environment in which the fingerprint was left (Hinners et al., 2020). Hinners, et al. (2020), looked specifically at the use of MALDI analysis for determining the age of fingerprints through the ozonolysis of triacylglycerols found in latent fingerprint residues. Through their analysis, they determined that unsaturated triacylglycerols tend to be more abundant than saturated triacylglycerols when a fingerprint is fresh (Hinners et al., 2020). They

also reported that these unsaturated compounds decreased over time and that at seven days, the amounts of saturated triacylglycerols became higher than the number of unsaturated triacylglycerols (Hinners et al., 2020). One important impact that the authors noted was that in addition to the standard application of MALDI-MS to determine fingerprint age, when forensic development techniques were applied to the prints, triacylglycerols and their products could still be detected (Hinners et al., 2020). This is important as it allows for further analysis of latent prints, despite the application of development powders or chemicals.

Triacylglycerols have also been used to determine information regarding diet, exercise, and health in individuals (O'Neill et al., 2020). A study conducted by O'Neill et al. (2020) employed the use of MALDI-MSI to analyze triacylglycerols found in the latent fingerprint residues and see how they related to lifestyle (O'Neill et al., 2020). The researchers determined that there is much variability in triacylglycerol levels among individuals of different lifestyles (O'Neill et al., 2020). For example, vegetarian diets were shown to increase the amounts of saturated triacylglycerol levels when compared to those with standard diets (O'Neill et al., 2020). Researchers also noted that the amount of exercise affected triacylglycerol levels in males (O'Neill et. al., 2020). Their conclusions showed that while female triacylglycerol levels were not particularly impacted by exercise, male participants were shown to have lower saturated triacylglycerol levels if they were more active (O'Neill et. al, 2020).

In some cases, in this experiment, individuals had compounding lifestyle factors which prevented the identification of the contributing factor. For example, while the authors noted that there was a general trend suggesting individuals with type 2 diabetes had a higher level of saturated triacylglycerols, they also stated that some of these individuals did not exercise regularly, or at all (O'Neill et. al., 2020). This potentially impacted the results as it is difficult to determine what is

the true cause of the differences in triacylglycerol levels (O'Neill et. al., 2020). O'Neill et. al. also noted that due to a small sample size used during experimentation, further research would be needed regarding their results.

These same researchers also used MALDI-MS and Tandem Mass Spectrometry to determine lifestyle by looking for chemical compounds in latent fingerprints (Hinners et al., 2018). In this experiment, the authors looked for chemicals commonly found in various bug sprays, sunscreens, as well as other compounds such as oils, alcohols, and citrus (Hinners et al., 2018). Through their analysis, they investigated how different consumer products would produce an exogenous compound composition in latent fingerprints and how this could be utilized as a method for determining lifestyle (Hinners et al., 2018). For example, the presence of bug spray and sunscreen reflected the idea that the user would have a lifestyle where they spent a lot of time outdoors (Hinners et al., 2018). The idea behind this analysis was that if a fingerprint was deposited at a crime scene, and contained components of a certain product, investigators could look for that same brand of product in the belongings of a suspect (Hinners et al., 2018). Similar investigations were done with oils, alcohol, and citrus fruits (Hinners et al., 2018).

One final important characteristic which was investigated by Huynh et. al. (2015) was determining gender from latent fingerprint residues. This method is slightly different from those previously discussed as it does not employ the use of mass spectrometry. This method instead looked at the use of a bioassay as a method for distinguishing gender (Huynh et al., 2015). The researchers found that they were able to extract and separate amino acids from the lipid-based components of volunteers' latent fingerprint residues (Huynh et al., 2015). Results indicated that when using mimicked fingerprint samples created based on a randomized concentration of amino acids, researchers were able to correctly determine gender with a 99% chance of accuracy (2015).

This, as with other methods, would allow for valuable information to be determined from the chemical analysis of latent fingerprint residues.

Zhou and Zare (2017) investigated the idea of pulling personal information from latent fingerprints using desorption electrospray ionization mass spectrometry (DESI-MS). In their work, they stated that previous research showed that there were very few differences found in the sebum composition of individuals of different gender and age; however, there was a difference in the lipid concentrations in serum in blood (Zhou & Zare, 2017). They selected eight individuals to provide fingerprints for their study who were considered to be racially diverse and covered a span of ages (Zhou & Zare, 2017). In addition to the samples of latent fingerprints collected from the eight participants, lipid samples were obtained by swiping a glass slide across the forehead of a separate sample group of 203 individuals (Zhou & Zare, 2017). The collected fingerprints and forehead swipes were analyzed using DESI-MS, and a lipid profile was created from the mass spectrometry data (Zhou & Zare, 2017). The first significant finding was that lipid profiles taken from separate places on an individual's body were consistent but may differ from one individual to another (Zhou & Zare, 2017). When these samples were run under DESI-MS it was found that gender, age, and ethnicity could be determined to a relatively high degree of certainty (Zhou & Zare, 2017). Additionally, they were able to identify overlapping fingerprints by analyzing the lipid profiles of individuals and noting the differences which existed (Zhou & Zare, 2017). The research they conducted showed that it was possible to determine gender and age using an individual's fingerprint, however ethnicity was only identified in lipid profiles obtained from forehead swipes (Zhou & Zare, 2017).

Principal Component Analysis and Its Applications in Forensic Science

Principal component analysis, or PCA, is a type of statistical analysis which applies machine learning to large data sets in order to make it easier for patterns to be identified in the dataset. This is completed by reducing a large dataset into a smaller one which still possesses the majority of the information of the original data. While this process of principal component analysis can be performed by hand, it is often easier to perform by computer learning. When the data is re-graphed using the newly reduced points, it makes it possible to easily view any patterns that exist within the dataset. Within principal component analysis, multiple new variables (known as Principal Components) are identified by reducing the number of dependent variables in the original data and replacing them with the new intercorrelated variables. Each of these principal components are identified by using PC (meaning principal component) as well as a number. These numbers are given in order of decreasing variance - that is to say that PC1 will explain the most variance, PC2 the second most, and so on. These PCs are a measure of variance within a sample and do not represent any one specific compound or variable. The goal of principal component analysis is to be able to quickly scan a dataset to identify correlations or lack there of in a large dataset.

Several forensic researchers have applied principal component analysis as a means of analyzing their data during research. Authors Riva and Champod applied principal component analysis to view impressions made by various firearms from the manufacturer Sig Sauer. Their PCA plot (shown in Figure 5) shows the distribution of firearms impressions made by the same firearm (shown in gray) and different firearms (depicted in black) (Riva & Champod, 2014). It can be seen that for the two firearms depicted (W1 on the left and W2 on the right) the breech face impressions have a fairly good separation (Riva & Champod, 2014). For both of these firearms, there is a good separation, meaning that it would be easy to distinguish the impressions from these

firearms. Additionally, PCA was performed on the firing pin impressions. The firing pin impressions were less distinct, and thus did not have as much separation in the PCA (Riva & Champod, 2014). This means that on their own, it would be easier to distinguish the firearms in the study based on breech face impressions, rather than firing pin impressions. The researchers then examined the PCA for both the breech face and firing pin impressions together. The PCA for the combined data was even more distinct than the separated data. The authors determined that based on this, using both firing pin impressions and breech face impressions together, it is much easier to distinguish between firearms manufactured by the same company (Riva & Champod, 2014).

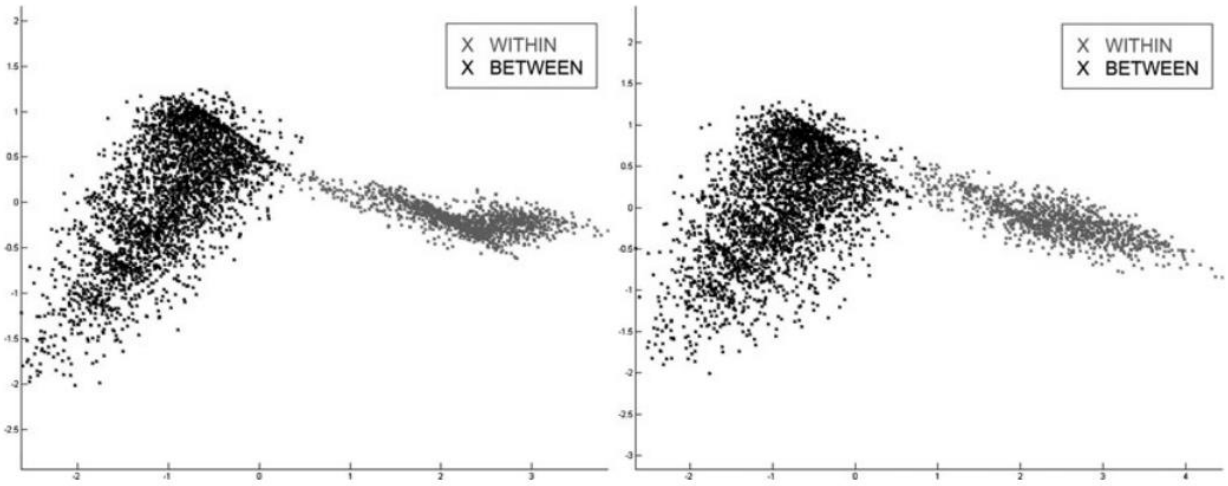


Figure 5 PCA plot of breech face impressions from Riva and Champod (2014).

Additionally, authors Gal, Oravec, Gemeiner, and Ceppan use PCA to compare various inks from three different types of inkjet printers. The PCA they obtained is shown in *Figure 6*.

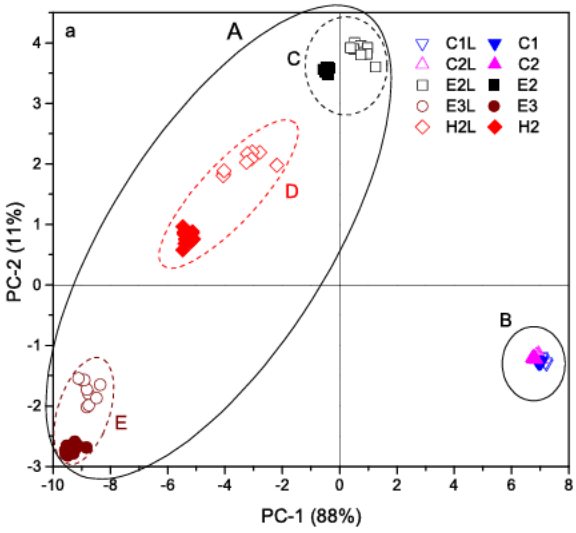


Figure 6 PCA for various inks and printers (Gal, Oravec, Gemeiner, and Ceppan, (2015)).

For these authors, inks represented by C were from a Cannon printer, H represented HP printers, and E denoted Epson. The authors analyzed the first 10 principal components but found that the first 3 accounted for 99.76% of the variability in the spectral data which was analyzed (Gál et al., 2015). The principal components shown in *Figure 6* denoted wavelengths between 500-700 and 750-1000 nm. The authors note that using *Figure 6* it can be possible to determine similarities and differences in samples based on clustering. Clusters found closer together are more similar than those which are further apart. From this graphic, we can see that there was a cluster (Group B) which contained inks from Cannon printers only, while the Epson and HP printers were more spread out within Group A (Gál et al., 2015). They note that the inks found in Group B may contain mostly carbon black as a colorant. Additionally, it can be seen that Groups C, D, and E each form their own smaller clusters containing distinct samples E2, H2, and E3 respectively. Based on this PCA, as well as additional study performed by the authors, they determined that the use of PCA in conjunction with Vis-NIR fibre optics reflection spectra (FORS) of inkjet inks, it would be possible to create a classification of these inks which could be used to distinguish inks for forensic use.

Geographic and Cultural Ancestry Based Metabolism

As previously shown, there are many differences in the chemistry of an individual's latent fingerprints. There are also other chemical differences in individuals of varying geographic and cultural ancestries. Metabolism is one of the biological processes which have been shown to differ based on an individual's racial background. One of the most documented is the concept of "blushing" in individuals of Asian descent. Studies have shown in general that persons of Asian descent are less likely to experience alcoholism and are also more likely to be abstinent compared to individuals of other racial backgrounds (Wall & Ehlers, 1995). One reason is believed to be related to physiological reasons known as flushing or "blushing (Wall & Ehlers, 1995). This reaction has been shown to be experienced by about half of individuals of Asian descent, specifically those with Chinese, Japanese, and Korean backgrounds (Wall & Ehlers, 1995). This reaction is specifically caused by an increase in blood flow to the skin of the upper body, specifically the face, chest, and neck (Wall & Ehlers, 1995). Additional symptoms may include tachycardia, hypertension, headache, nausea, and vomiting (Wall & Ehlers, 1995). This flushing response has been shown to be related to the alcohol metabolism pathway (shown in Figure 5) (Wall & Ehlers, 1995). Specifically, researchers believe that elevated levels of acetaldehyde are responsible for the flushing reaction (Wall & Ehlers, 1995).

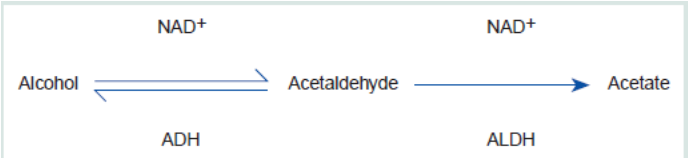


Figure 7 Pathway for Alcohol Metabolism (Wall & Ehlers, 1995)

This so called "Asian Glow" results due to an inherited deficiency related to aldehyde dehydrogenase 2 (Brooks et al., 2009). This enzyme is responsible for converting the acetaldehyde metabolism product into acetate (Wall & Ehlers, 1995).

Another example of metabolism which differs based on geographic and cultural ancestry is the breakdown of lactose from dairy products. A study was performed in 1980, in which adolescents from ages 14 to 19 were given lactose and studied to determine intolerance or non-persistence. The studies showed that 83% of black, 62% of Hispanic and 32% of white subjects were determined to be lactose non-persistent (Lapides & Savaiano, 2018). This means that even when consuming something with low lactose, such as skim milk, individuals that are lactose non-persistent will have symptoms such as bloating, flatulence, cramps, or laxative effects as early as 1 to 2 hours following consumption (Lapides & Savaiano, 2018). Another study released in a 1966 article in the *Journal of American Medical Association* stated that 70% of African Americans experienced lactose and milk intolerances in comparison to about 5% of White individuals (Wall & Ehlers, 1995).

Identification of Geographic and Cultural Ancestry in Forensic Science

Identification of perpetrators involved in crimes is one of the major goals of forensic science. Over the years, many methods have evolved which allow for the identification of key characteristics of those committing crimes, such as the use of security cameras and facial recognition, as well as DNA analysis. Methods like DNA analysis and fingerprint examination have become valuable for identifying a suspect related to a crime.

However, in some cases, there is no primary suspect or suspects. These are the cases that end up going unsolved for years or are perhaps never solved. While an exact identification is desired, it may not be feasible. Instead, it may be possible to identify characteristics and obtain a starting place for investigators to narrow down a list of potential suspects. One form of preliminary identification would be the identification of geographic and cultural ancestry.

Identification of geographic and cultural ancestry is different from racial profiling in that it uses some type of measurement to identify the geographic and cultural ancestry of an individual – a more scientific form of behavioral analysis. It should be noted that this type of identification can very easily slip into racial profiling, and it should not always be taken as a 100% fact. Instead, this form of analysis should be used as a tool to aid investigators in identifying the characteristics of a suspect. While it may be a useful tool during the investigation stage of forensic science, this type of identification should not be admissible in a court of law.

Current Use of Geographic and Cultural Ancestry Identification in Forensic Science

One of the most widely recognized forms of geographic and cultural ancestry identification is in the form of eyewitness testimony (Nayak & Khajuria, 2019). In these cases, an individual who either witnesses the crime or the victim themselves tries to recall any information that they can remember about the perpetrator – to include geographic and cultural ancestry (Nayak & Khajuria, 2019). While it seems that this would be a rather straightforward way of identifying an individual's geographic and cultural ancestry, it can often prove to be inaccurate. According to the Constitutional Rights Foundation, a wrong eyewitness testimony will lead to a wrongful conviction in about half of the cases where it is presented in court (Constitutional Rights Foundation, n.d.) Additionally, there is the issue of the Cross-Race Effect in eyewitness testimony (Wilson et al., 2013) This is a psychological phenomenon that occurs when the eyewitness and the identified suspect are of two different geographic and cultural ancestries (Wilson et al., 2013). The Cross-Race Effect shows that when individuals are of the same geographic and cultural ancestry, an individual's face is much more easily recognized than when individuals are of two different ancestries (Wilson et al., 2013). This can be a major issue as misidentification may lead to an individual being wrongly convicted and imprisoned.

Another form of geographic and cultural ancestry identification is done through DNA analysis. DNA is a fairly accurate method of identification, and it is also a very commonly encountered type of evidence at crime scenes, due to the type of samples which contain DNA (Houck & Siegel, 2010). Research has shown that using the Y-chromosome it may be possible to establish a male lineage and an individual's ethnicity (M'charek et al., 2020) While this can be helpful, it does have a small downfall due to the use of the Y-chromosome. The Y-chromosome is only found in male individuals; thus, this type of identification is useless when the perpetrator is female (National Human Genome Research Institute, 2021). Mitochondrial DNA (also known as mtDNA) is another form of DNA which can provide information about an individual's geographic and cultural background. This type of DNA traces the female lineage and provides information about an individual's family history as it is passed down from the mother.

The microscopic examination of hair can provide some racial characteristics of hair, although this is not a definitive approach. Using microscopic analysis, specific characteristics such as color, density, and the form of the shaft can be used to identify ancestry (M'charek et al., 2020) This type of identification typically allows individuals to fall into one of three subgroups – African, Asian, or European (de la, 2007). The problem with this type of identification is the fact that these three categories are overly broad (de la, 2007). Additionally, there is some problem with this type of identification due to unnatural application of products such as hair dye or chemical curling or straightening which change the characteristics of the hair (M'charek et al., 2020). In addition to problems created by unnatural products used on hair, there are also identification concerns when it comes to mixed-race individuals (Takahashi, 2019). Individuals of mixed race may have different hair characteristics on different strands of hair or have a combination of hair types on a single hair and their hair type may not fall within a distinct category (Takahashi, 2019).

The identification of dental morphology has also been used as a method of identifying geographic and cultural ancestry. These types of identification are common in cases of mass disasters or the discovery of bodies that have been decomposing for some time (Krishan et al., 2015). The use of dental characteristics has shown a few characteristics which seem to be present in individuals of different ancestries (Krishan et al., 2015). For example, Aboriginal people, American Indians, and Eskimos tend to have teeth that are larger in scale and have a wider crown than races such as Lapps or Bushmen (Krishan et al., 2015). Additionally, there may be cultural practices or jobs which result in distinct wear and tear that may be able to help identify the ancestry or ethnicity of an individual (Krishan et al., 2015). Dental identification may have some downfall as there is not an exact science to bite mark/dental analysis (Krishan et al., 2015).

Finally, there has been some research indicating that geographic and cultural ancestry can be determined by the morphology of latent fingerprints. It has been shown that African ancestral groups have more bifurcations and fewer ridge endings when compared to individuals of European or Asian descent (Ford, 2020). It has also been shown that Asian and African groups have lower averages of short ridges, while Native American and Hispanic groups have higher averages of dots than their African and European counterparts (Ford, 2020).

Gaps in the Research

The chemical composition of latent prints has been investigated as a means to determine gender, age, and lifestyle. However, no research to date has explored the chemical analysis of latent fingerprint residues for assessing geographic and cultural ancestry. Based on current research models, there is potential for mass spectrometry to be a method for determining geographic and cultural ancestry based on fingerprint residue composition. The idea that geographic and cultural ancestry can be determined from latent fingerprint residues may

potentially be a helpful starting place for investigators as a means of inclusion and exclusion of suspects. The use of mass spectrometry in the analysis of latent prints may be able to help forensic investigators by creating a potential profile (gender, geographic and cultural ancestry, lifestyle), much like criminal profiling.

Methodology

Sampling

This study utilized purposeful sampling (Bui, 2009). Latent fingerprint samples used in this study were obtained from individuals at the University of Central Oklahoma and Oklahoma City University due to the need to control for age – all participants were ages 18-28. All participation was voluntary, and no incentive was given for participation. Any personal and/or identifiable information (other than gender and geographic and cultural ancestry) was excluded from the study to protect the privacy of the participants. Due to the nature of human involvement in this experiment, approval from the Institutional Review Board at the University of Central Oklahoma was received. The goal of this research was to collect 20 right index fingerprints (10 males and 10 females) from each defined ancestry category:

- 10 Hispanic descent males, 10 Hispanic descent females
- 10 Asian descent males, 10 Asian descent females
- 10 Native American descent males, 10 Native American descent females
- 10 African descent males, 10 African descent females
- 10 European descent males, 10 European descent females

While these were the desired numbers, due to cost and time constraints on testing, a total of 15 individuals provided samples as follows:

- 1 Asian Male (A)
- 1 Black Female* (B)
- 7 European Females* (E)
- 5 Hispanic Females* (H)
- 1 Middle Eastern Female (ME)

Geographic and cultural ancestry was self-identified and defined by the individual participant's responses to a questionnaire prior to fingerprint collection. Self-identification may result in some error in the study based on the survey shown in Appendix A. Future research could benefit from genealogical measures.

Methods

The methodology for this study was based on several similar studies that utilized DESI-mass spectrometry imaging. Individuals were asked to clean their hands to remove any outside chemicals that may be present and dry using chemical-free, unbleached paper towels. Participants were then asked to put a pair of polyethylene gloves on for 30 minutes to allow for sweat accumulation (in case of allergy, other types of gloves were provided). After the allotted time had passed, participants deposited an impression of their right index finger onto a glass microscope slide. Participants pressed prints onto the glass slide for a total of 10 seconds. Prints were analyzed the day following sample collection. Additionally, a mid-course change was made in the number of prints taken from each individual in order to increase the statistical significance. Therefore, the first five individuals B1, E1, E2, H1, and H2 provided only one right index fingerprint, while all other participants provided a total of 3 prints from the right index finger on the same slide.

The print impressions were then analyzed by the Waters Synapt G2-Si coupled with DESI XS stage. This instrumentation was used in collaboration with the University of Oklahoma Mass Spectrometry, Proteomics & Metabolomics (MSPM) Core Facility. The samples were analyzed using standard protocol for the OU MSPM as follows: a solvent of 98:2 (%v/v) methanol:water with 1% formic acid was employed as the solvent mixture and dispensed at a flow rate of 1 µL min^{-1}. Samples were analyzed using a full scan mode (m/z 100 to 1200). Positive ion mode was used with a 0.75 KV voltage capacity and a capillary temperature of 150°C. The scan was set at

150 µm x 150 µm, and resolution of 300 µm/s was used during scans. The resolution is the movement speed of the DESI stage, meaning 2 spectra can be obtained in 1 second. To reduce scan time, only a portion of the prints were analyzed, focusing on the core and area in the center of the fingerprint. This ensured that the primary location for sweat was analyzed, and an optimal sample could be obtained. Average analysis time for each individual slide was about 4 hours per scan. During preliminary analysis to determine the optimal instrument settings, it was determined that an untargeted study would be applied meaning that no particular compound or molecule would be targeted for detection. Initially it was planned to focus analysis on lipid profiles. However, it was observed that degradation of the samples occurred due to temperature during transport from the IRB site at the University of Central Oklahoma to the analysis site at University of Oklahoma. Lipid profiles were deemed to be unsatisfactory for experimentation; instead, the totality of the sample was analyzed, and no specific chemical profiles were targeted.

Following the acquisition of mass spectrometry data, four regions of interest (ROIs) were chosen per print and the individual and group data was analyzed using the MetaboAnalyst website, focusing on principal component analysis (PCA). PCA is a type of statistical analysis which makes it possible to easily analyze a large data set by performing a linear transformation into a new set of variables and simplifying the data set (Sewell, 2007). In this type of analysis, the characteristic with the greatest variance is assigned as the first principal, the second greatest variance as the second principal, and so on (Sewell, 2007). The principal components are a combination of variables that are produced through the process of machine learning. This means that thousands of datapoints entered into the machine learning system are condensed into a series of variables that are easier to analyze for patterns. In the case of mass spectra, the variables are intensities at defined masses. Geena 2 was used to align data from different scans, as each scan had slightly different

scan times. Geena 2 is a tool for filtering, averaging, and aligning mass spectra and is designed to assist in the analysis of high volumes of data and support them for comparative studies (Romano et. al., 2018). Following analysis using PCA, additional analysis was performed using MetaboAnalyst in the ANOVA function. The ANOVA is a statistical comparison which uses the variance of data groups and compares them. This comparison is then defined as either significant ($p<0.05$) or nonsignificant. The values can be viewed graphically or by written table and allows the analyst to determine the most significant contributing values in a dataset.

Data

Individual Data

The first stage of data analysis focused on PCA for each individual. Each print had four ROIs which were chosen and analyzed. For each print, the ROIs were placed as close to the core of the print as possible. The data obtained was input into Microsoft Excel, before being analyzed using the MetaboAnalyst website. PCA graphs for each individual were obtained. Note that each figure represents data for the three prints collected from a single individual.

As seen in Figures 10 and 15, DESI allows for the creation of a visual image in addition to collecting chemical data. This image collection is beneficial in the field of fingerprint analysis because it allows for the physical analysis of latent fingerprints (e.g., analyzing pattern type and Galton details) as well as providing chemical information. The chemical data was analyzed using machine learning software. The obtained data was placed into Microsoft Excel and the prints of each individual were analyzed. Principal component analysis condenses the large data set obtained from each scan into a format that allows patterns in data to be easily identified. Substantial overlap in the obtained data indicates similarities between the latent fingerprints of an individual, meaning that there is a similar chemical profile every time an individual leaves their fingerprint. As seen in Figure 8, there is a high degree of overlap between the three prints from European female 4 (Group 1 corresponding to Print 1, Group 2 corresponding to Print 2, and Group 3 corresponding to Print 3). A high variance is observed in PC 1 (82.9%) and a lower variance in PC 2 (16.2%). Though these two numbers are very different, they are still closer than the variances shown in Figure 11 (97.3% for PC 1 and 2.4% for PC 2). From this, it can be seen that the principal components account for 99.1% and 99.7% of the total variance, respectively. In most cases the total variance is expected to be 70-80% when viewing all the principal components. In this case, we are getting

a higher value for the variance for just the first two principal components when looking at overall accepted criteria. We can see in our loading plot for European Female 4 (Figure 9) that the m/z values of 304.2941 and 332.3244 had the highest influence on the variability of PC 1, while PC 2 was most influenced by 284.3257 and 304.2941. This loading plot represents the values which give the most influence on each principal component, thus influencing the graphical principal component representation. These values can be selected as the most prominent as they are located to the far sides of the X-axis and Y-axis, rather than being found near the more clustered center. Principal components 1 and 2 in Hispanic Female 6 was most influenced by 200.2361 and 301.2824 respectively as seen in Figure 12. The ANOVA for European Female 4 returned that there were no significant m/z points and thus is not shown, while Hispanic Female 6 had a most significant m/z of 673.3442. The ANOVA represents a statistical analysis of the data. The top value found in Figure 14 for the Hispanic individual is the most significant for that data. This means that machine learning identifies it as the species which most influences the overall data. Additional analysis could be performed on this species to potentially determine its structure. These individuals were chosen for discussion as their graphical representations show the two different extremes found in the data. The European Female 4 shows a great degree of overlap, while Hispanic Female 6 shows almost no overlap. All other individuals analyzed using these methods have graphics which fall somewhere between these two discussed individuals. Individuals B1, E1, E2, H1, and H2 only deposited one print each. This means that there was not enough data provided by these individuals to perform an analysis on them alone. Overall results for all individuals are discussed in the conclusion section and additional graphics for each individual can be found in Appendix A.

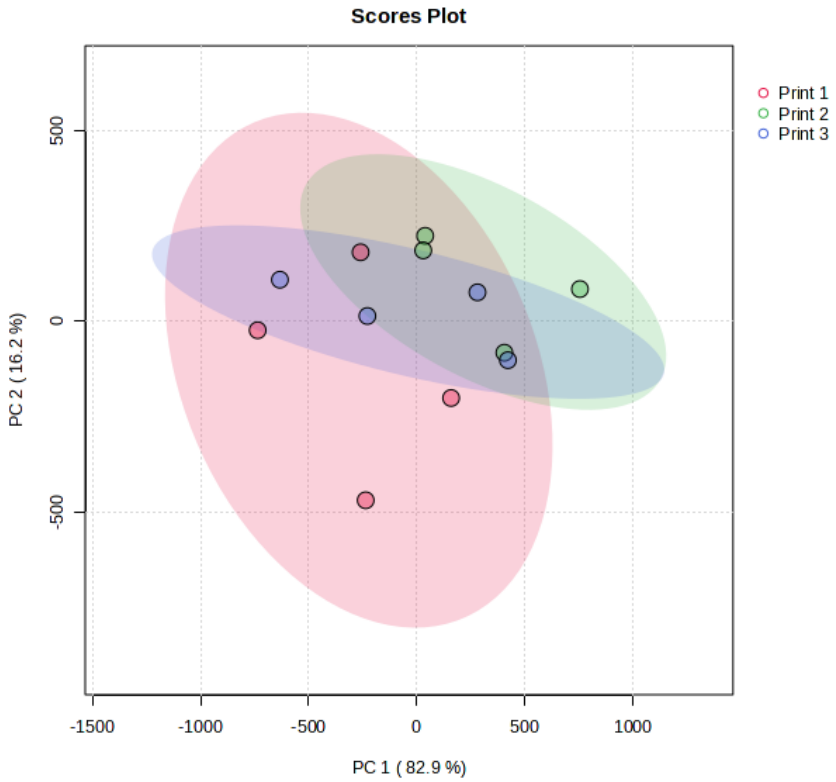


Figure 8 European Female 4. From the graphic representing this European Female, it can be seen that there is a high degree of overlap between the three prints. It can be determined that there is a great deal of similarity in this individual's prints. A total variance of 99.1% is shown between these two principal components.

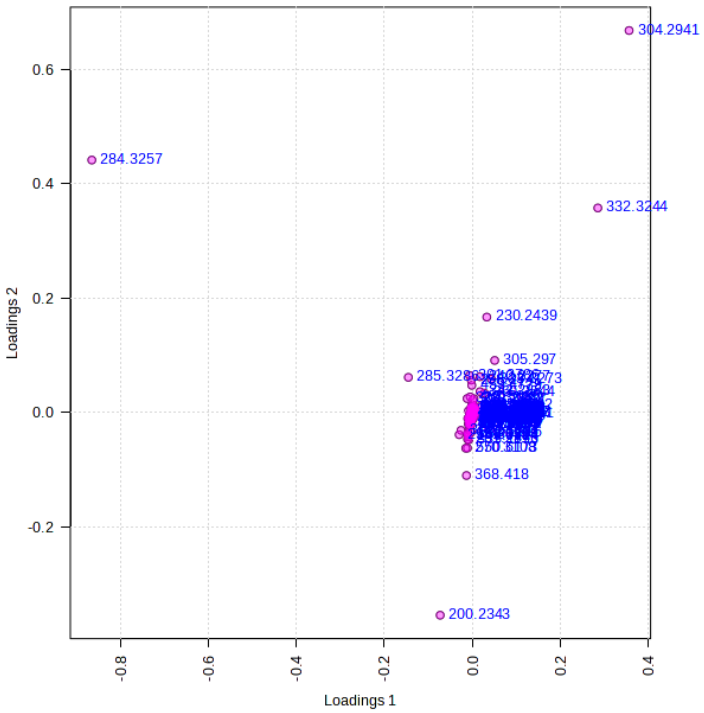


Figure 9 Loading Plot for European Female 4. From these loadings it can be determined that the values of 304.2941 and 332.3244 had the highest influence on the variability of PC 1 (shown on the X-axis), while PC 2 (shown on the Y-axis) was most influenced by 284.3257 and 304.2941.

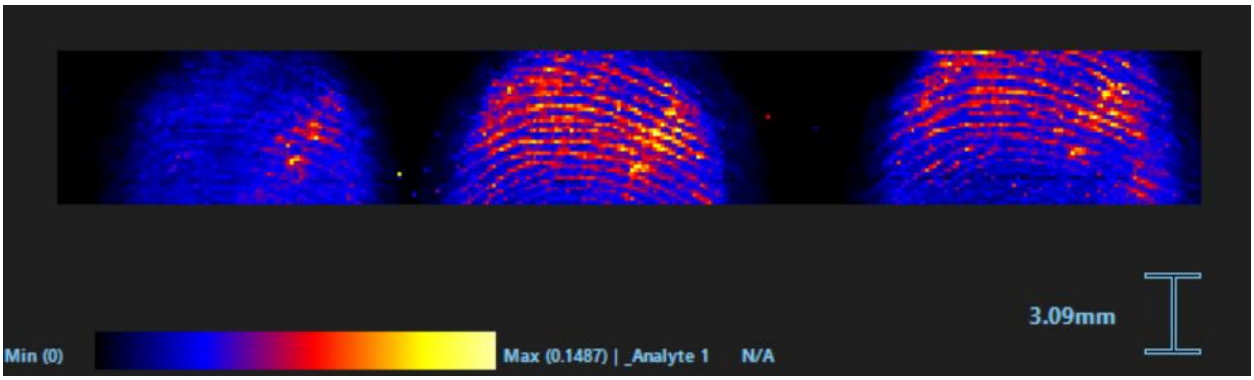


Figure 10 European Female 4 HDI Image. This shows a general idea of what fingerprints analyzed using MSI will look like. The ridges are not well defined, and their precise endings are unclear using these specific parameters for analysis. Increasing the sharpness of the image is possible by increasing the number of pixels at each scan point, at the cost of an increased scan time. The brighter the color of the print, the greater the concentration at that specific M/Z.

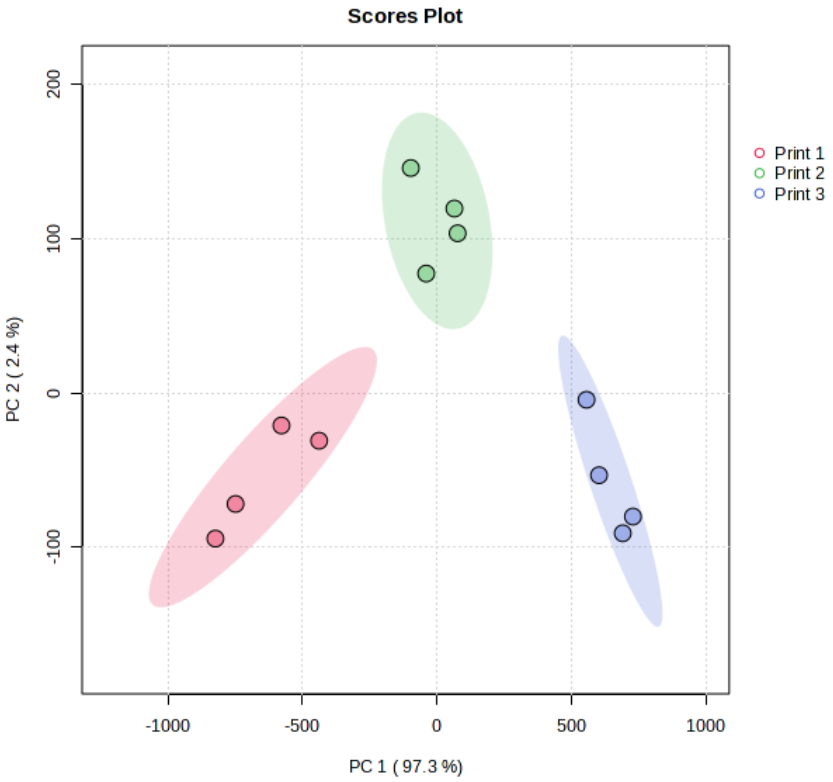


Figure 11 Hispanic Female 6. From the clusters in this group, it can be seen that there is no overlap between these three prints. The variance is totaled at 99.7%. From this graphic alone, it would not be possible to determine if these three prints came from the same individual.

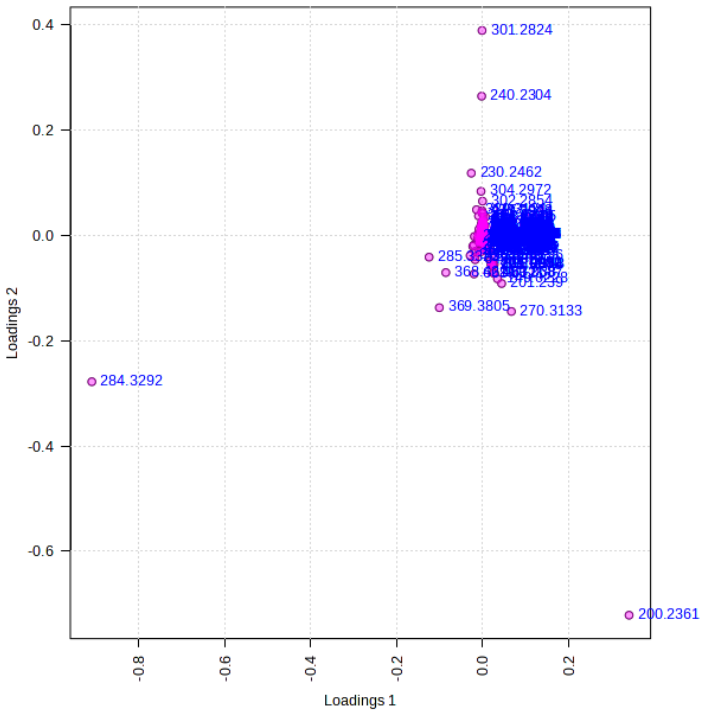


Figure 12 Loading Plot for Hispanic Female 6. Principal components 1 and 2 in Hispanic Female 6 was most influenced by 200.2361 and 301.2824, respectively.

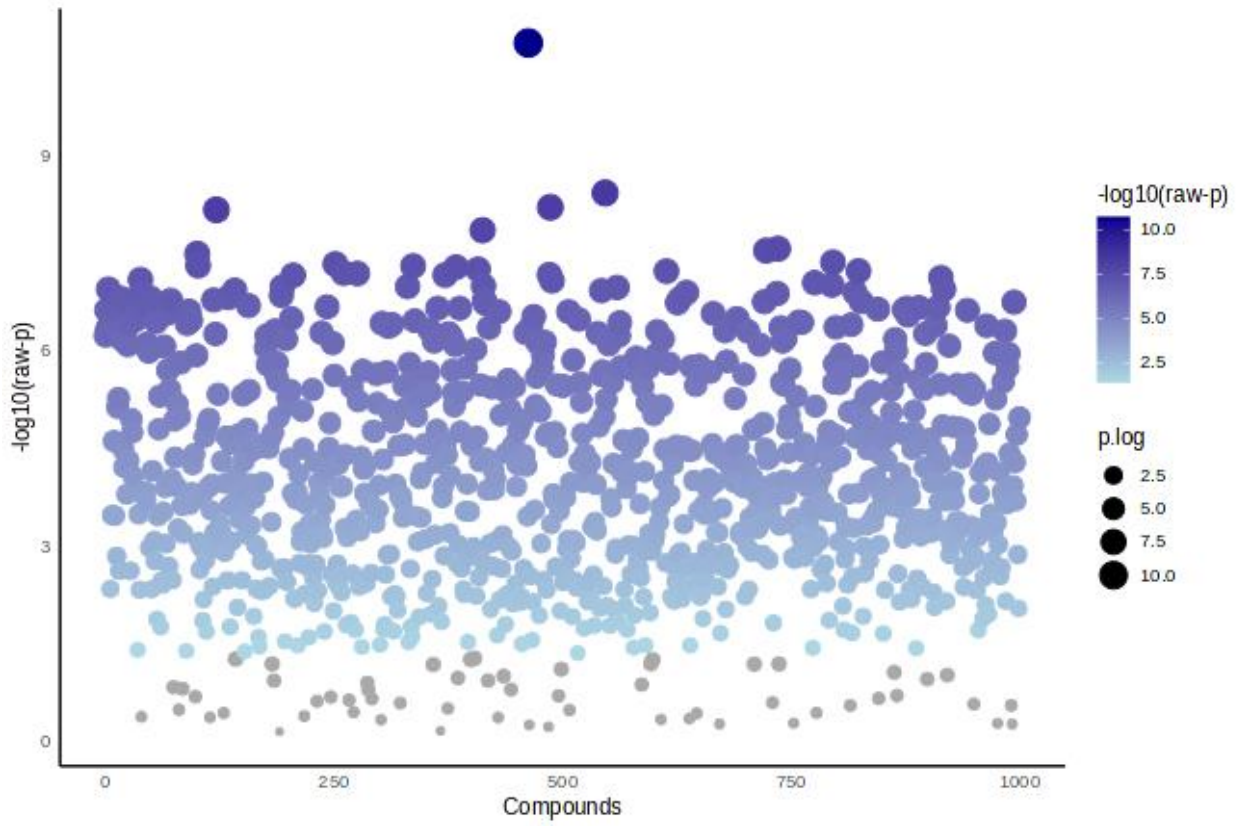


Figure 13 ANOVA for Hispanic Female 6. This depicts the most significant m/z value as determined by machine learning. This graphic shows a visual representation of the data found in Figure 14. No values can be assigned directly from this graphic, but a general understanding of the number of significant vs non-significant values can be observed. The darker the blue-purple color appearing in each dot, the more significant the value, while values shown in gray are nonsignificant for this dataset.

A1 m/z

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD													
2	673.3442	1080.7	1.9036e-1	10.72	1.9036e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
3	227.1004	330.05	3.7965e-0	8.4206	1.7196e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
4	239.0812	294.54	6.2903e-0	8.2013	1.7196e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
5	279.0954	288.66	6.8785e-0	8.1625	1.7196e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
6	727.4515	245.23	1.4152e-0	7.8492	2.8304e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
7	706.4362	211.56	2.7156e-0	7.5661	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
8	193.1198	208.49	2.8963e-0	7.5381	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
9	223.0923	203.18	3.2449e-0	7.4888	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
10	135.0436	189.7	4.3883e-0	7.3577	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
11	595.3793	187.08	4.6657e-0	7.3311	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
12	301.138	184.73	4.9316e-0	7.307	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
13	233.0688	184.13	5.0033e-0	7.3007	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
14	629.318	182.41	5.2134e-0	7.2829	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
15	175.1311	179.26	5.6274e-0	7.2497	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
16	190.1573	176.92	5.9615e-0	7.2246	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
17	585.289	176.6	6.0089e-0	7.2212	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
18	241.178	173.91	6.4279e-0	7.1919	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
19	425.2813	173.37	6.5169e-0	7.186	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													
20	165.0793	171.68	6.8017e-0	7.1674	3.3052e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2													

Figure 14 Hispanic Female 6 ANOVA Values Chart. M/Z can be viewed in the far-left hand column in order of decreasing significance. From this chart, it can be seen that the M/Z represented by 673.3442 is the most significant in this data set.

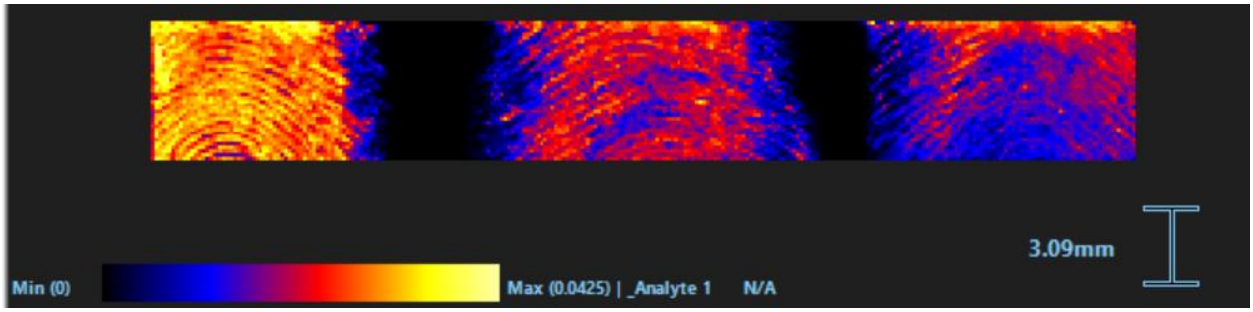


Figure 15 Hispanic Female 6 HDI Image. This shows a general idea of what fingerprints analyzed using MSI will look like. The ridges are not well defined, and their precise endings are unclear using these specific parameters for analysis. Increasing the sharpness of the image is possible by increasing the number of pixels at each scan point, at the cost of an increased scan time. The brighter the color of the print, the greater the concentration at that specific M/Z.

Data for Individuals with the same Geographic and Cultural Ancestry

Following the analysis of individual data, individuals from the same geographic and cultural ancestry were analyzed as a group. In order to make this data into a usable format, the data had to be aligned using Geena 2 technology. For this stage of the analysis, only one print from each individual was used, meaning only 4 ROIs of the obtained 12 for each individual were present during each analysis. PCA graphs for each set of prints were obtained.

Again, the principal components account for a high degree of variability in both of these graphs. When looking at the data in Figure 16, a cluster can be viewed in the lower left area, as well as two data sets that appear to be outliers in comparison to the other data. In this graphic, the principal components account for about 71.7% of the variance. In Figure 20, the data is much more clustered in the center, though most individuals have their own area in which their data can be found. In this graphic, it can be seen that 80.3% of the variance data is explained by the principal components. When we view these total variances, we see that they are lower than those discussed for the single individuals discussed previously (Hispanic Female 6 and European Female 4). This is somewhat expected as we are now viewing different individuals rather than the same individual. However, while these total variances are lower than those of the single individual, they are still

within an acceptable range. In this case, it would just take more principal components to account for more variance. The loading plots for European Print 1 show that the m/z values of 284.326, 200.234, and 368.417 had the most influence on the principal components while the Hispanic Print 1's principal components were most influenced by 200.234, 284.327, and 385.390. These values for the loading plots are the data points which have the most influence on the Principal Component Analysis and most contribute to the variance. ANOVA was used to show the most significant m/z for each PCA plot. For the individuals of European descent, the most significant m/z was 892.809 while for individuals of Hispanic descent, the most significant m/z was 960.83. If desired, these M/Z values could be analyzed further to determine the contributing species which is present and eventually allow for potential identification of an individual's background as this is the most prevalent species. The first print only was selected for analysis here as it allows for the greatest number of individuals to be represented as some participants only provided a single fingerprint. Additionally, some groups, such as Asian and African American descent individuals were not analyzed as there was only one participant for each of these categories. Graphics depicting this phenomenon can be found in Appendix A along with other data graphs for European and Hispanic Prints 2 and 3. Additional data for each print is discussed in the conclusion.

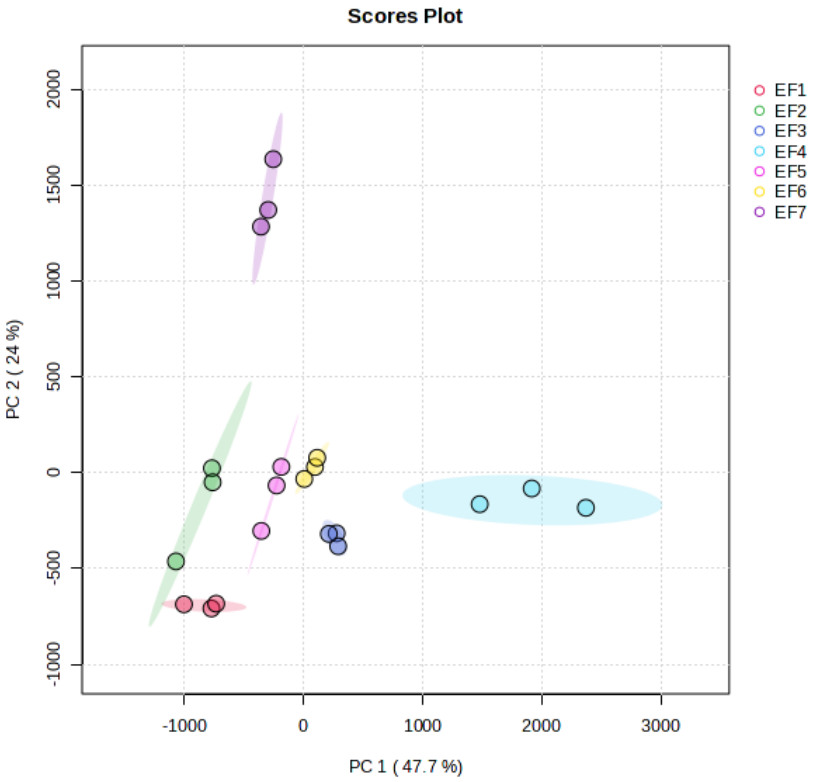


Figure 16 PCA for the first fingerprint taken from each European participant. Here there are not as tight of clusters as compared to other groups. There is a cluster in the lower left side of the graph, containing individuals 1, 2, 3, 5, and 6. Meanwhile individuals 4 and 7 are further out from this cluster. This means these two latter individuals are noted as being more variable when compared to the other individuals. The total variance is 71.1% for these two principal components.

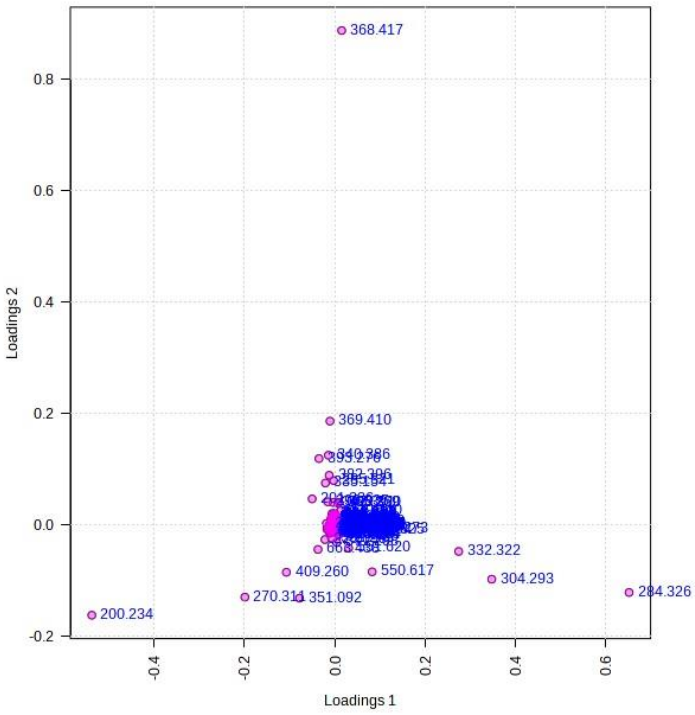


Figure 17 Loading Plot for European Print 1. The values 284.326, 200.234, and 368.417 had the most influence on the Principal Components for the European individuals. These values are identified as they are found at the far sides of the X-axis and Y-axis.

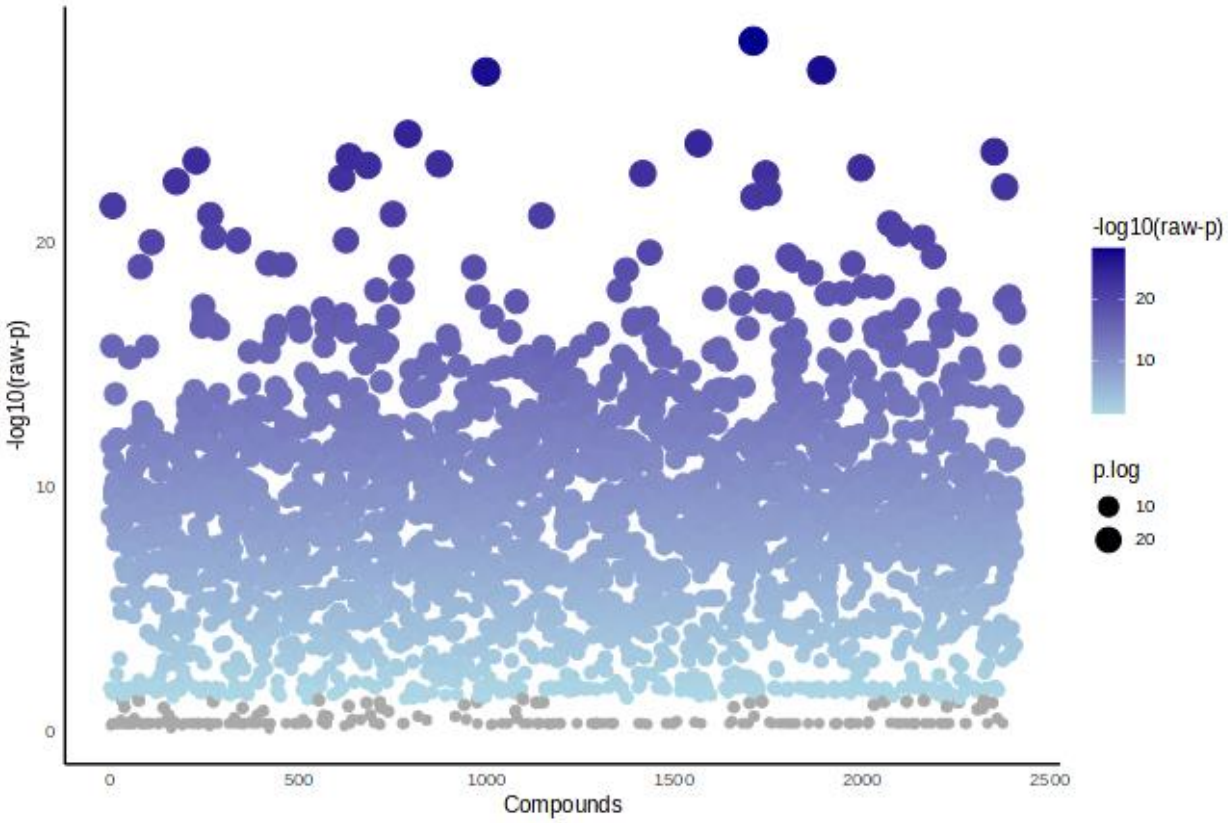


Figure 18 ANOVA For the first fingerprint taken from each European participant. This graphic shows a visual representation of the data found in Figure 19. No values can be assigned directly from this graphic, but a general understanding of the number of significant vs non-significant values can be observed. The darker the blue-purple color appearing in each dot, the more significant the value, while values shown in gray are nonsignificant for this dataset.

A1 m/z

	A	B	C	D	E	F
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD (top 1000)
2	892.809	41882	5.9942e-2	28.222	1.4434e-2	EF6 - EF1; EF6 - EF2; EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7
3	960.877	28249	9.436e-28	27.025	8.6783e-2	EF6 - EF1; EF6 - EF2; EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7
4	624.065	27705	1.0812e-2	26.966	8.6783e-2	EF1 - EF2; EF1 - EF3; EF1 - EF4; EF1 - EF5; EF1 - EF6; EF1 - EF7
5	522.284	12018	3.7373e-2	24.427	2.2498e-2	EF5 - EF1; EF5 - EF2; EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7
6	838.726	10559	9.2465e-2	24.034	4.4531e-2	EF3 - EF1; EF4 - EF1; EF3 - EF2; EF4 - EF2; EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7; EF4 - EF5; EF4 - EF6; EF4 - EF7
7	1174.899	9451.4	2.0079e-2	23.697	8.0583e-2	EF3 - EF1; EF3 - EF2; EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
8	428.258	8822.2	3.2515e-2	23.488	1.1185e-2	EF6 - EF1; EF6 - EF2; EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7
9	206.097	8367.5	4.7088e-2	23.327	1.4173e-2	EF2 - EF1; EF2 - EF3; EF2 - EF4; EF2 - EF5; EF2 - EF6; EF2 - EF7
10	566.345	8012.3	6.3789e-2	23.195	1.7067e-2	EF7 - EF1; EF7 - EF2; EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
11	451.31	7862.2	7.281e-24	23.138	1.7533e-2	EF7 - EF1; EF7 - EF2; EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
12	1006.802	7593.7	9.2853e-2	23.032	2.0326e-2	EF3 - EF1; EF3 - EF2; EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
13	778.446	7043.4	1.5719e-2	22.804	2.9953e-2	EF4 - EF1; EF4 - EF2; EF4 - EF3; EF4 - EF5; EF4 - EF6; EF4 - EF7
14	905.711	7014.9	1.6171e-2	22.791	2.9953e-2	EF3 - EF1; EF3 - EF2; EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
15	417.23	6631.5	2.3962e-2	22.62	4.1215e-2	EF5 - EF1; EF5 - EF2; EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7
16	176.467	6329	3.3221e-2	22.479	5.333e-21	EF1 - EF2; EF1 - EF3; EF1 - EF4; EF1 - EF5; EF1 - EF6; EF1 - EF7
17	1187.894	5878.9	5.566e-23	22.254	8.3768e-2	EF3 - EF1; EF3 - EF2; EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
18	908.334	5452	9.4329e-2	22.025	1.3361e-2	EF5 - EF1; EF5 - EF2; EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7
19	893.337	5144.1	1.4167e-2	21.849	1.8952e-2	EF5 - EF1; EF5 - EF2; EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7
20	102.127	4568.3	3.2503e-2	21.488	4.1194e-2	EF2 - EF1; EF2 - EF3; EF2 - EF4; EF2 - EF5; EF2 - EF6; EF2 - EF7

Figure 19 European Print 1 ANOVA Values Chart. M/z can be viewed in the far-left hand column in order of decreasing significance. From this table, it can be determined that the chemical substance represented by M/Z valued at 892.809 is the most significant value in this data set. Further analysis could be performed to determine what this substance may be.

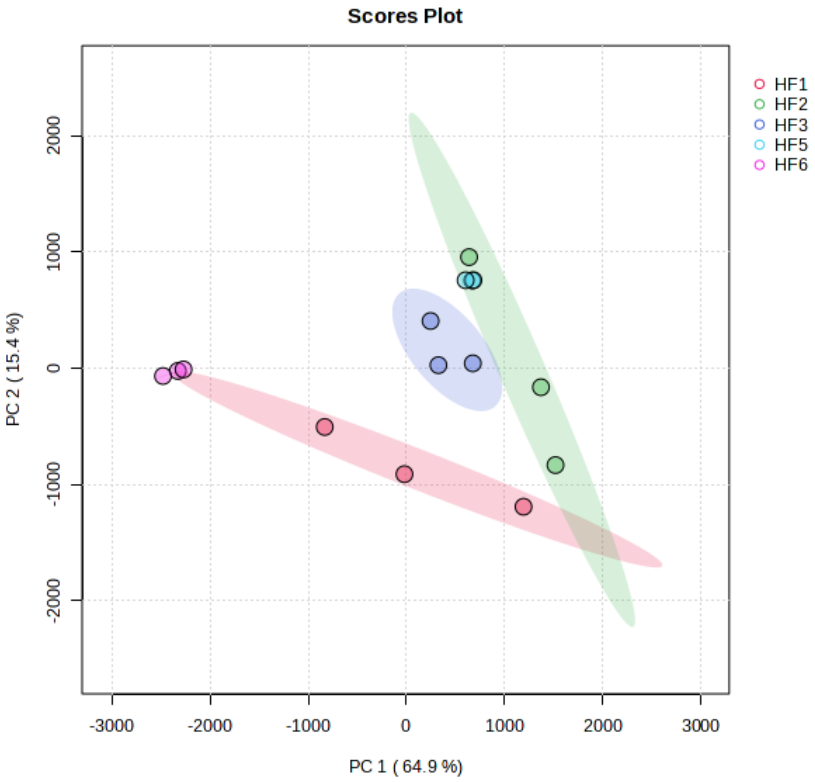


Figure 20 Principal Component Analysis for the first print obtained for all Hispanic individuals. This graphic allows for a quick overview of the data. It can be seen that some of the individuals (represented by a single-color circle) are more similar to certain individuals than others. For example, the tight blue cluster representing Hispanic Female 5, perfectly overlaps with Hispanic Female 2 (represented by green), while Hispanic Female 6 (shown in pink) does not overlap with the previously mentioned individuals at all. Additionally, the axis values for PC 1 and PC 2, show that overall, 80.3% of the overall variance in the data is represented.

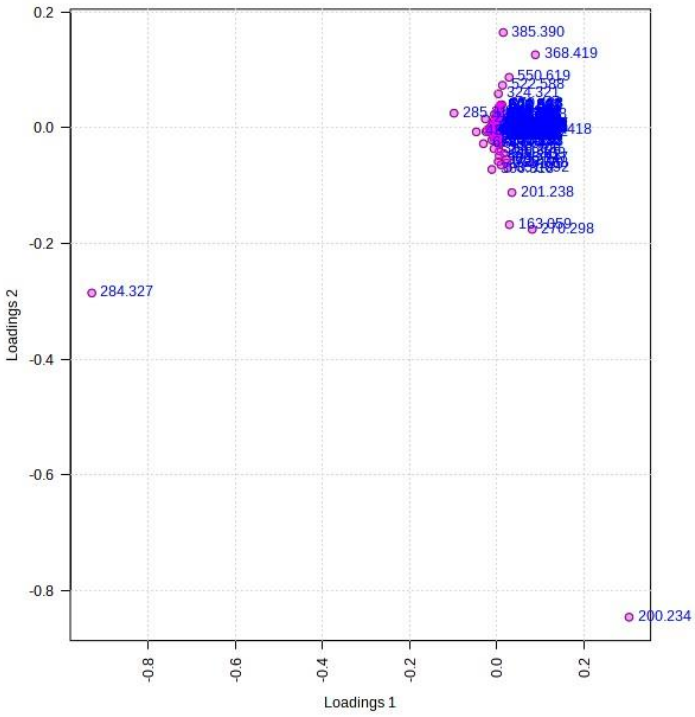


Figure 21 Loading Plot for Hispanic Print 1. The values located at the far sides of the X-axis and Y-axis (200.234, 284.327, and 385.390) contributed most to the Principal Component graphic.

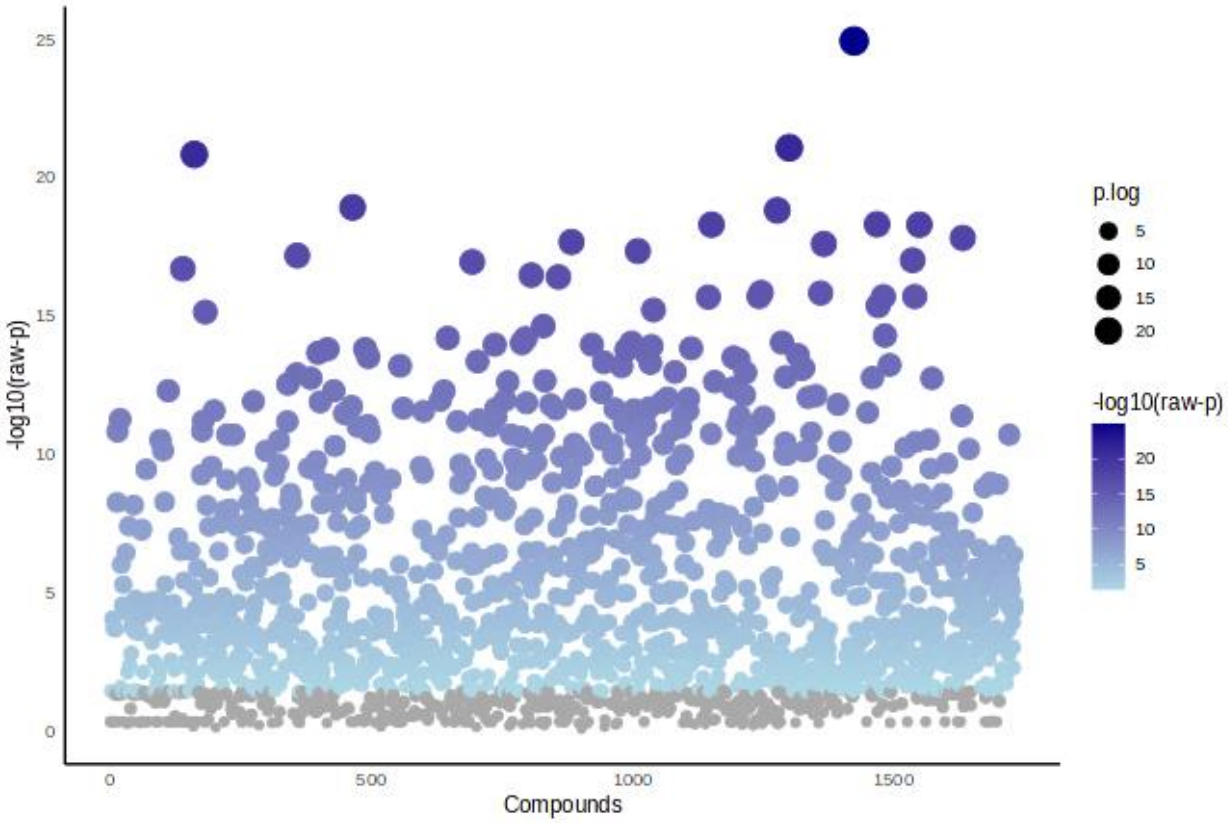


Figure 22 ANOVA For Hispanic Print 1. This graphic shows a visual representation of the data found in Figure 23. No values can be assigned directly from this graphic, but a general understanding of the number of significant vs non-significant values can be observed. The darker the blue-purple color appearing in each dot, the more significant the value, while values shown in gray are nonsignificant for this dataset.

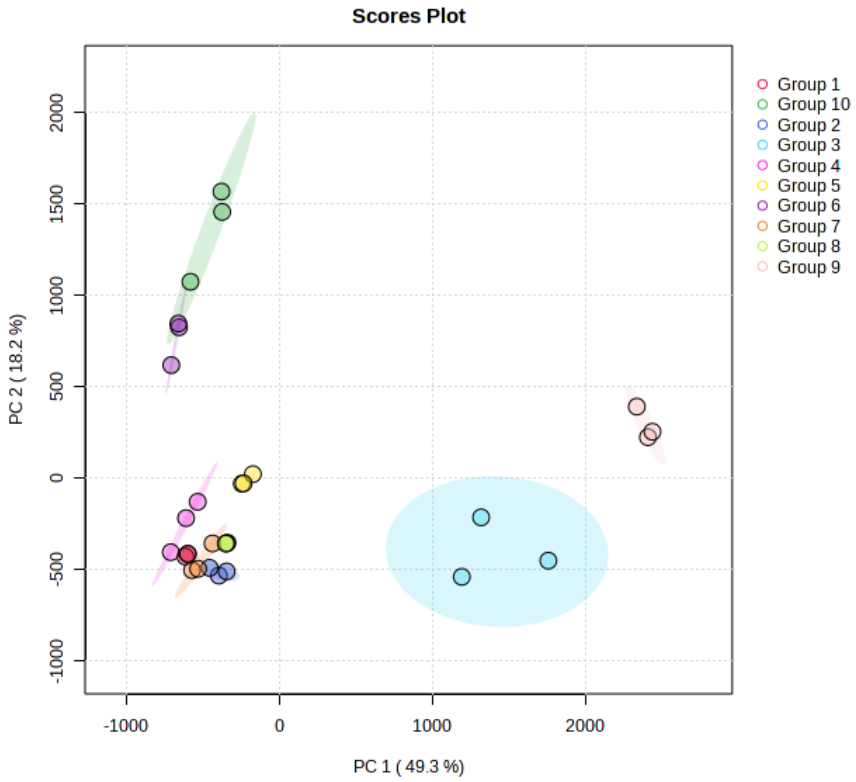
m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD (top 1000)
960.83	348100	1.1463e-2	24.941	1.9866e-2	HF6 - HF1; HF6 - HF2; HF6 - HF3; HF6 - HF5
882.603	58915	8.253e-22	21.083	7.1513e-1	HF5 - HF1; HF5 - HF2; HF5 - HF3; HF5 - HF6
231.25	52686	1.443e-21	20.841	8.3355e-1	HF6 - HF1; HF6 - HF2; HF6 - HF3; HF6 - HF5
460.487	21837	1.1791e-1	18.928	5.1086e-1	HF6 - HF1; HF6 - HF2; HF6 - HF3; HF6 - HF5
871.544	20786	1.5091e-1	18.821	5.2306e-1	HF5 - HF1; HF5 - HF2; HF5 - HF3; HF5 - HF6
988.831	16485	4.8081e-1	18.318	1.0775e-1	HF6 - HF1; HF6 - HF2; HF6 - HF3; HF6 - HF5
1040.901	16374	4.9742e-1	18.303	1.0775e-1	HF6 - HF1; HF6 - HF2; HF6 - HF3; HF6 - HF5
802.031	16374	4.9742e-1	18.303	1.0775e-1	HF5 - HF1; HF5 - HF2; HF5 - HF3; HF5 - HF6
1105.982	13132	1.4989e-1	17.824	2.8862e-1	HF6 - HF1; HF6 - HF2; HF6 - HF3; HF6 - HF5
681.996	12275	2.1004e-1	17.678	3.64e-16	HF5 - HF1; HF5 - HF2; HF5 - HF3; HF5 - HF6
922.553	11883	2.4697e-1	17.607	3.8909e-1	HF5 - HF1; HF5 - HF2; HF5 - HF3; HF5 - HF6
734.019	10551	4.475e-18	17.349	6.4626e-1	HF5 - HF1; HF5 - HF2; HF5 - HF3; HF5 - HF6
387.385	9800.7	6.4703e-1	17.189	8.6254e-1	HF5 - HF1; HF5 - HF2; HF5 - HF3; HF5 - HF6
1032.859	9064.5	9.5596e-1	17.02	1.1833e-1	HF6 - HF1; HF6 - HF2; HF6 - HF3; HF6 - HF5
597.376	8809.8	1.1023e-1	16.958	1.2735e-1	HF2 - HF1; HF2 - HF3; HF2 - HF5; HF2 - HF6
214.213	7861.6	1.9476e-1	16.711	2.1095e-1	HF1 - HF2; HF1 - HF3; HF1 - HF5; HF1 - HF6
647.33	7069.3	3.3119e-1	16.48	3.3762e-1	HF6 - HF1; HF6 - HF2; HF6 - HF3; HF6 - HF5
670.425	6858.3	3.8533e-1	16.414	3.7099e-1	HF5 - HF1; HF6 - HF1; HF5 - HF2; HF6 - HF2; HF5 - HF3; HF6 - HF3; HF5 - HF6
851.551	5295.6	1.4031e-1	15.853	1.2729e-1	HF3 - HF1; HF3 - HF2; HF3 - HF5; HF3 - HF6

Figure 23 Hispanic Print 1 ANOVA Values Chart. M/z can be viewed in the far-left hand column in order of decreasing significance. From this table it can be seen that the M/Z value of 960.83 has the most significance in this dataset for the first fingerprints of Hispanic individuals.

Comparing Different Geographic and Cultural Ancestries

The third and final stage of analysis focused on individuals of different geographic and cultural ancestries.

In this final data set, analysis included data from all individuals. It can be seen that Groups 1, 2, 4, 5, 7, and 8 are all clustered closely together in the lower left corner. These groups represent the individuals A1, E4, E5, E6, H3, and H5. Groups 10, 2, and 6 also created a cluster, representing ME1, E3, and E7. Finally, Group 3 and Group 9 are shown as being apart from either cluster, representing E4 and H6. This shows that there is some agreement between individuals of the same geographic and cultural ancestry, while also showing that there are some similarities among individuals of different geographic and cultural ancestries. In this large data set, the principal components were most influenced by 284.236, 368.417, and 200.234. These values as previously stated are those which contribute most to the variance within the principal components. An attempt was made to reanalyze the data following initial analysis in order to obtain ANOVA information for the dataset, however, an error occurred with MetaboAnalyst, and the data could not be obtained. This data is discussed here as all participants are represented in this data due to some participants only providing one fingerprint for analysis. Additional graphics showing all of the participants in the study can be found in Appendix A, and all data is discussed in the conclusion section.



Key:
Group 1 – A1
Group 10 – ME1
Group 2 – E3
Group 3 – E4
Group 4 – E5
Group 5 – E6
Group 6 – E7
Group 7 – H3
Group 8 – H5
Group 9 – H6

Figure 24 All Geographical and Cultural Ancestries Print 2. It can be observed that 1, 2, 4, 5, 7, and 8 are all clustered closely together in the lower left corner while Groups 10, 2, and 6 create a cluster in the upper left side. Groups 3 and 9 are not near the other clusters, but do not make a cluster of their own either. From this graphic, it is not possible to differentiate groups based on their ancestry completely. While some individuals such as Group 3, 4 and 5 (all individuals of European descent) occur within the same group, this cluster also contains individuals of Asian and Hispanic ancestry.

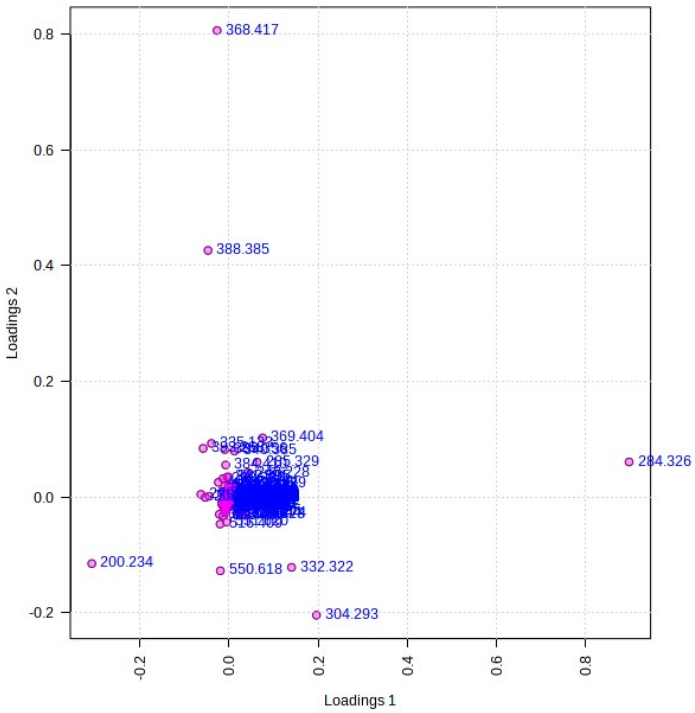


Figure 25 Loading Plot for All Geographical and Cultural Ancestries Print 2. From this graphic it can be seen that the values of 284.236, 368.417, and 200.234 are contributing most to the values shown in the PCA. This can be determined as they are the values furthest to the outside of the X-axis and Y-axis in both the positive and negative directions.

Discussion

As this type of information has never been studied before, there are three separate types of criteria that must be evaluated in order to determine the validity of the work and the need for future studies. The first criteria that must be demonstrated is the correlation of fingerprints within a single individual. It can be seen in *Figure 8*, there is a great deal of chemical overlap in this European Female 4's latent fingerprints. When looking at the first principal component, it can be seen that the variance is 82.9%, while the second principal component is 16.2%. The difference between these two principal components (66.7) contributes to the fact that these two principal components have a significant degree of overlap. In comparison, *Figure 11* (the fingerprint of Hispanic female 6) has significantly less overlap. As shown on the horizontal axis, the first principal component accounts for 97.3% of the variance of the sample. This means that it individually accounts for the greatest degree a single variable will provide. In comparison, the second principal component only contributes 2.4% to the overall variance. When comparing *Figure* 8 and *Figure* 11 directly, it can be determined that the more variance present in the first principal component, the less overlap will exist. As this was an untargeted study, a greater degree of variance is expected to exist. That is to say, if a specific component of study, such as lipids, would have been identified, a lesser degree of variability may exist. An untargeted study was conducted to identify if any broad chemical patterns existed between geographic and cultural ancestry and latent fingerprint chemistry. A targeted study would require the exact identification of a compound and require a control to compare an analyzed sample to.

The second criteria that we must establish is a similarity in the chemical residues of individuals of the same ancestral background. When viewing the graphics containing the data for all European and Hispanic individuals, we can see that in some cases we have overlap, meaning

that similarities exist, while other graphics depict all the latent fingerprint residues to have chemistry that does not contain any similarities. In *Figure 16*, for example, it can be seen that each European individual has their own individual location which does not intersect with any of the other European descent participants. While a lower total variance exists (71.1%) it is still within an acceptable range. An addition of more of the principal components could provide a greater degree of overlap as the variance decreases with each subsequent principal component. In comparison, graphics like *Figure 20*, show a greater degree of overlap, though not a significant amount. While only a few of the individuals exist in the same space (such as H2 and H4), others differ greatly. Overall, the data for all the groups shown in *Figure 20* exist in a similar space as none are significant outliers. Due to this, it cannot be ascertained if there are similarities when comparing individuals of the same ancestral background. Only the Hispanic and European individuals were analyzed at this stage as the other ancestral backgrounds (African American, Asian, and Middle Eastern) only contained one individual, meaning no comparisons could be made. Increasing the number of principal components could show a greater agreement between individuals of the same geographic and cultural ancestry, as well as increasing the number of participants in the study.

The final criteria that is essential for this type of analysis is the existence of varying latent fingerprint chemistries among individuals of different ancestral backgrounds. As can be seen in the PCA graphs found above, as well as in Appendix A, it is not possible to determine if there are significant differences between individuals of different geographic and cultural ancestries. This is evident when viewing the PCA plots. If significant differences were to exist among individuals of different geographic and cultural ancestries clusters would be seen which would have all of the European individuals together as well as all the Hispanic individuals together while the outlier

geographic and cultural ancestries, African (represented by B), Asian (represented by A) and Middle Eastern (represented by ME) would exist separately from these clustered groups. Instead, it can be seen that some individuals of the same geographic and cultural ancestry exist clustered together while also being intermingled with individuals of different geographic and cultural ancestries. It can be seen in *Figure 24* that the total variance of the two principal components represented in this graphic are only 67.5%, slightly lower than the accepted 70% variance. It is possible that part of this may be due to the small sample size. With an increase in the number of participants, it is possible that the corresponding ancestral groups would become more clustered and the expected differences among ancestral groups would become apparent. As mentioned previously, the use of a targeted study could produce a more reliable picture of the potential differences between these groups. From this particular study, there is not sufficient evidence to say that distinct differences exist among varying ancestral backgrounds.

Additionally, outside of the obtained chemical data and data analysis, it has been shown that imaging mass spectrometry, specifically DESI-MS could greatly benefit a latent fingerprint examiner. This type of data acquisition has shown to be beneficial when using a glass substrate as the instrument allows the analyst to obtain an image of the latent fingerprint without disturbing the print or requiring the addition of chemicals or powders in order to visualize the latent fingerprint. Using imaging mass spectrometry would allow an analyst to obtain a digital image of a latent fingerprint which can immediately be processed on a computer. Using DESI-MS can be beneficial as the non-destructive technique perimeters are set by the user. If an initial scan is performed and a sufficient scan is not obtained, a secondary scan can be performed in order to obtain a better image. The pixel rate is chosen as well. This allows the analyst to increase or decrease the sharpness of the image depending on the requirements. It should be noted that a sharper image

requires more pixels and will increase the scan time. However, if the goal of the analyst is to obtain chemical data, a lower pixel rate can be used. The scan rate is always set to be double that of the pixels, meaning that accurate data can be obtained as two scans are occurring at every pixel and being averaged together. Additionally, the HDI application allows the user to adjust the image in order to obtain the best image quality simply by sliding a digital scale.

While it is not possible to determine from this study whether geographic and cultural ancestry can be determined from an individual's latent fingerprint, a protocol for this type of study has been determined and has been found to be valid. Based on the results of this study no definitive conclusions can be made about an individual's geographic and cultural ancestry based on the chemical information obtained from their fingerprints. However, the methodology established in this study does allow for chemical information to be obtained from an individual's latent fingerprints. The potential for identification of an individual's geographic and cultural ancestry background from latent fingerprints would allow for investigators to include or exclude individuals on a list of suspects and further research should be conducted. This would assist forensic scientists as it would assist in creating more evidence with which to identify an individual who may have committed a crime.

Limitations

There are several limitations in this study. The most important thing to note is the relatively small sample size. Due to the cost of performing a single scan, as well as time constraints, a small sample size was utilized. Further study would warrant a larger sample size in order to increase the significance of the study. As a result of the small sample size, there were other limitations in regard to the participants in the study. In most cases, the participants in this study were female, thereby only representing one gender within a geographic and cultural ancestry. While one participant was male, there is not enough gender diversity to represent the whole population.

Limitations were also placed on the ancestral background of participants. The geographic and cultural ancestries were selected prior to sample collection, meaning that if individuals did not hear their background during recruitment, they may have not deemed themselves to be valid for the study. Further investigation would benefit from expanding the ancestral backgrounds which are studied. It may also be beneficial to study regionality as well. Since individuals who live in different areas may have different diets, it may be possible that an individual of African descent living on an African diet may have a different chemistry than an individual of African descent living in the United States. Additionally, participants were asked to self- identify their geographic and cultural background. While this information served the needs of the study, future studies would be aided by using ancestry testing to help increase accuracy of the information presented.

Another major limitation of this study was the sample substrate. All prints analyzed in this study were taken on glass slides which were clean and completely sterile. In the world of forensics, glass is not often the only object which fingerprints are found upon. More research would be needed to identify whether other backgrounds, as well as additional chemical components from

that background, would have any effect on the chemical data obtained from an individual's latent fingerprint chemistry.

In this study, a specific chemical component in the latent prints was not focused on for detection. Initially it was planned to focus on lipids, however, sample degradation during transportation did not allow for a proper lipid profile to be obtained. The data obtained from the lipid profile was not at a high enough concentration to perform data analysis. Instead of focusing on a compound such as fats or lipids, an overall chemical profile was obtained. Future investigation could benefit from isolating specific compounds in order to increase the value of the data obtained. Finally, due to the time required to travel to use instrumentation, not all prints were scanned within the same amount of time following acquisition. This additionally would add to degradation as a freezer at a low enough temperature to properly store samples was not available during experimentation and transport. Due to this, it is possible that some differences in chemistry could be the result of evaporation of the sample as prints were stored at room temperature in containers with lids.

Future Research

There are multiple ways in which this study could be continued in the future. The first, and potentially most important, would be to have a larger sample size. This would help to establish the true connection between latent fingerprint chemistry and an individual's geographic and cultural background. It is also recommended that some type of genealogical testing be utilized to increase the accuracy of results so the study would not be based on self-identification of geographic and cultural ancestry. This would also be helpful to gauge how individuals of mixed backgrounds fit into the overall study in comparison to those who are not of mixed backgrounds. If a difference in the chemical residues of individuals of different geographic and cultural ancestries could be established, it would be helpful to determine what chemicals in latent fingerprints are responsible for that difference. Additionally, a time-course study could help to establish if there is any change in fingerprint chemistry the longer a print remains on a surface. As of right now, DESI technology is only established to be performed on glass surfaces. In the future, it would be beneficial to develop a way for this technology to be performed on other types of surfaces, such as plastic or metals. If this type of technology were developed, all of the previously mentioned studies would need to be performed again on these new surfaces in order to establish whether or not the information previously discovered applied to all surfaces. Finally, it would be beneficial to investigate the ability to obtain all the characteristics of an individual from latent print residues (i.e., geographic, and cultural ancestry, gender, age, etc.) using a single methodology and instrumental analysis approach. This type of study, however, would require establishing the exact m/z at which these distinctive characteristics can be differentiated.

Resources

Brooks, P. J., Enoch, M. A., Goldman, D., Li, T. K., & Yokoyama, A. (2009). The alcohol flushing response: An unrecognized risk factor for esophageal cancer from alcohol consumption. *PLoS Medicine*, *6*(3), 0258–0263. https://doi.org/10.1371/journal.pmed.1000050

Bui, Y. (2009). *How to Write a Master's Thesis* (V. Knight, S. Connelly, & T. Bigelow, Eds.). SAGE Publications.

Constitutional Rights Foundation. (n.d.). *How Reliable Are Eyewitnesses?* Retrieved November 18, 2021, from https://www.crf-usa.org/bill-of-rights-in-action/bria-13-3-c-how-reliable-are-eyewitnesses

Croxton, R. S., Baron, M. G., Butler, D., Kent, T., & Sears, V. G. (2010). Variation in amino acid and lipid composition of latent fingerprints. *Forensic Science International, 199*(1–3), 93–102. https://doi.org/10.1016/j.forsciint.2010.03.019

de Hoffmann, E., & Stroobant, V. (2007). *Mass Spectrometry Principles and Applications*.

de la, R. (2007). *Shape Variability and Classification of Human Hair: A Worldwide Approach*.

Dwairi, M. O., Hindi, A., & Alqadi, Z. (2020). *Analysis of Procedures used to build an Optimal Fingerprint Recognition System*. https://www.researchgate.net/publication/339285960

Flowers, P., Theopold, K. (Klaus H., Langley, R., Robinson, W. R., Blaser, M., Bott, S. (Simon G., Carpenetti, D., Eklund, A., El-Giar, E., Frantz, D. (Donald D., Hooker, P., Kaminski, G., Look, J., Martinez, C., Milliken, T., Moravec, V., Powell, J. D., Sorensen, T., Soult, A., … Open Textbook Library. (n.d.). *Chemistry*.

Ford, J. R. (2020). *Fingerprints and Ancestry*.

Gál, L., Oravec, M., Gemeiner, P., & Ceppan, M. (2015). Principal component analysis for the forensic discrimination of black inkjet inks based on the Vis-NIR fibre optics reflection spectra. *Forensic Science International*, *257*, 285–292. https://doi.org/10.1016/j.forsciint.2015.09.011

Gross, J. H. (2017). *Mass Spectrometry*.

Hinners, P., O'Neill, K. C., & Lee, Y. J. (2018). Revealing Individual Lifestyles through Mass Spectrometry Imaging of Chemical Compounds in Fingerprints. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-23544-7

Hinners, P., Thomas, M., & Lee, Y. J. (2020). Determining Fingerprint Age with Mass Spectrometry Imaging via Ozonolysis of Triacylglycerols. *Analytical Chemistry*, *92*(4), 3125–3132. https://doi.org/10.1021/acs.analchem.9b04765

Holder, E. H., & Laub, J. H. (2012). *THE FINGERPRINT SOURCEBOOK*. www.nij.gov

Houck, M. M., & Siegel, J. A. (2010). *Fundamentals of Forensic Science* (Second Edition). Elsevier.

Huynh, C., Brunelle, E., Halámková, L., Agudelo, J., & Halámek, J. (2015). Forensic Identification of Gender from Fingerprints. *Analytical Chemistry*, *87*(22), 11531–11536. https://doi.org/10.1021/acs.analchem.5b03323

Jones, M. (2019, June 8). *Gas Chromatography-Mass Spectrometry*. American Chemical Society.

Krishan, K., Kanchan, T., & Garg, A. K. (2015). Dental Evidence in Forensic Identification-An Overview, Methodology and Present Status. In *The Open Dentistry Journal* (Vol. 9).

Lapides, R. A., & Savaiano, D. A. (2018). Gender, age, race, and lactose intolerance: is there evidence to support a differential symptom response? A scoping review. In *Nutrients* (Vol. 10, Issue 12). MDPI AG. https://doi.org/10.3390/nu10121956

M'charek, A., Toom, V., & Jong, L. (2020). The Trouble with Race in Forensic Identification. *Science Technology and Human Values*, *45*(5), 804–828. https://doi.org/10.1177/0162243919899467

National Human Genome Research Institute. (2021, July 22). *Y-Chromosome Infographic*.

Nayak, B. P., & Khajuria, H. (2019). Eyewitness testimony: probative value in criminal justice system. In *Egyptian Journal of Forensic Sciences* (Vol. 9, Issue 1). Springer Berlin Heidelberg. https://doi.org/10.1186/s41935-018-0109-z

O'Neill, K. C., Hinners, P., & Lee, Y. J. (2020). Potential of triacylglycerol profiles in latent fingerprints to reveal individual diet, exercise, or health information for forensic evidence. *Analytical Methods*, *12*(6), 792–798. https://doi.org/10.1039/c9ay02652e

Popa, G., Potorac, R., & Preda, N. (2010). Method for fingerprints age determination. *Romanian Journal of Legal Medicine*, *18*(2), 149–154. https://doi.org/10.4323/rjlm.2010.149

Riva, F., & Champod, C. (2014). Automatic comparison and evaluation of impressions left by a firearm on fired cartridge cases. *Journal of Forensic Sciences*, *59*(3), 637–647. https://doi.org/10.1111/1556-4029.12382

Romano, P., Profumo, A., Facchiano, A. (2018) Pre-processing MALDI/TOF mass spectra by using Geena 2. Current Protocols in Bioinformatics, 64(1).

Sewell, M. (2007). Principal Component Analysis. In *Journal of Educational Psychology* (Vol. 24, Issue 7). Springer-Verlag.

Takahashi, T. (2019). Unique hair properties that emerge from combinations of multiple races. *Cosmetics*, *6*(2). https://doi.org/10.3390/COSMETICS6020036

Wall, T. L., & Ehlers, C. L. (1995). *Alcohol Health and Research World Volume 19 Number 3 1995 The Genetics of Alcoholism*.

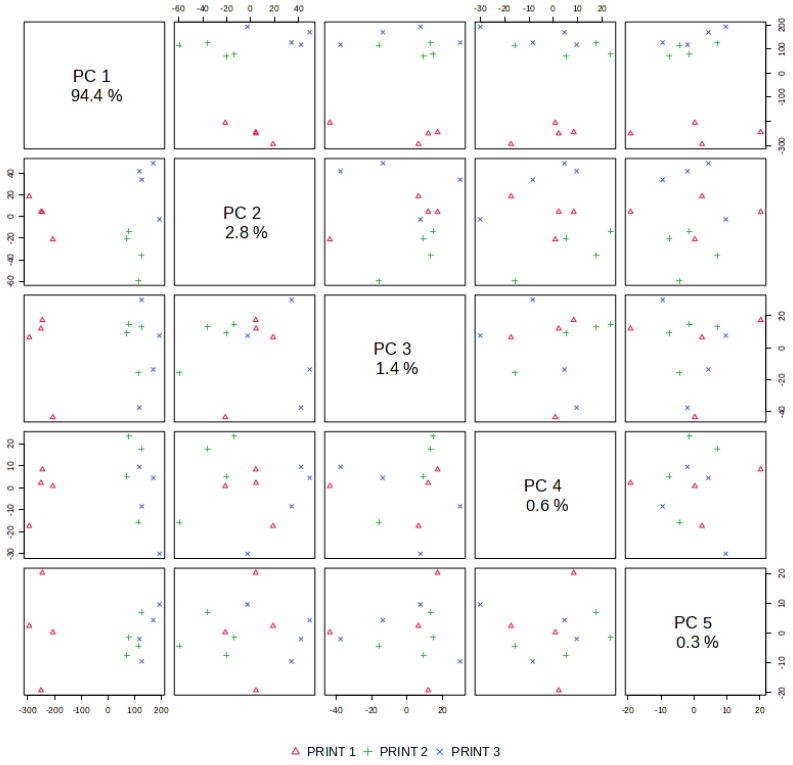
Wilson, J. P., Hugenberg, K., & Bernstein, M. J. (2013). The Cross-Race effect and eyewitness identification: How to improve recognition and reduce decision errors in eyewitness situations. *Social Issues and Policy Review*, *7*(1), 83–113. https://doi.org/10.1111/j.1751-2409.2012.01044.x

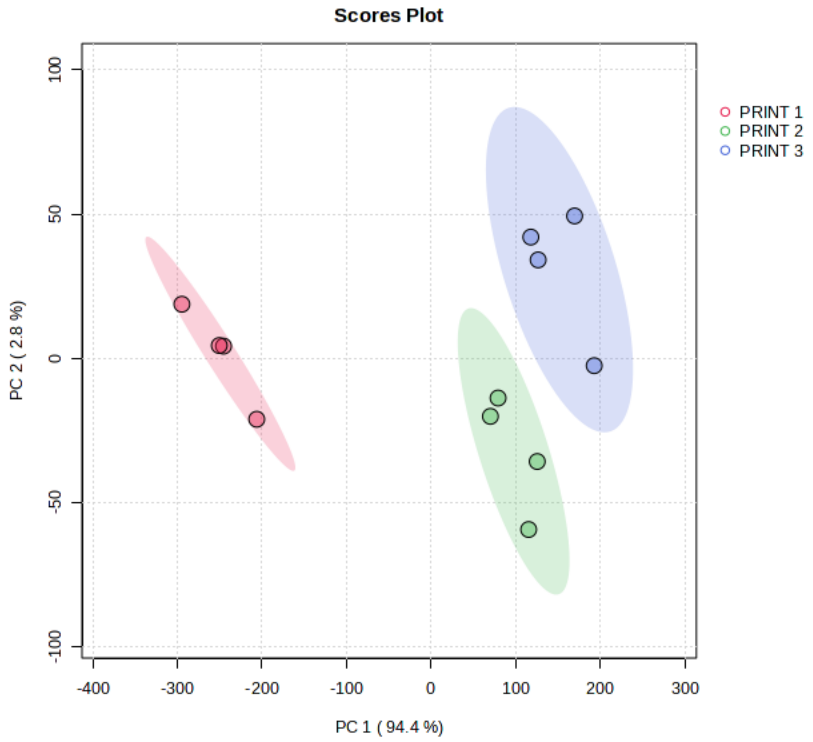
Zhou, Z., & Zare, R. N. (2017). Personal information from latent fingerprints using desorption electrospray ionization mass spectrometry and machine learning. *Analytical Chemistry*, *89*(2), 1369–1372. https://doi.org/10.1021/acs.analchem.6b04498

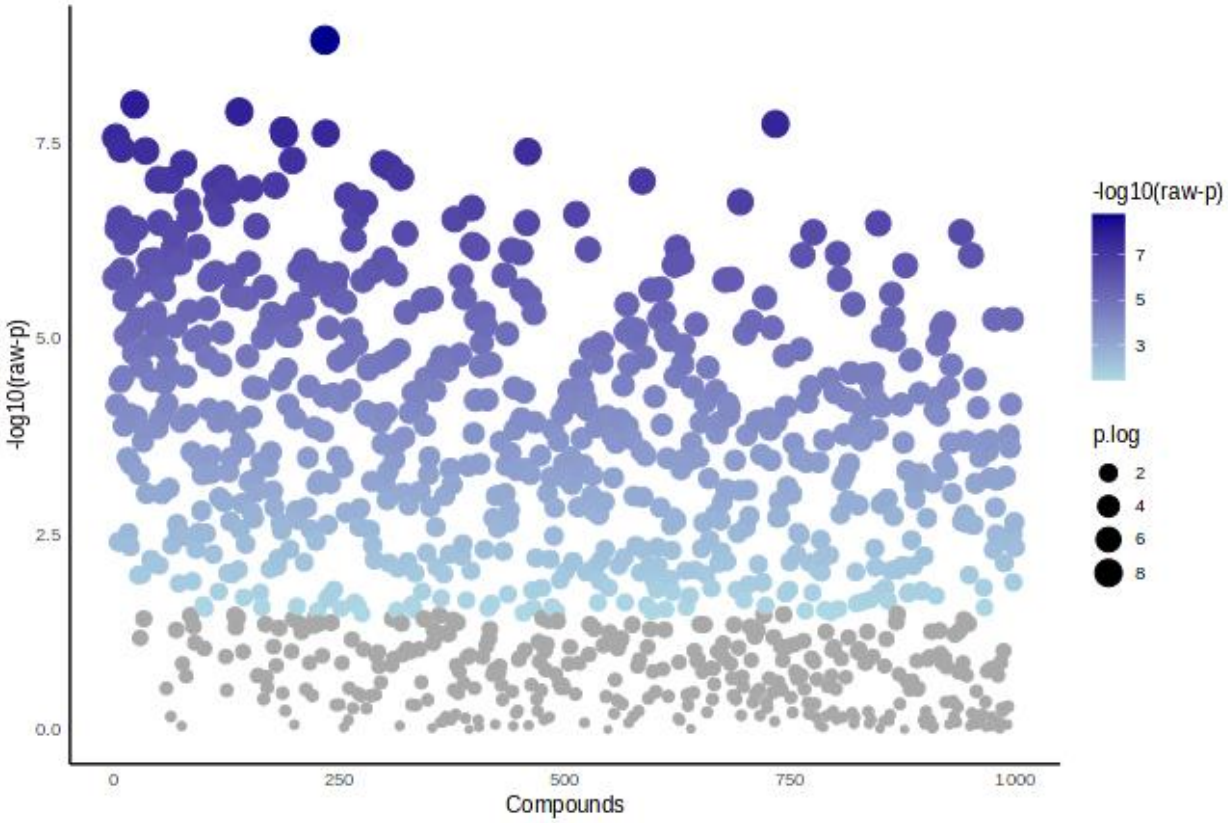
Appendix A

Individual Data Graphs

Asian Male 1 Data

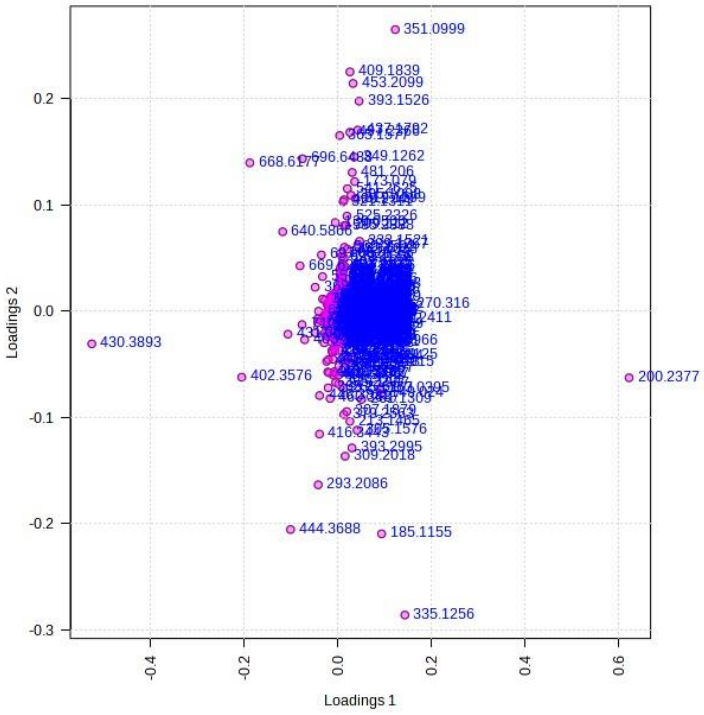



Scores Plot
PRINT 1
PRINT 2
PRINT 3
PC 2 (2.8 %)
PC 1 (94.4 %)
100
50
0
-50
-100
-400
-300
-200
-100
0
100
200
300

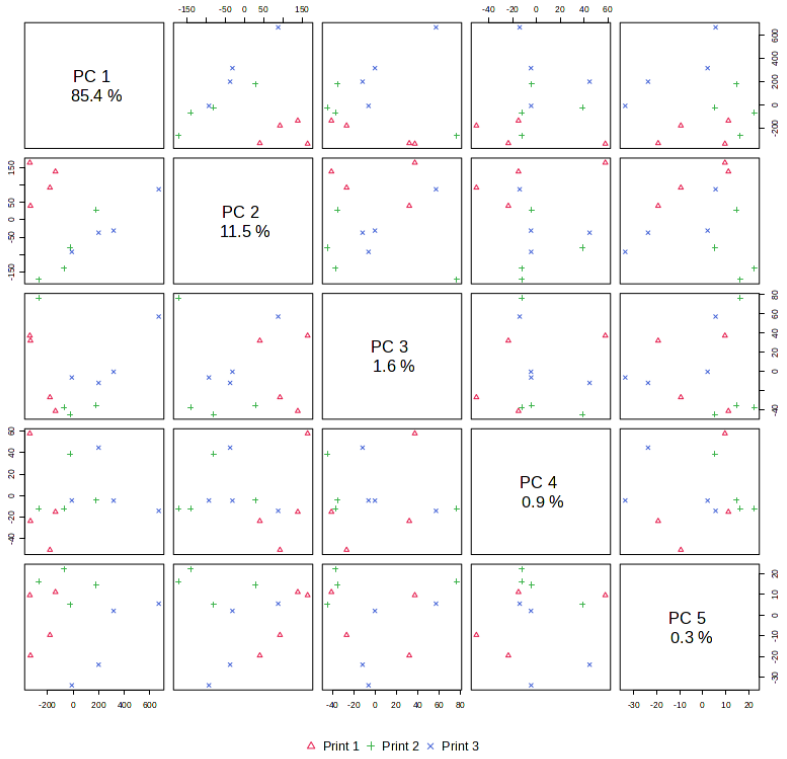

-log10(raw-p)
Compounds
7.5
5.0
2.5
0.0
0
250
500
750
1000
-log10(raw-p)
7
5
3
p.log
2
4
6
8

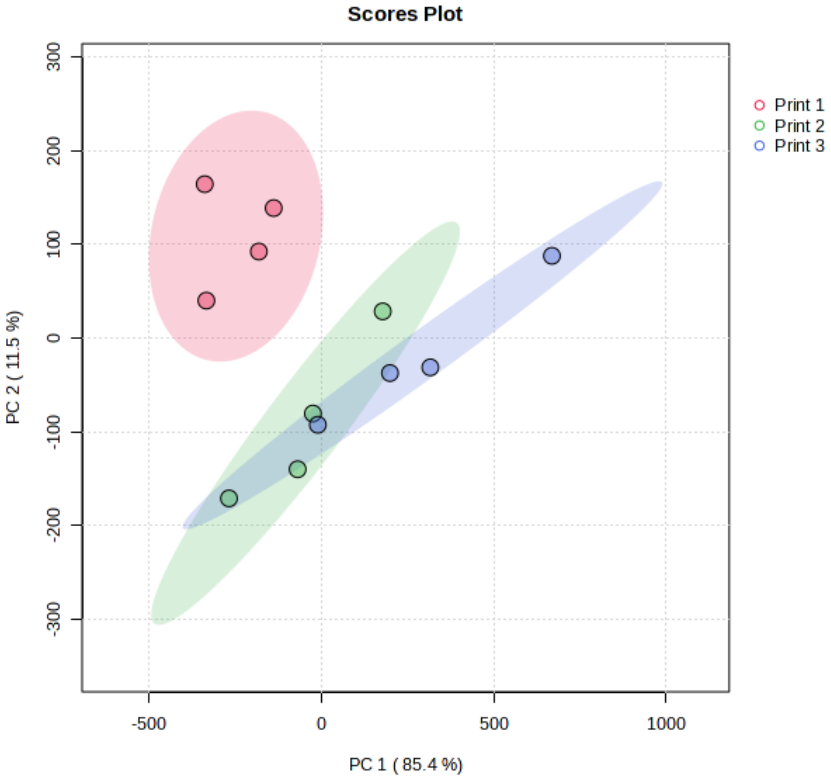
N13

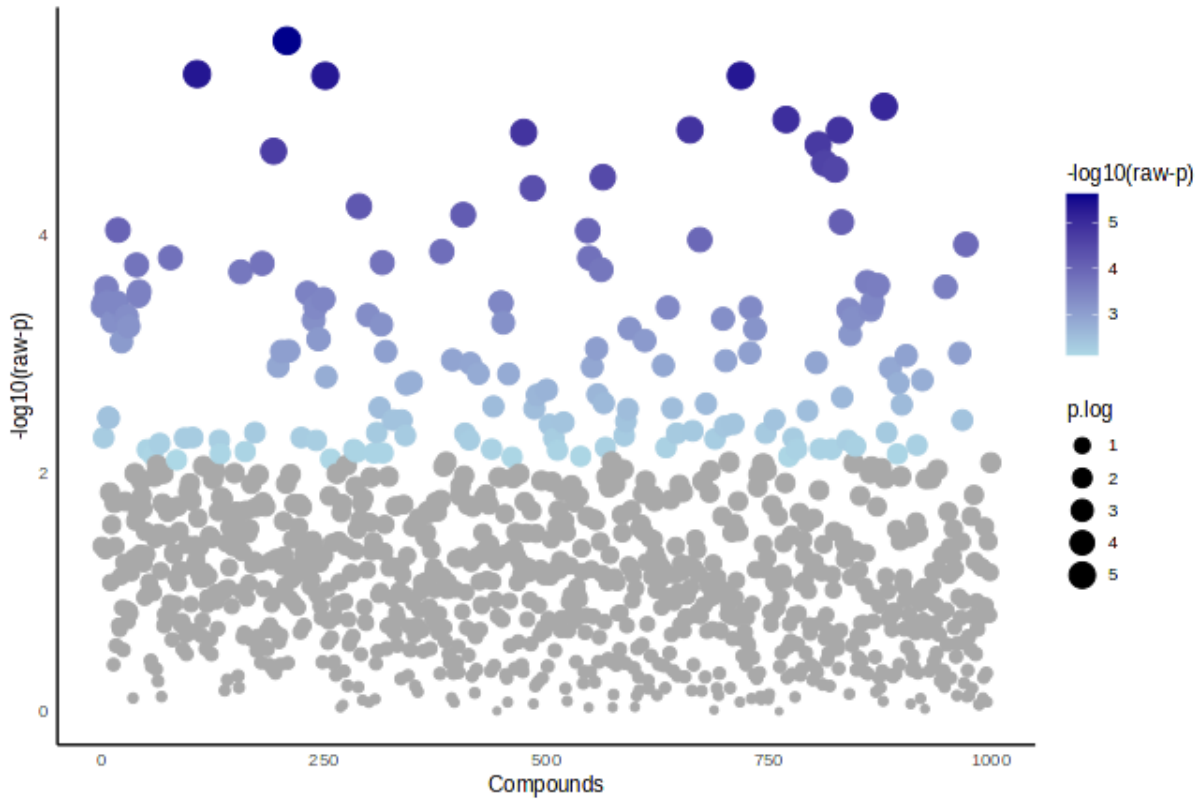
	A	B	C	D	E	F
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD
2	672.5723	403.92	1.547e-09	8.8105	1.547e-06	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
3	431.3919	263.49	1.0302e-0	7.9871	3.3988e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
4	428.3754	251.18	1.273e-08	7.8952	3.3988e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3
5	596.596	231.47	1.8265e-0	7.7384	3.3988e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3
6	730.5935	221.66	2.2108e-0	7.6554	3.3988e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
7	715.5927	217.45	2.406e-08	7.6187	3.3988e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
8	432.3918	216.17	2.4693e-0	7.6074	3.3988e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
9	430.3893	211.5	2.7191e-0	7.5656	3.3988e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
10	402.3576	196.1	3.7931e-0	7.421	3.7509e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
11	684.6111	193.37	4.0343e-0	7.3942	3.7509e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
12	688.5635	192.38	4.126e-08	7.3845	3.7509e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3
13	686.5745	181.37	5.3463e-0	7.2719	4.2122e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
14	758.6132	177.44	5.8849e-0	7.2303	4.2122e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
15	714.5932	177.36	5.8971e-0	7.2294	4.2122e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3
16	386.3622	173.15	6.5519e-0	7.1836	4.3679e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3
17	337.1349	162.08	8.7527e-0	7.0579	4.9038e-0	PRINT 2 - PRINT 1; PRINT 3 - PRINT 1; PRINT 3 - PRINT 2
18	333.1521	161.09	8.9919e-0	7.0461	4.9038e-0	PRINT 2 - PRINT 1; PRINT 3 - PRINT 1; PRINT 3 - PRINT 2
19	261.1309	159.6	9.3653e-0	7.0285	4.9038e-0	PRINT 2 - PRINT 1; PRINT 3 - PRINT 1; PRINT 2 - PRINT 3
20	403.3609	159.05	9.5072e-0	7.0219	4.9038e-0	PRINT 1 - PRINT 2; PRINT 1 - PRINT 3; PRINT 2 - PRINT 3



European Female 3 Data

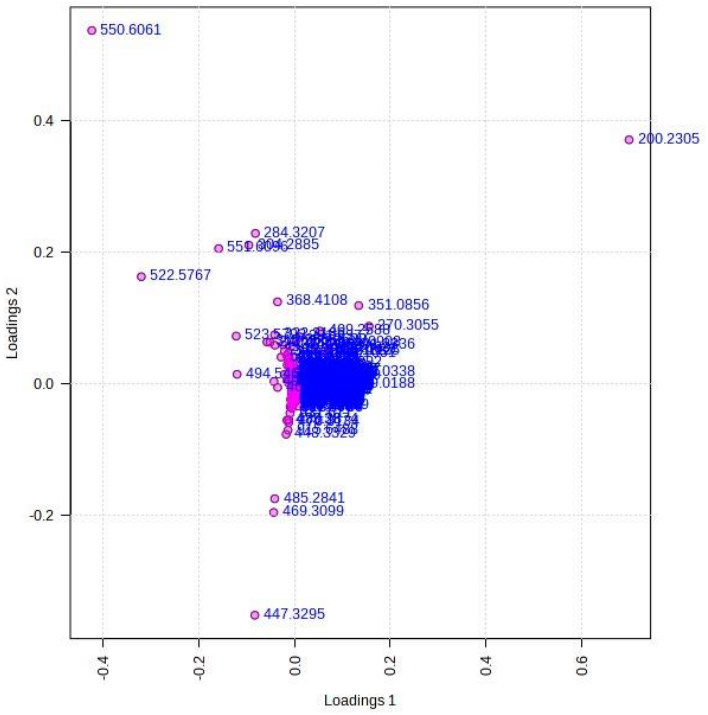




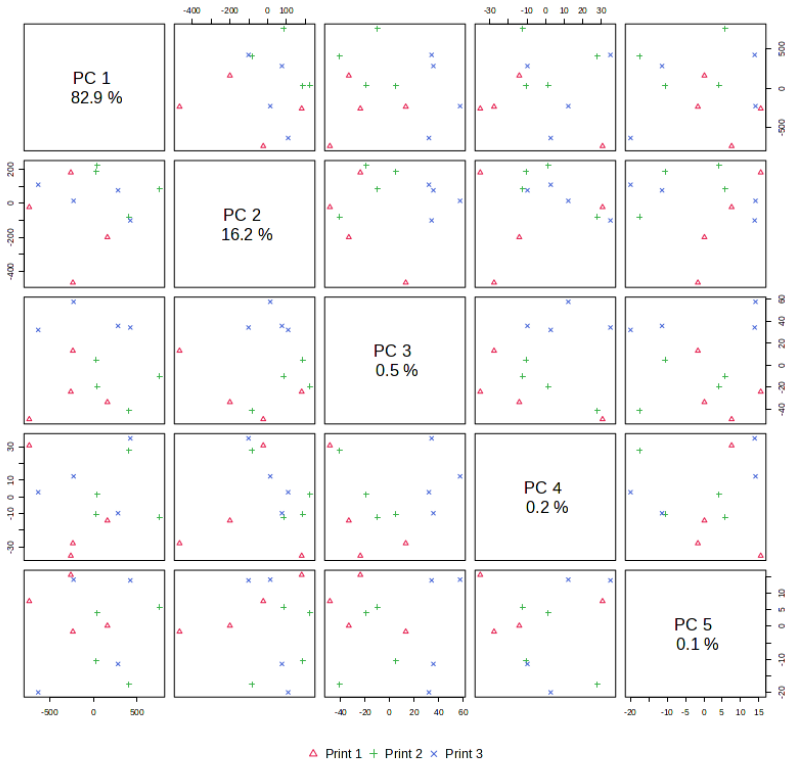


A1 m/z

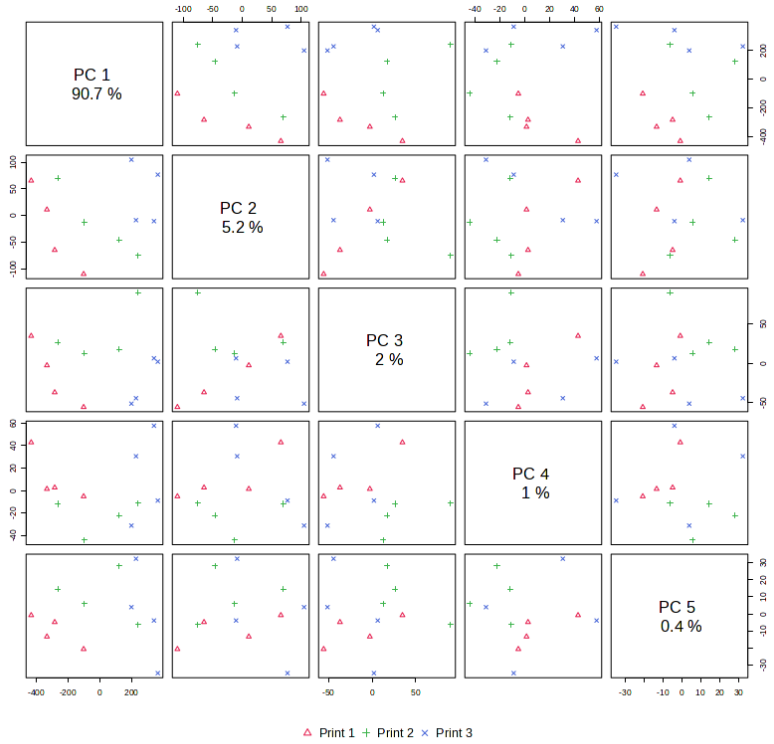
	A	B	C	D	E	F
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD
2	399.3418	75.515	2.3724e-0	5.6248	0.001165	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
3	427.3726	64.871	4.5099e-0	5.3458	0.001165	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
4	607.2084	64.396	4.6515e-0	5.3324	0.001165	Print 1 - Print 2; Print 1 - Print 3; Print 3 - Print 2
5	268.2537	64.365	4.6607e-0	5.3316	0.001165	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
6	296.2838	55.854	8.4389e-0	5.0737	0.001545	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
7	651.2327	52.574	1.0851e-0	4.9645	0.001545	Print 1 - Print 2; Print 1 - Print 3; Print 3 - Print 2
8	563.1842	50.094	1.3253e-0	4.8777	0.001545	Print 1 - Print 2; Print 1 - Print 3; Print 3 - Print 2
9	428.373	50.041	1.3311e-0	4.8758	0.001545	Print 1 - Print 2; Print 1 - Print 3
10	519.1606	49.518	1.39e-05	4.857	0.001545	Print 1 - Print 2; Print 1 - Print 3; Print 3 - Print 2
11	324.314	46.744	1.7622e-0	4.7539	0.001762	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
12	503.1855	45.328	1.999e-05	4.6992	0.001817	Print 1 - Print 2; Print 1 - Print 3
13	1186.933	42.856	2.5135e-0	4.5997	0.002095	Print 2 - Print 1; Print 3 - Print 1
14	531.457	41.594	2.8382e-0	4.547	0.002183	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
15	722.4825	40.077	3.2995e-0	4.4816	0.002357	Print 2 - Print 1; Print 3 - Print 1
16	459.1616	38.012	4.0846e-0	4.3888	0.002723	Print 1 - Print 2; Print 1 - Print 3
17	1018.747	34.846	5.7866e-0	4.2376	0.003617	Print 2 - Print 1; Print 2 - Print 3
18	1085.783	33.437	6.8186e-0	4.1663	0.004011	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
19	695.2563	32.271	7.8462e-0	4.1053	0.004359	Print 1 - Print 2; Print 1 - Print 3
20	368.4108	31.031	9.1568e-0	4.0383	0.004636	Print 1 - Print 2; Print 1 - Print 3

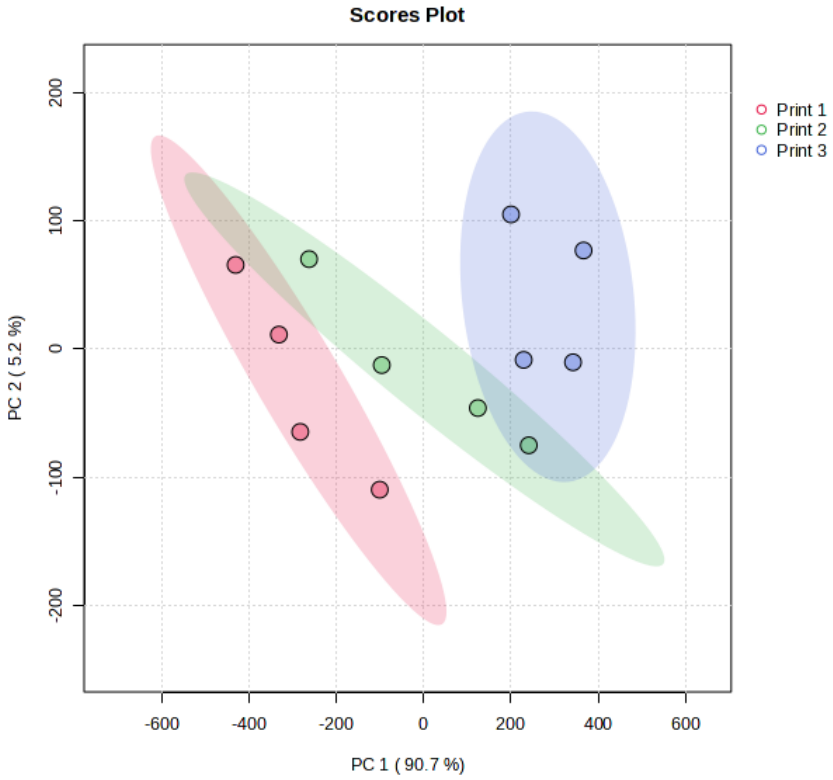


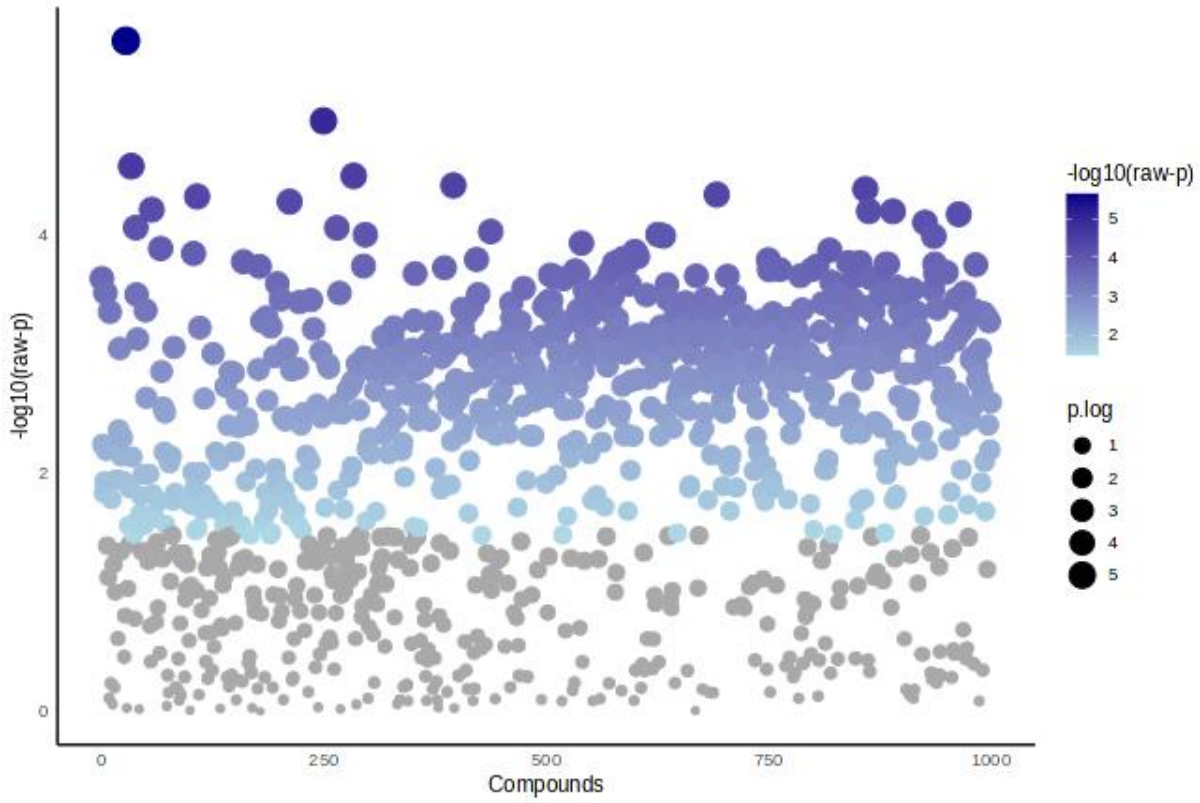
European Female 4 Data



European Female 5 Data

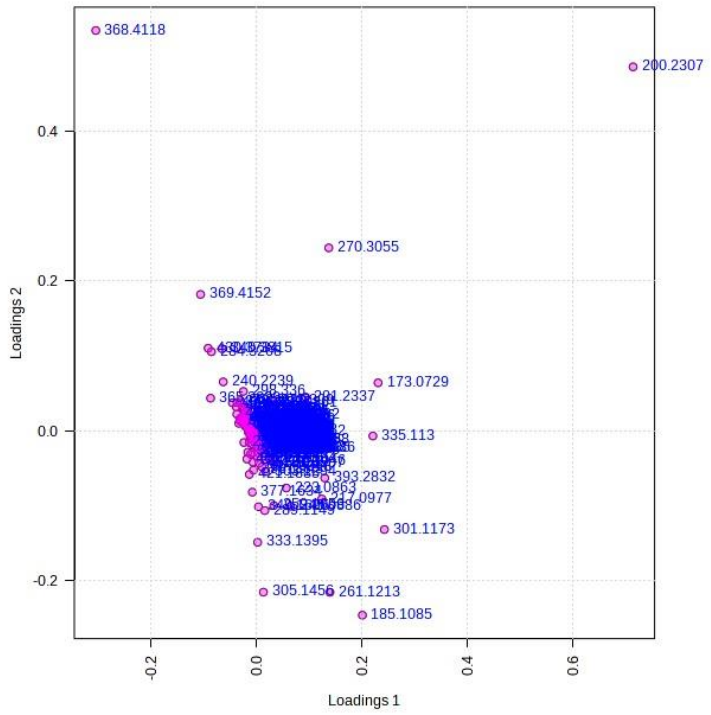




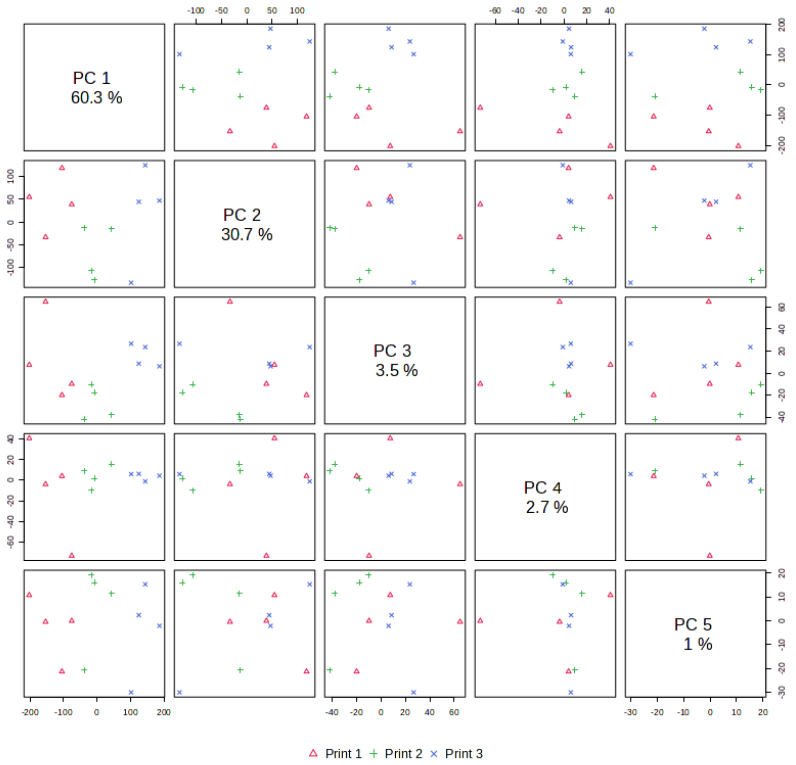


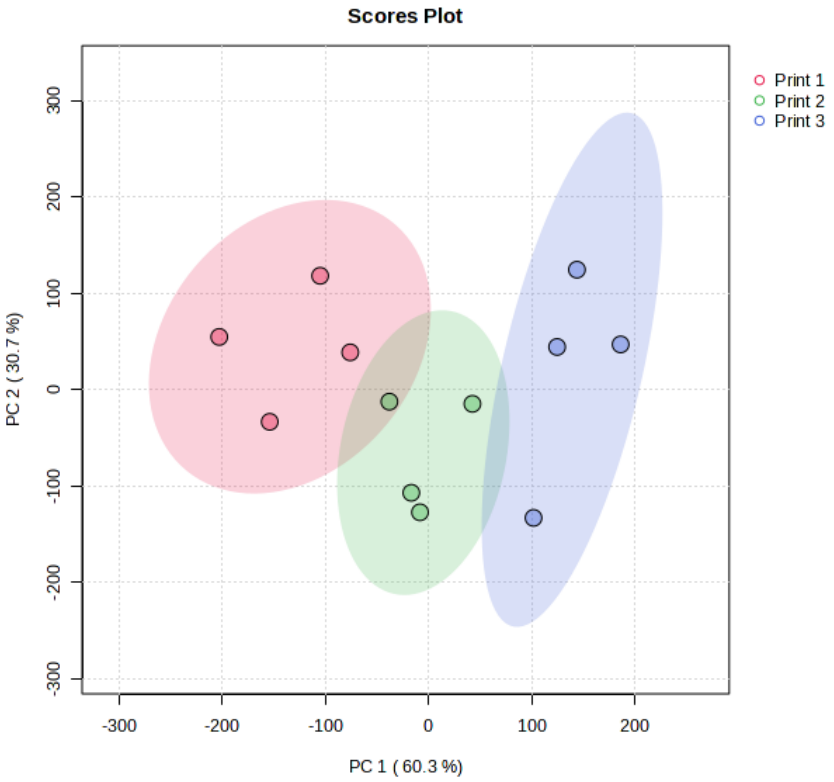
A1 m/z

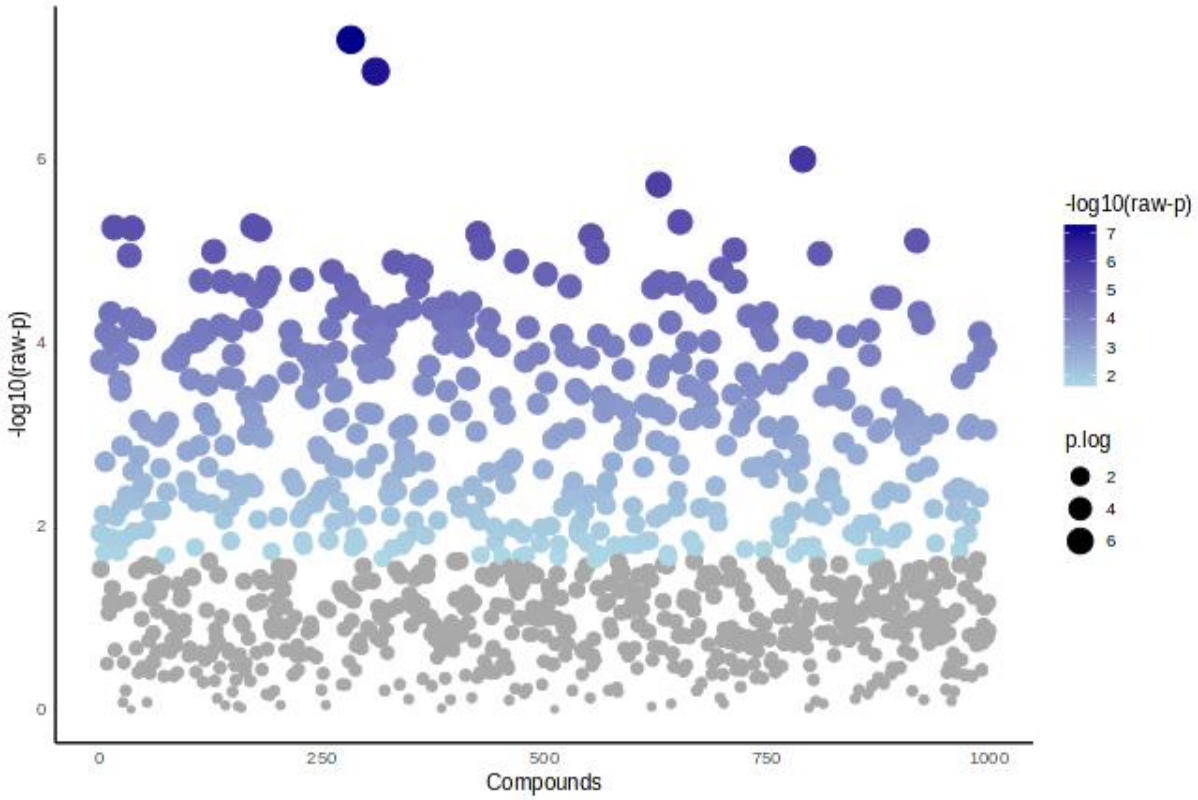
	A	B	C	D	E	F
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD
2	465.2125	76.164	2.2877e-0	5.6406	0.002288	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
3	466.2159	52.653	1.0784e-0	4.9672	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
4	509.2375	42.52	2.5954e-0	4.5858	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
5	211.0892	40.598	3.1313e-0	4.5043	0.003454	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
6	852.2718	38.786	3.7662e-0	4.4241	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
7	281.0414	38.089	4.0517e-0	4.3924	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
8	1119.366	37.107	4.4999e-0	4.3468	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
9	283.1646	36.77	4.6675e-0	4.3309	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
10	599.304	35.858	5.1613e-0	4.2872	0.003454	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
11	256.2906	34.488	6.0293e-0	4.2197	0.003454	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
12	1150.38	34.247	6.2e-05	4.2076	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
13	1120.376	34.217	6.2218e-0	4.2061	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
14	1148.833	33.731	6.5857e-0	4.1814	0.003454	Print 1 - Print 2; Print 1 - Print 3
15	1156.856	32.397	7.7264e-0	4.112	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
16	271.309	31.591	8.5343e-0	4.0688	0.003454	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
17	355.0568	31.561	8.5657e-0	4.0672	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
18	1059.355	30.993	9.2005e-0	4.0362	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
19	1121.382	30.539	9.7488e-0	4.011	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
20	510.2409	30.5	9.7981e-0	4.0089	0.003454	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3



European Female 6 Data

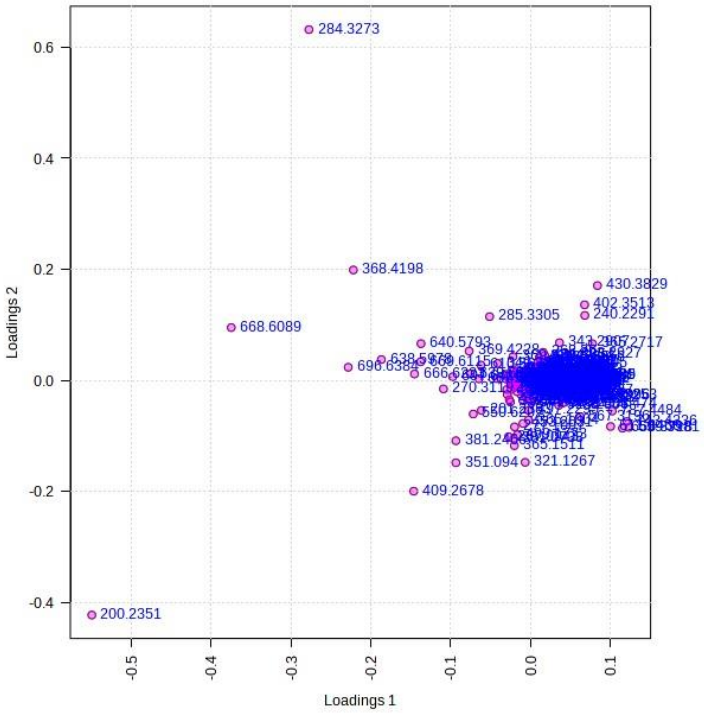



Scores Plot
Print 1
Print 2
Print 3
PC 2 (30.7 %)
PC 1 (60.3 %)
300
200
100
0
-100
-200
-300

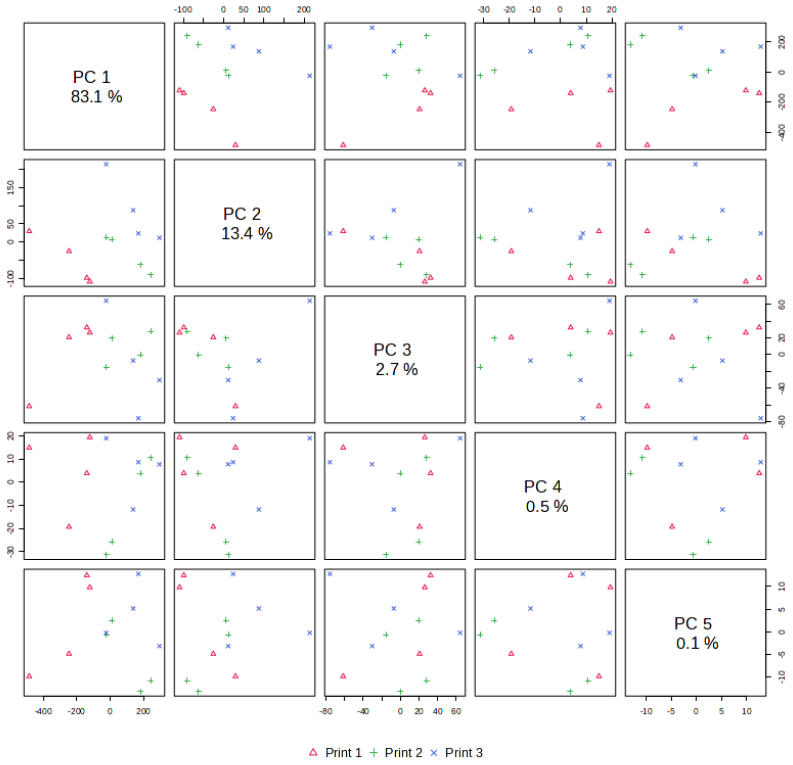

-log10(raw-p)
Compounds
0
250
500
750
1000
0
2
4
6
-log10(raw-p)
7
6
5
4
3
2
p.log
2
4
6

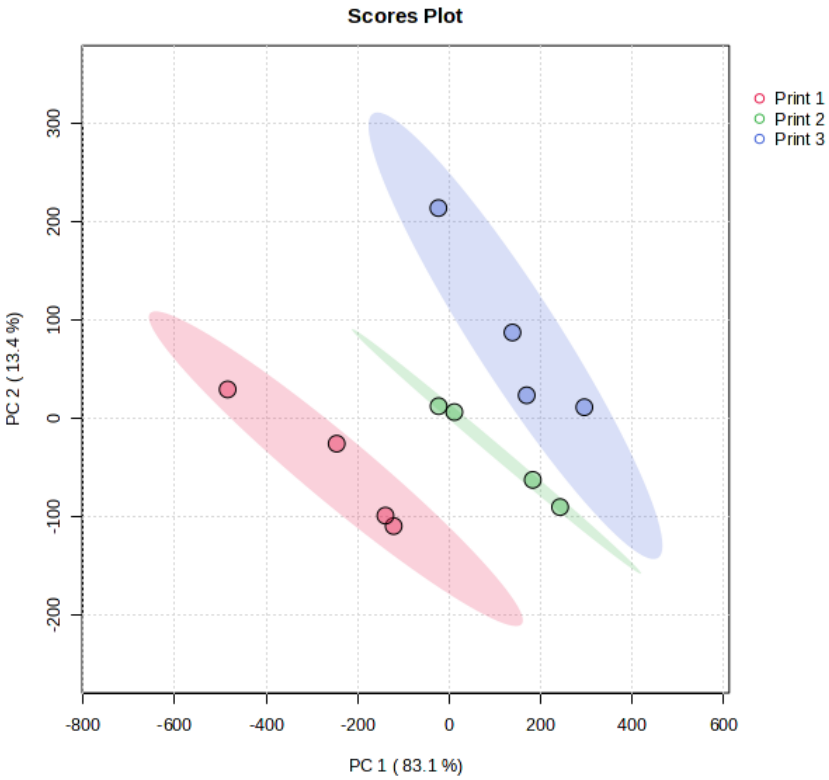
A1 m/z

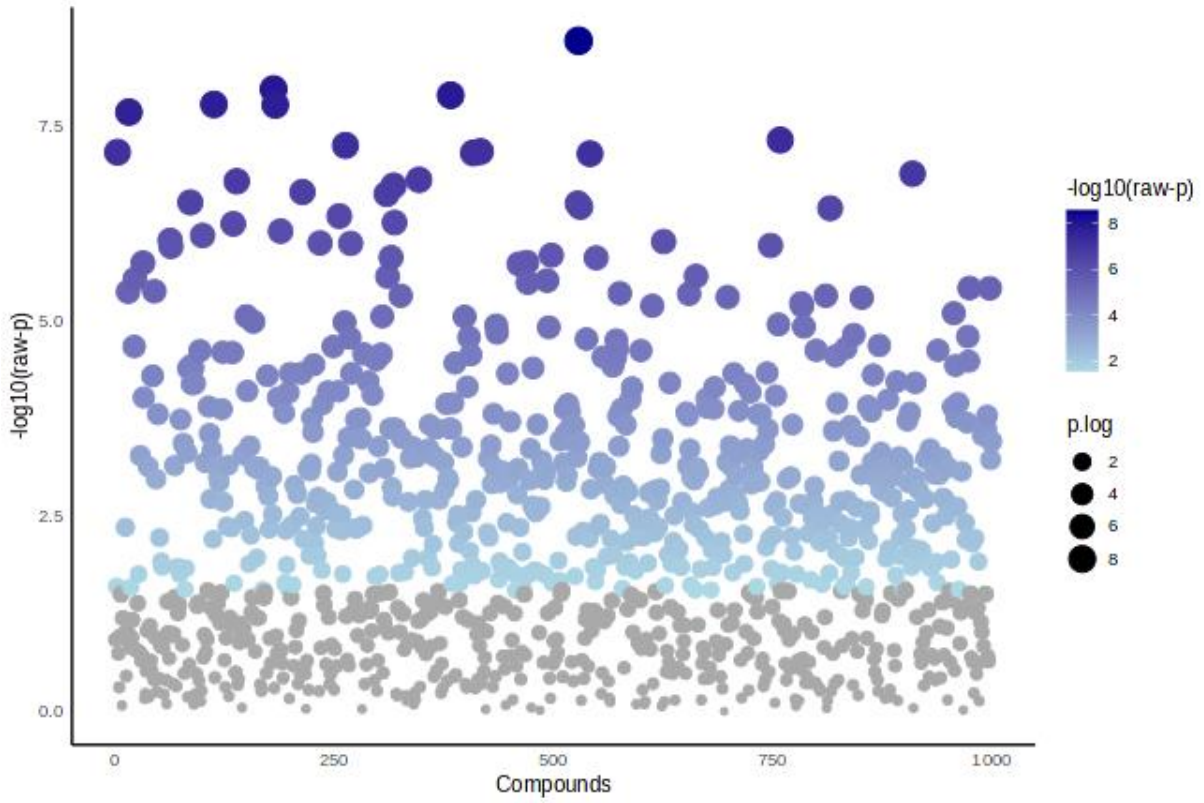
	A	B	C	D	E	F
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD
2	920.8314	183.71	5.0534e-0	7.2964	5.0534e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
3	948.8491	153.3	1.1169e-0	6.952	5.5843e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
4	976.8711	92.14	1.0145e-0	5.9937	0.000338	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
5	921.8348	79.523	1.904e-06	5.7203	0.000476	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
6	949.8522	63.701	4.8686e-0	5.3126	0.000626	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
7	463.302	62.081	5.4248e-0	5.2656	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
8	666.6223	61.643	5.5884e-0	5.2527	0.000626	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
9	301.1327	61.366	5.6948e-0	5.2445	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
10	495.2841	60.971	5.851e-06	5.2328	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
11	493.2719	59.385	6.5335e-0	5.1849	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
12	483.2792	58.59	6.9126e-0	5.1604	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
13	449.2447	56.944	7.7855e-0	5.1087	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
14	581.3271	54.532	9.3232e-0	5.0304	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
15	538.3126	53.831	9.838e-06	5.0071	0.000626	Print 2 - Print 1; Print 3 - Print 1
16	507.3153	53.365	1.02e-05	4.9914	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
17	892.8092	53.16	1.0364e-0	4.9845	0.000626	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
18	508.3255	52.69	1.0753e-0	4.9685	0.000626	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
19	667.6259	52.101	1.1265e-0	4.9483	0.000626	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
20	537.2983	50.246	1.3088e-0	4.8831	0.000663	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2



European Female 7 Data

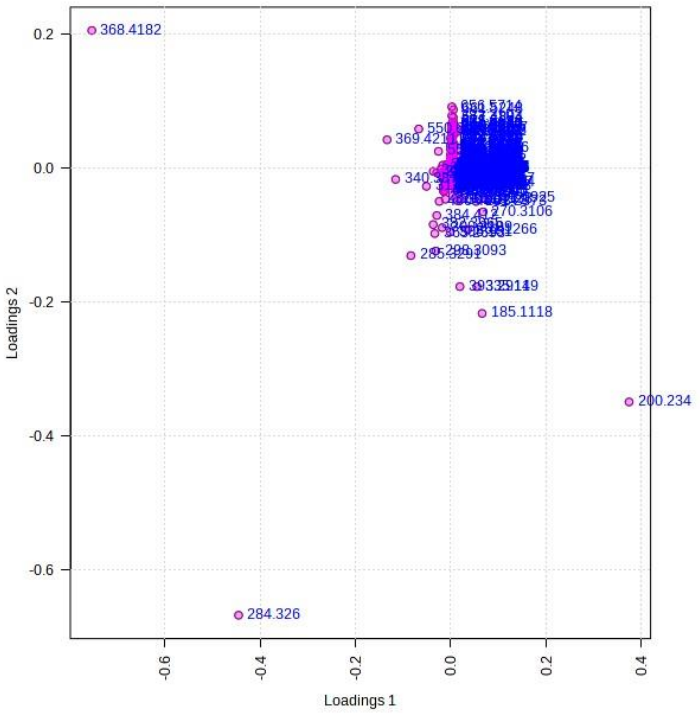




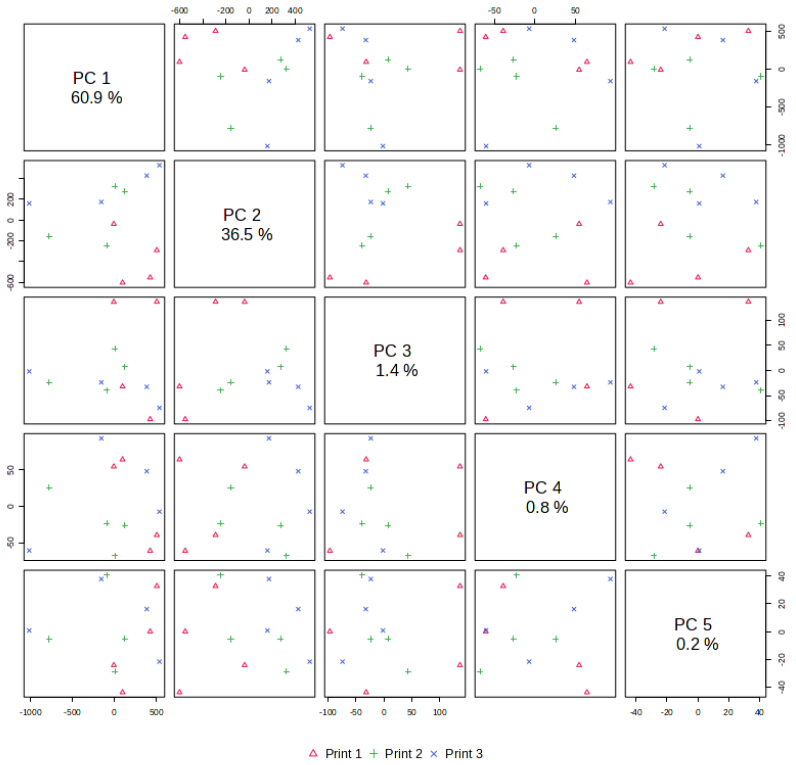


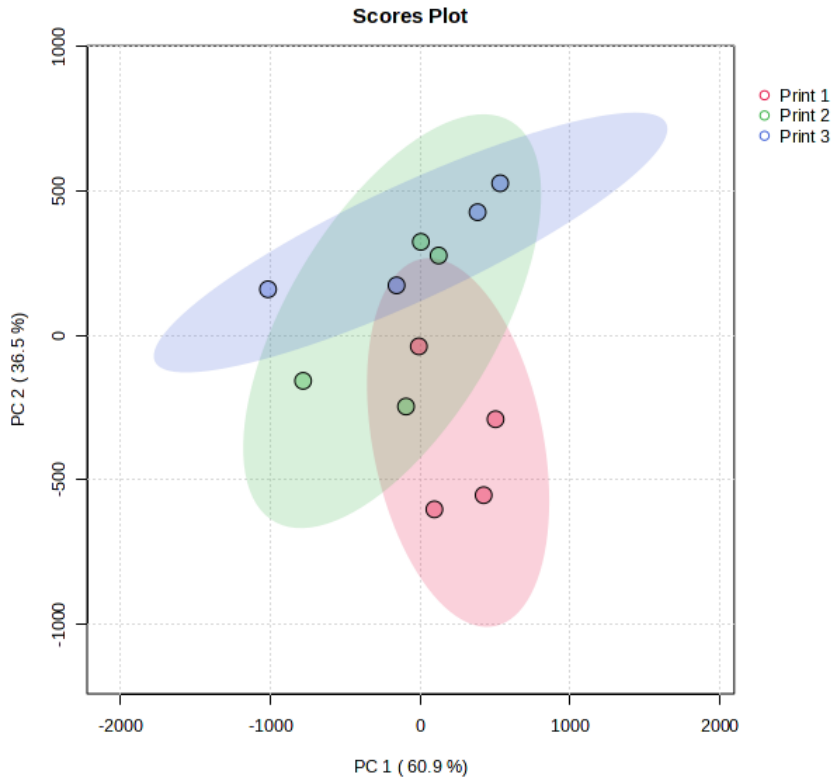
A1 m/z

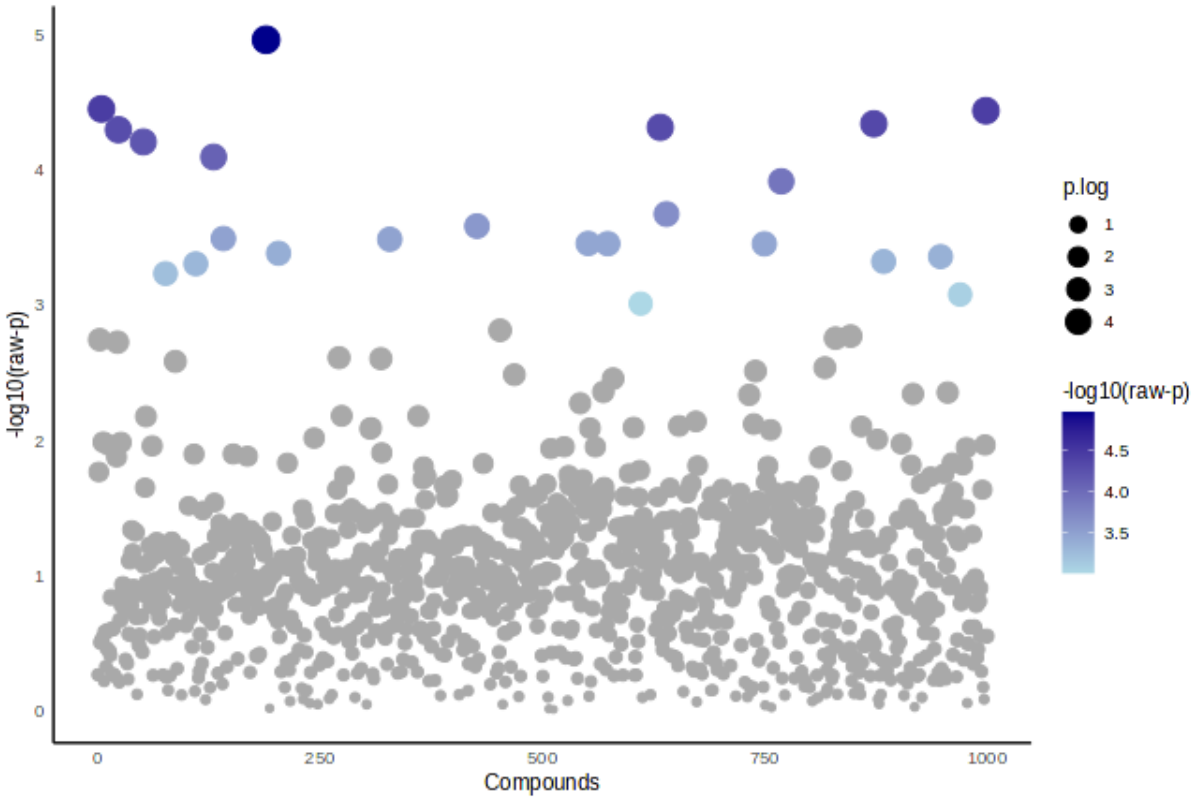
	A	B	C	D	E	F
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD
2	759.4219	360.98	2.5501e-0	8.5934	2.5501e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
3	653.3916	261.74	1.0611e-0	7.9742	3.3804e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
4	715.3966	251.21	1.2722e-0	7.8954	3.3804e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
5	329.31	236.3	1.6673e-0	7.778	3.3804e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
6	657.3539	235.57	1.6902e-0	7.7721	3.3804e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
7	285.3291	224.46	2.0919e-0	7.6794	3.4866e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
8	386.3556	186.12	4.7723e-0	7.3213	5.9387e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
9	627.3402	179.54	5.5888e-0	7.2527	5.9387e-0	Print 2 - Print 1; Print 3 - Print 1
10	673.3323	172.3	6.6961e-0	7.1742	5.9387e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
11	284.326	171.52	6.8302e-0	7.1656	5.9387e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
12	629.3074	170.74	6.9675e-0	7.1569	5.9387e-0	Print 2 - Print 1; Print 3 - Print 1
13	342.3415	169.87	7.1265e-0	7.1471	5.9387e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
14	761.3884	148.36	1.2885e-0	6.8899	9.9119e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
15	671.3717	142.38	1.5419e-0	6.8119	1.0734e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
16	354.3653	140.98	1.6101e-0	6.7931	1.0734e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
17	430.3801	136.93	1.8283e-0	6.738	1.1427e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
18	609.3688	131.1	2.2094e-0	6.6557	1.2996e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
19	495.2603	129.38	2.34e-07	6.6308	1.3e-05	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
20	383.3993	122.21	2.9976e-0	6.5232	1.5476e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3



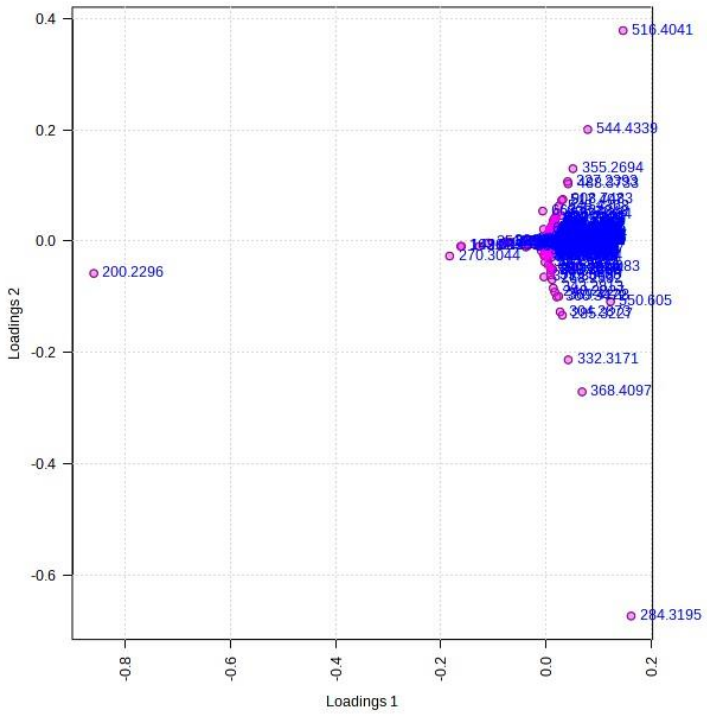
Hispanic Female 3 Data



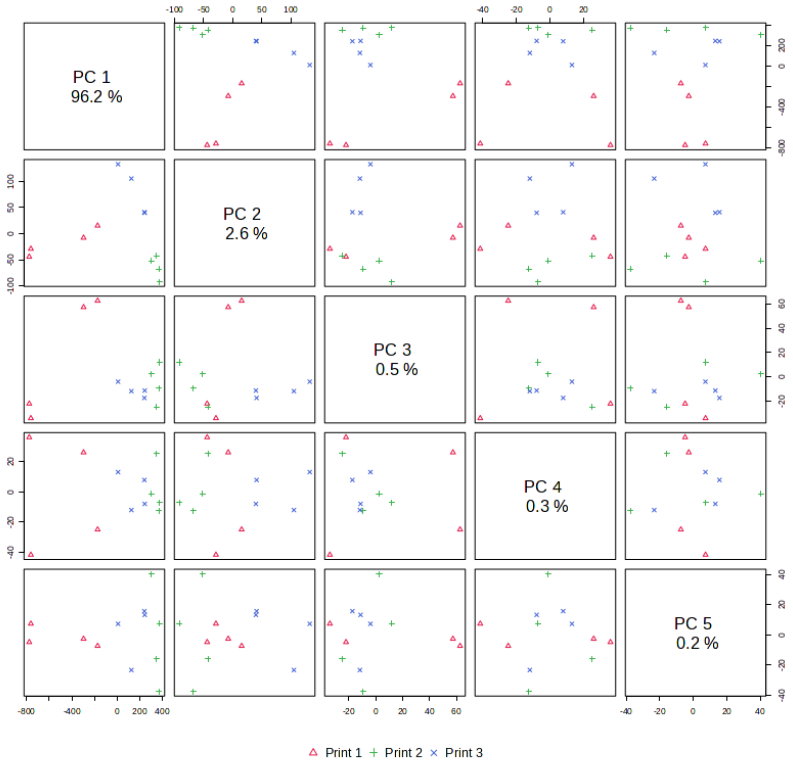

Scores Plot
Print 1
Print 2
Print 3
PC 2 (36.5 %)
PC 1 (60.9 %)
1000
500
0
-500
-1000
-2000
-1000
0
1000
2000

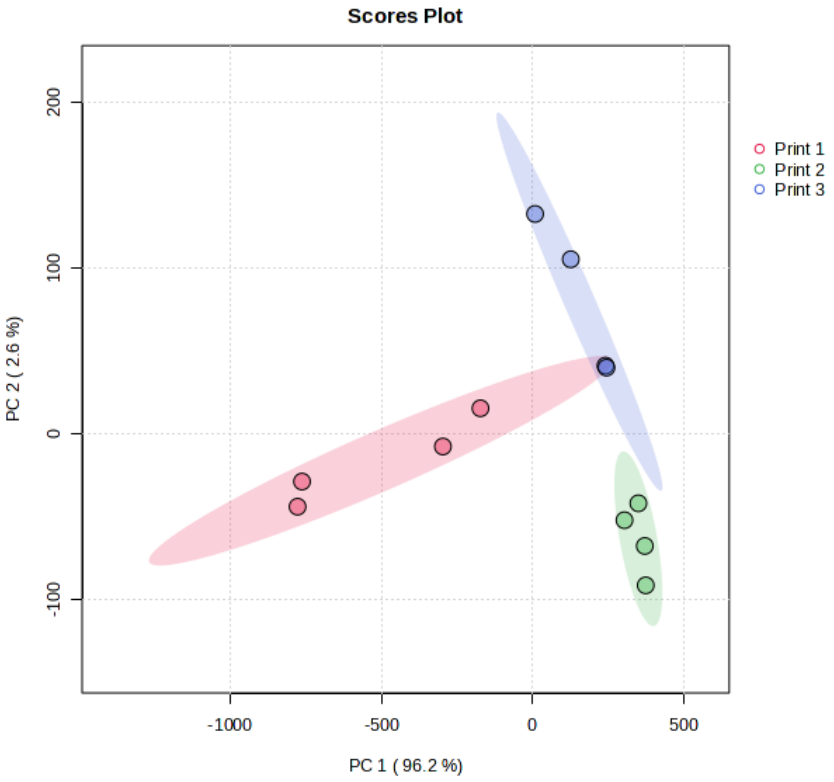

-log10(raw-p)
Compounds
0
250
500
750
1000
0
1
2
3
4
5
p.log
1
2
3
4
-log10(raw-p)
4.5
4.0
3.5

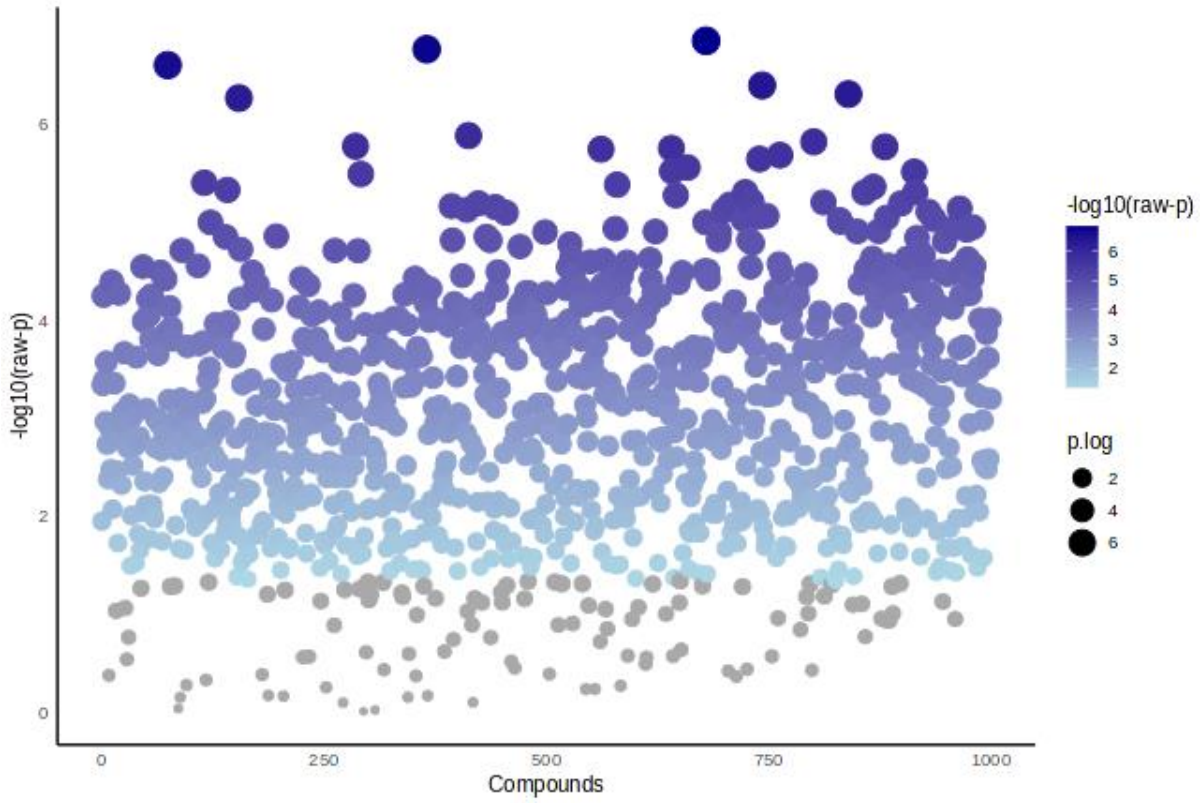
	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD
2	369.3676	52.638	1.0797e-0	4.9667	0.008337	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
3	368.4097	39.471	3.5091e-0	4.4548	0.008337	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
4	370.4153	39.175	3.6177e-0	4.4416	0.008337	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
5	370.3634	37.071	4.5174e-0	4.3451	0.008337	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
6	341.3385	36.51	4.8024e-0	4.3185	0.008337	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
7	369.4128	36.141	5.0019e-0	4.3009	0.008337	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
8	340.3794	34.315	6.151e-05	4.2111	0.008787	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
9	596.5699	32.157	7.9567e-0	4.0993	0.009946	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
10	341.3822	28.932	0.00012	3.9193	0.013381	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
11	597.5732	25.037	0.00021	3.6772	0.021029	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
12	384.4028	23.719	0.000258	3.588	0.021837	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
13	400.3675	22.42	0.000319	3.4959	0.021837	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
14	324.3127	22.345	0.000323	3.4904	0.021837	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
15	624.5964	21.926	0.000347	3.4597	0.021837	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
16	549.4849	21.904	0.000348	3.4581	0.021837	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
17	369.2497	21.886	0.000349	3.4567	0.021837	Print 2 - Print 1; Print 3 - Print 1
18	382.3876	20.963	0.00041	3.3872	0.024115	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
19	430.3709	20.643	0.000434	3.3624	0.024115	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
20	313.3084	20.185	0.000472	3.3265	0.024547	Print 1 - Print 2; Print 1 - Print 3



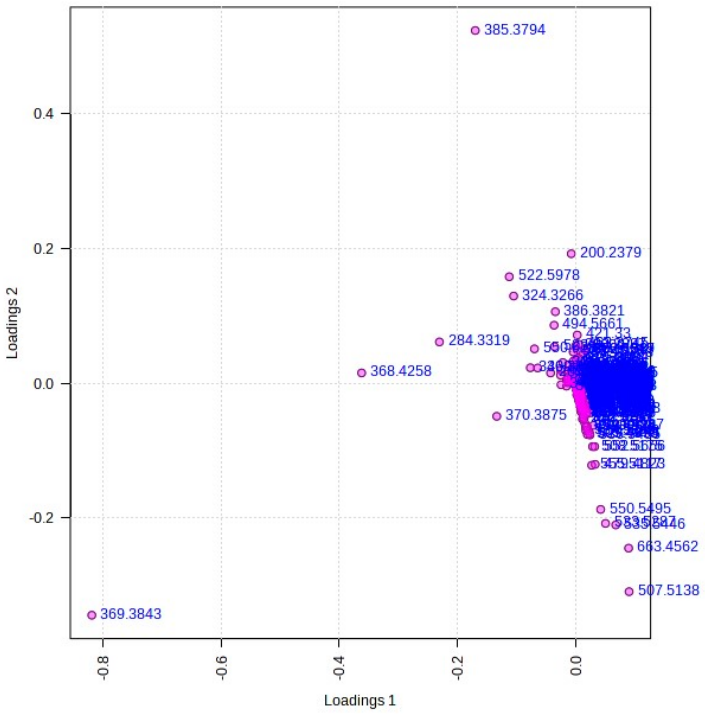
Hispanic Female 5 Data



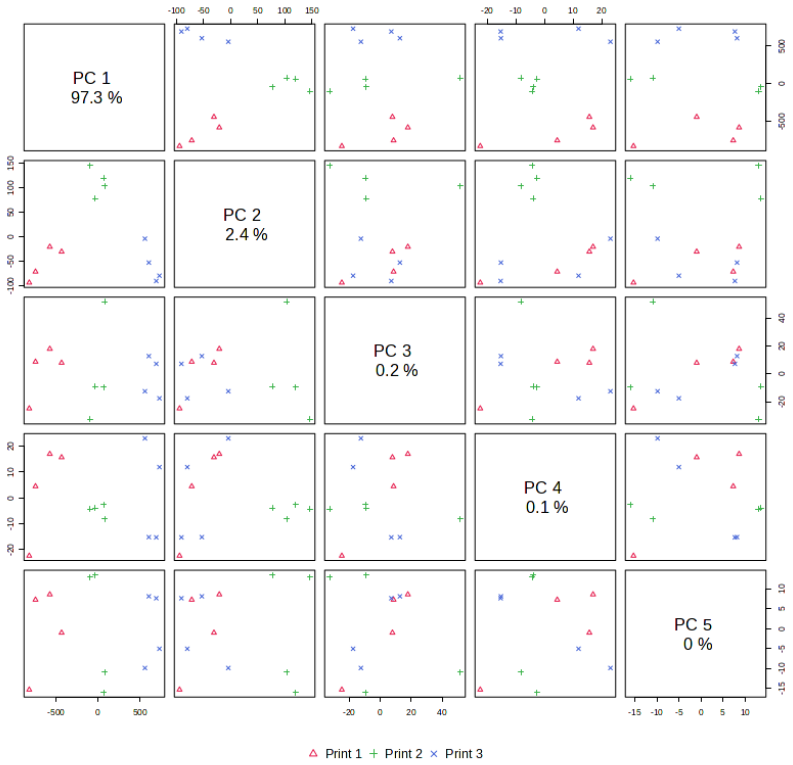




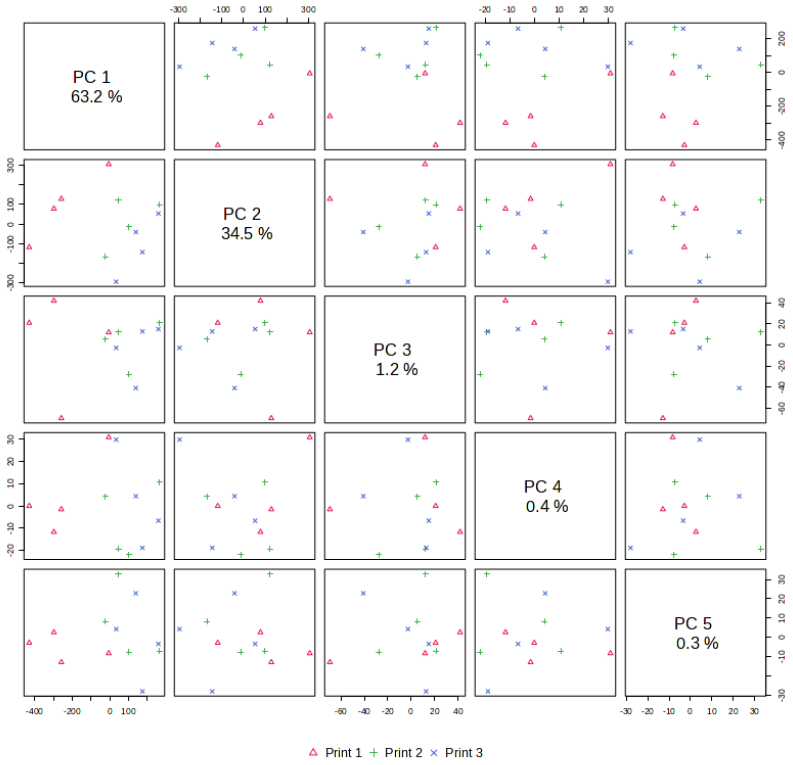
	A	B	C	D	E	F
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD
2	936.6083	145.52	1.4019e-0	6.8533	8.2823e-0	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
3	405.3349	139.05	1.7097e-0	6.7671	8.2823e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
4	635.4667	127.61	2.4847e-0	6.6047	8.2823e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
5	751.5318	114.37	3.9961e-0	6.3984	8.9653e-0	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
6	829.5173	109.11	4.8983e-0	6.31	8.9653e-0	Print 2 - Print 1; Print 3 - Print 1
7	561.3907	106.77	5.3792e-0	6.2693	8.9653e-0	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
8	433.365	86.987	1.2982e-0	5.8867	0.000149	Print 3 - Print 1; Print 3 - Print 2
9	803.5399	84.064	1.5025e-0	5.8232	0.000149	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
10	767.5286	82.048	1.6664e-0	5.7782	0.000149	Print 2 - Print 1; Print 3 - Print 1
11	920.6558	81.8	1.6882e-0	5.7726	0.000149	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
12	746.6222	81.104	1.7508e-0	5.7568	0.000149	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
13	886.7365	80.704	1.7881e-0	5.7476	0.000149	Print 2 - Print 1; Print 3 - Print 1
14	589.4203	78.1	2.0561e-0	5.687	0.000158	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
15	710.6166	76.468	2.2493e-0	5.648	0.000161	Print 2 - Print 1; Print 3 - Print 1
16	855.5781	72.967	2.7443e-0	5.5616	0.000178	Print 2 - Print 1; Print 3 - Print 1
17	705.4306	71.456	2.9987e-0	5.5231	0.000178	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
18	1008.719	71.24	3.0373e-0	5.5175	0.000178	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
19	923.5991	70.373	3.1988e-0	5.495	0.000178	Print 2 - Print 1; Print 3 - Print 1; Print 2 - Print 3
20	646.5365	67.039	3.9266e-0	5.406	0.000198	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2

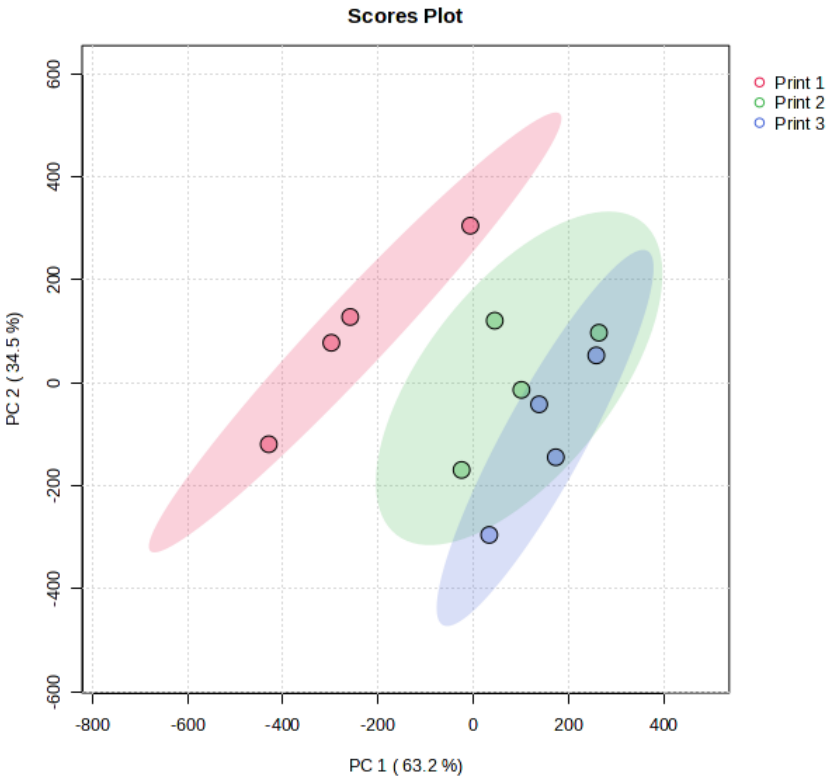


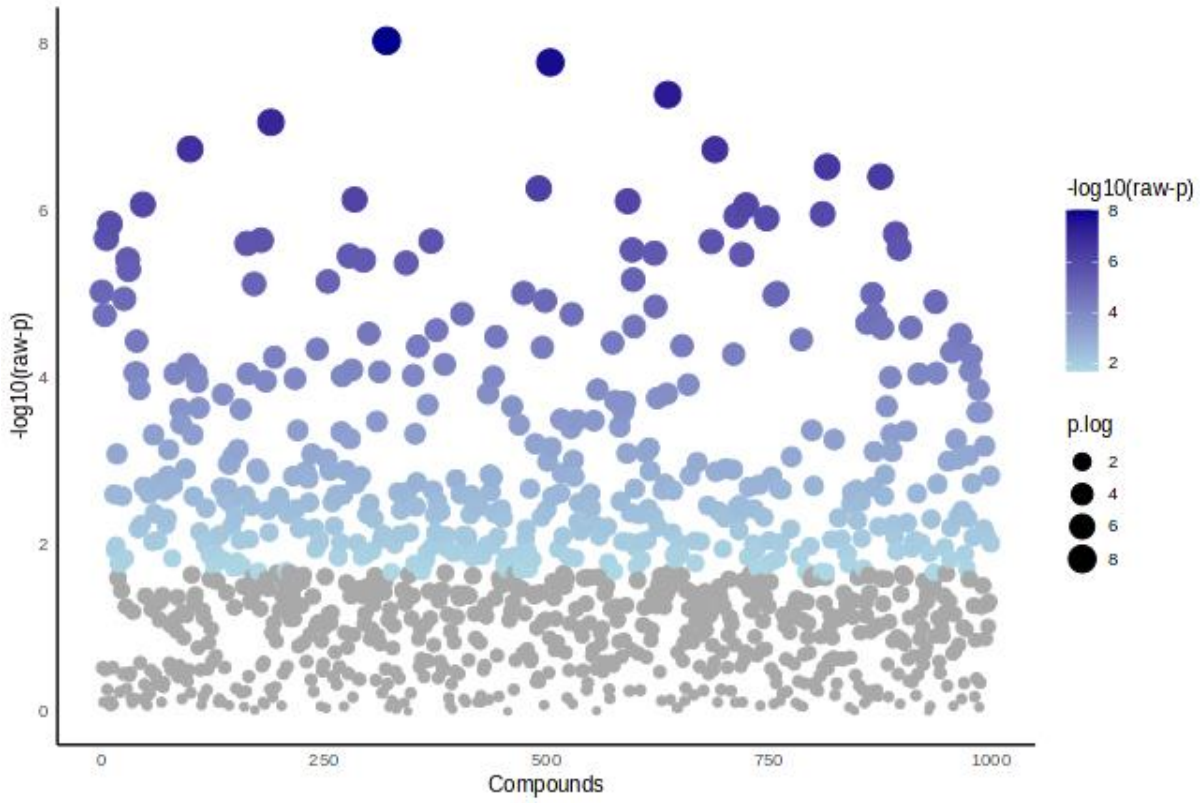
Hispanic Female 6 Data



Middle Eastern Female 1 Data

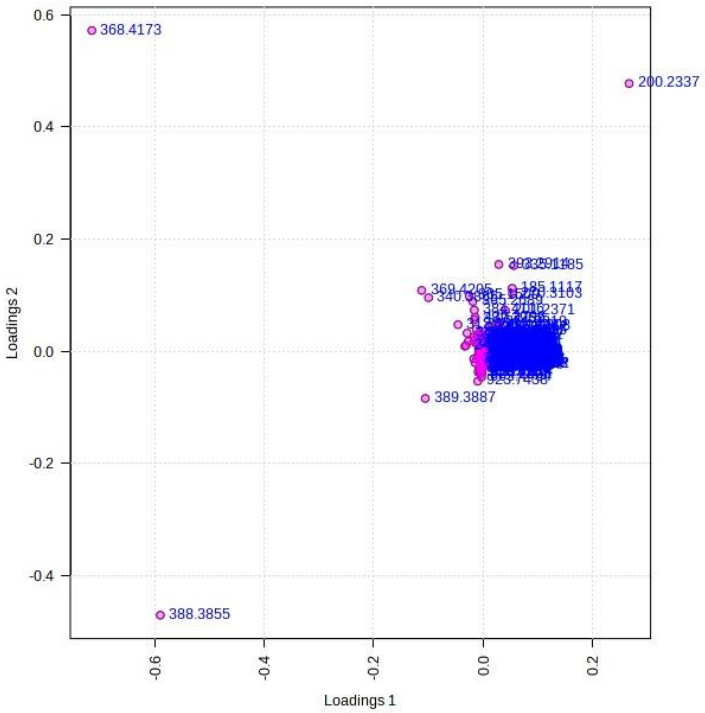






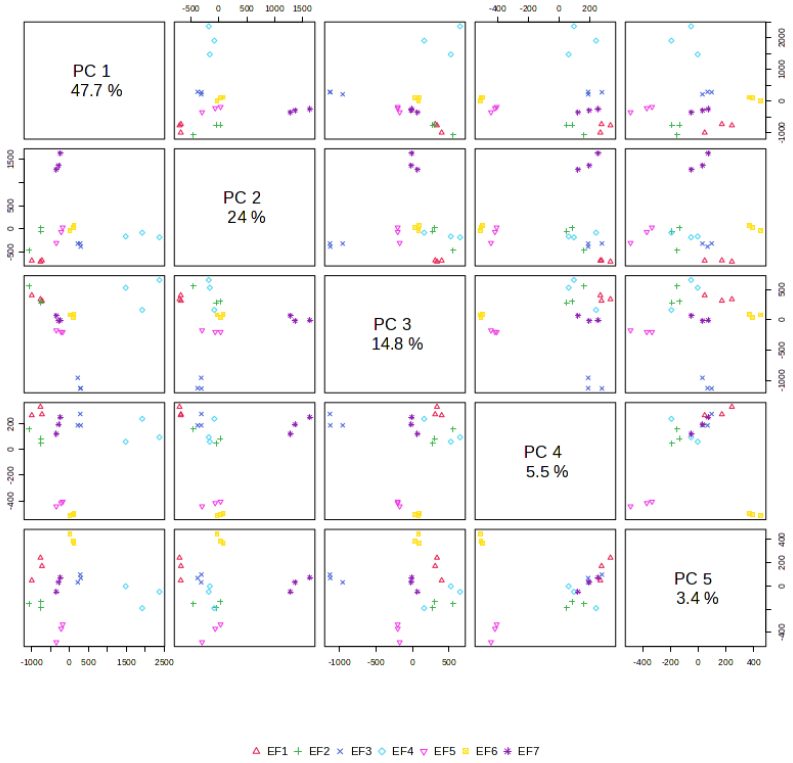
A1 m/z

	A	B	C	D	E	F
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD
2	859.5212	269.26	9.3609e-0	8.0287	8.4982e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
3	772.4729	235.27	1.6996e-0	7.7696	8.4982e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
4	816.4978	192.37	4.1273e-0	7.3843	1.3758e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
5	815.4946	161.75	8.8324e-0	7.0539	2.2081e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
6	771.4697	136.59	1.8483e-0	6.7332	3.1176e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
7	581.3392	136.21	1.8706e-0	6.728	3.1176e-0	Print 2 - Print 1; Print 3 - Print 1
8	860.5254	122.14	3.0057e-0	6.5221	4.2938e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
9	947.5727	114.75	3.9395e-0	6.4046	4.9244e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
10	851.5815	106.36	5.4687e-0	6.2621	6.0763e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
11	697.4406	99.144	7.4046e-0	6.1305	6.6842e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
12	903.5455	98.077	7.7576e-0	6.1103	6.6842e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
13	727.4438	96.075	8.4771e-0	6.0718	6.6842e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
14	741.4566	95.524	8.6895e-0	6.061	6.6842e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
15	669.3837	90.379	1.102e-06	5.9578	7.745e-05	Print 2 - Print 1; Print 3 - Print 1
16	657.3653	89.273	1.1618e-0	5.9349	7.745e-05	Print 2 - Print 1; Print 3 - Print 1
17	480.3268	87.82	1.2463e-0	5.9044	7.7893e-0	Print 2 - Print 1; Print 3 - Print 1; Print 3 - Print 2
18	312.3561	84.809	1.4469e-0	5.8396	8.5112e-0	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3
19	485.2843	79.486	1.9078e-0	5.7195	0.000106	Print 2 - Print 1; Print 3 - Print 1
20	340.3865	77.25	2.1541e-0	5.6667	0.000107	Print 1 - Print 2; Print 1 - Print 3; Print 2 - Print 3

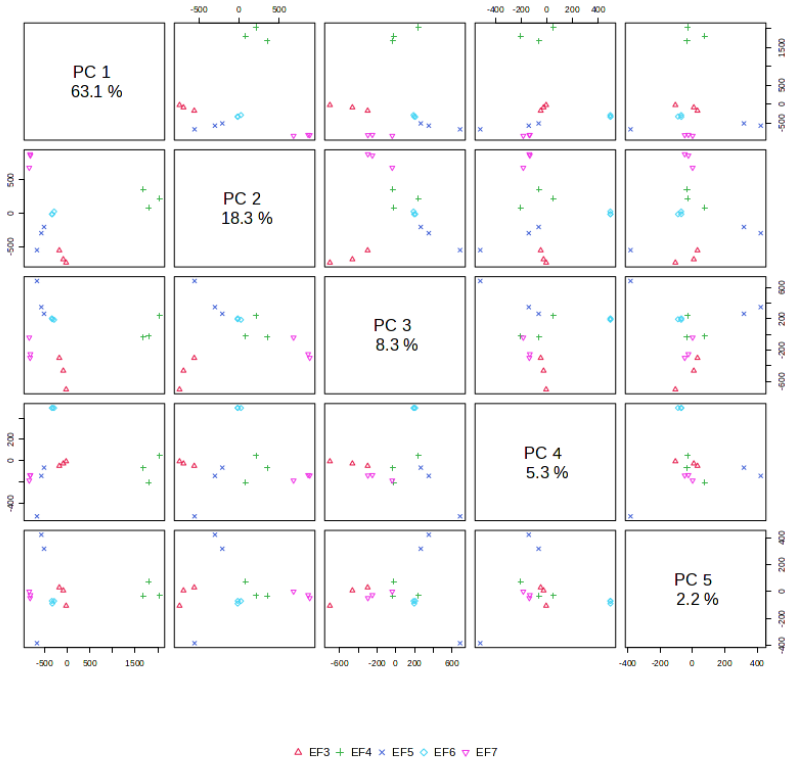

368.4173
200.2337
389.3887
388.3855
Loadings 2
Loadings 1
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.4
-0.2
0.0
0.2

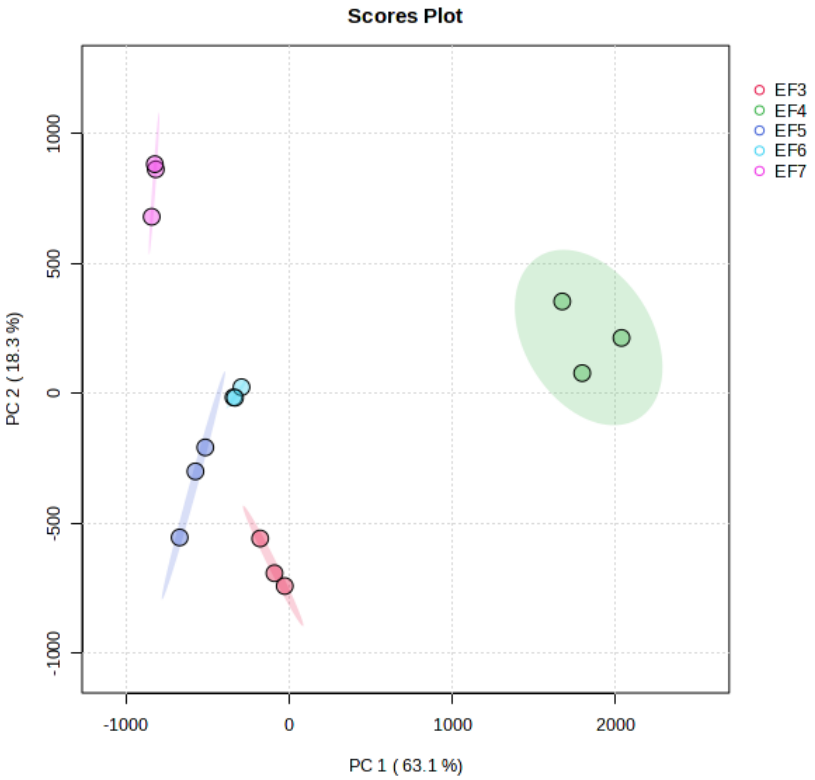
Same Geographical and Cultural Ancestry Data

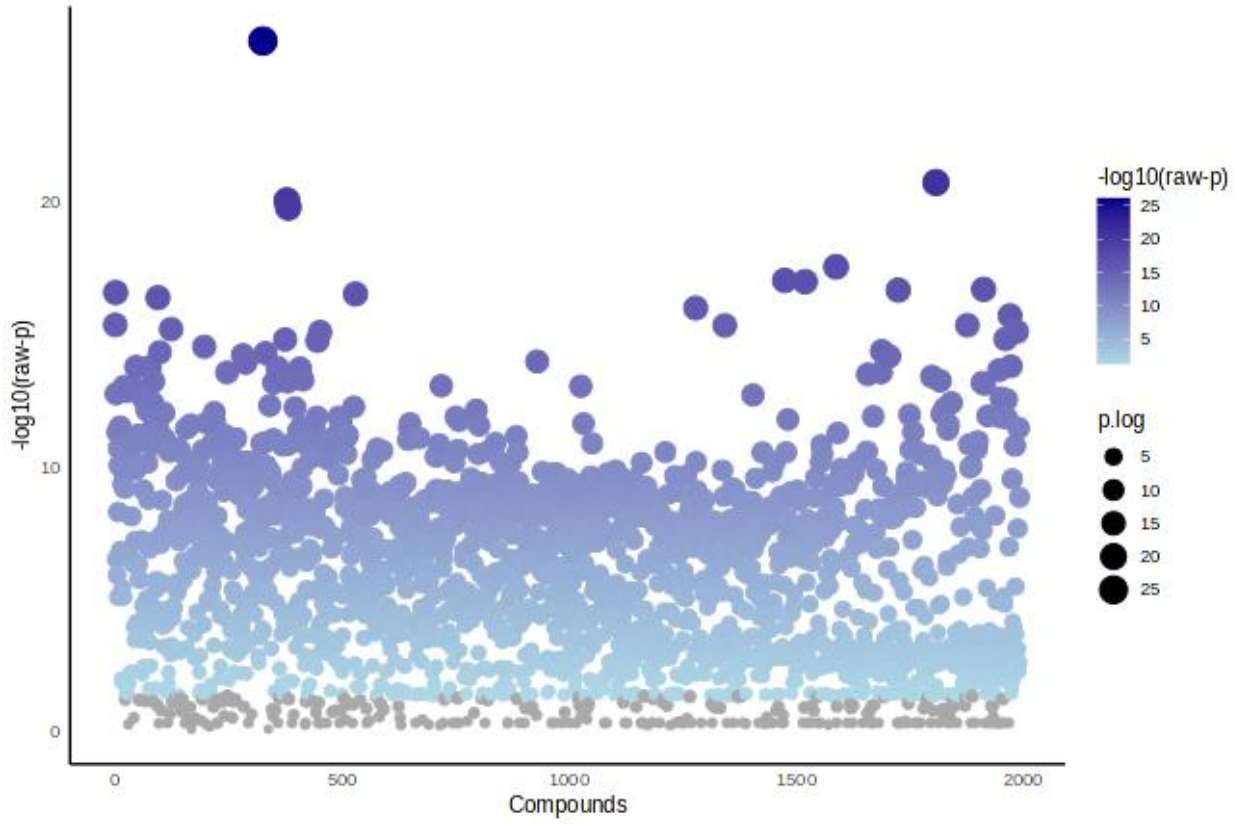
European Print 1 Data



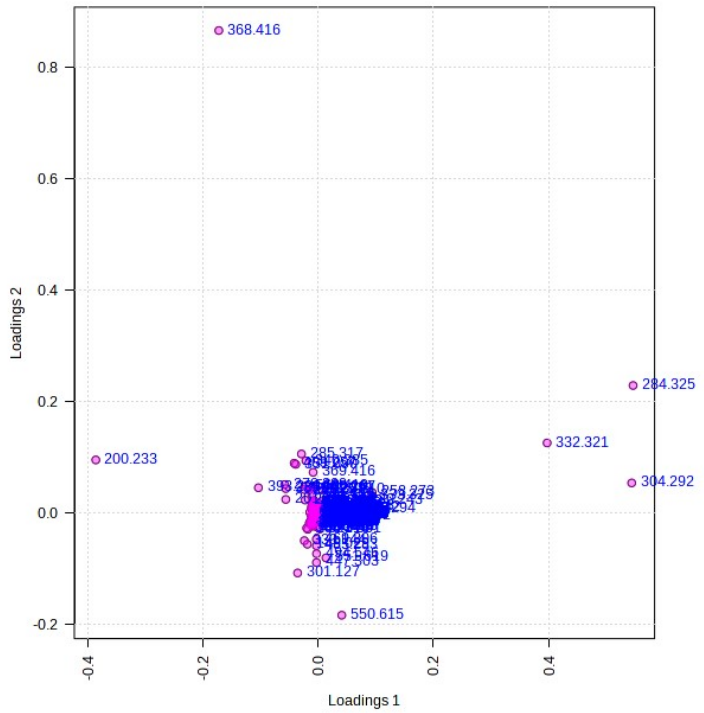
European Print 2 Data



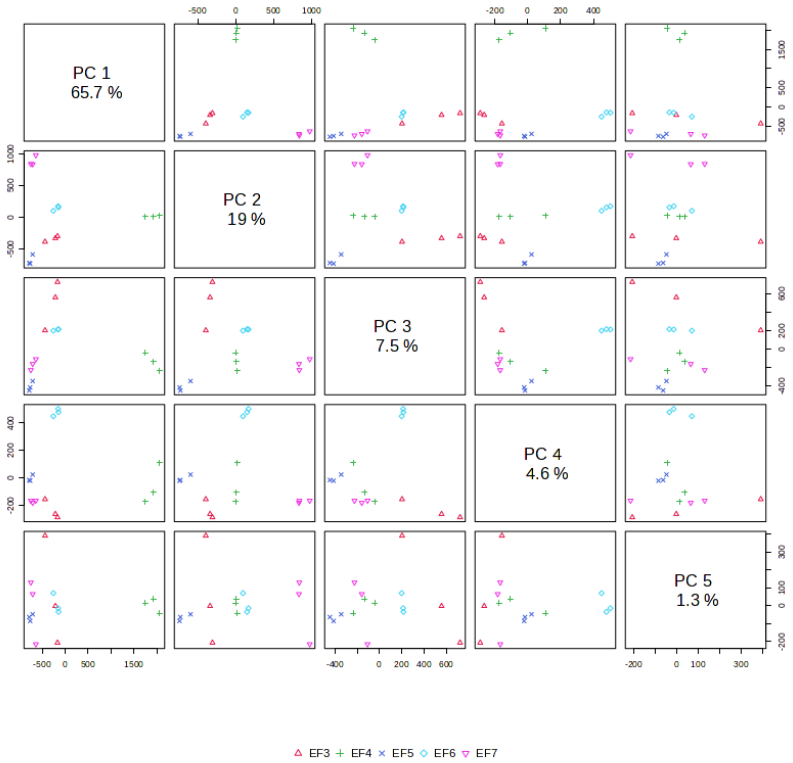


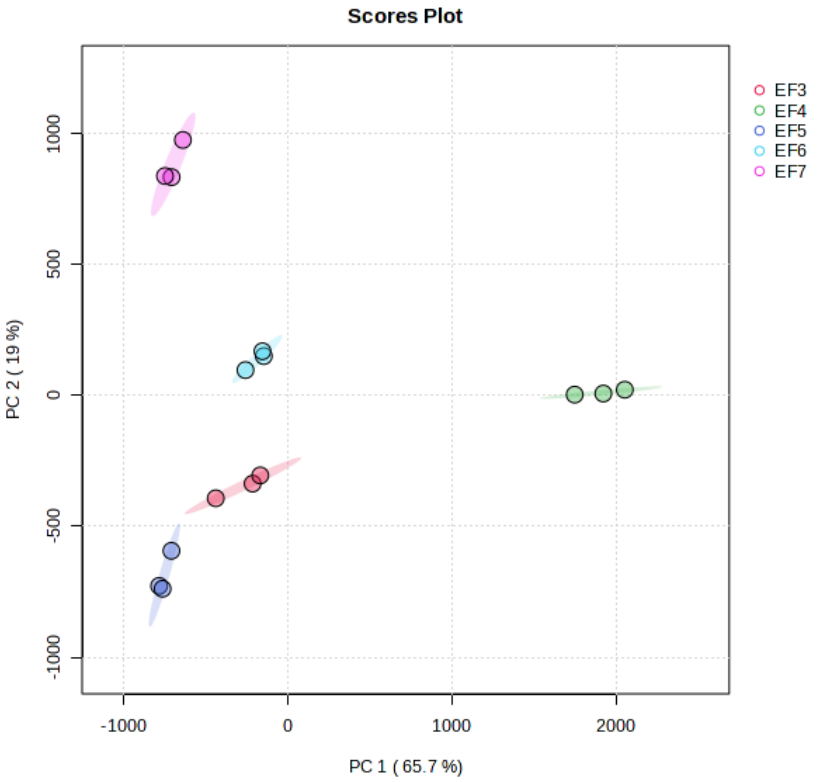


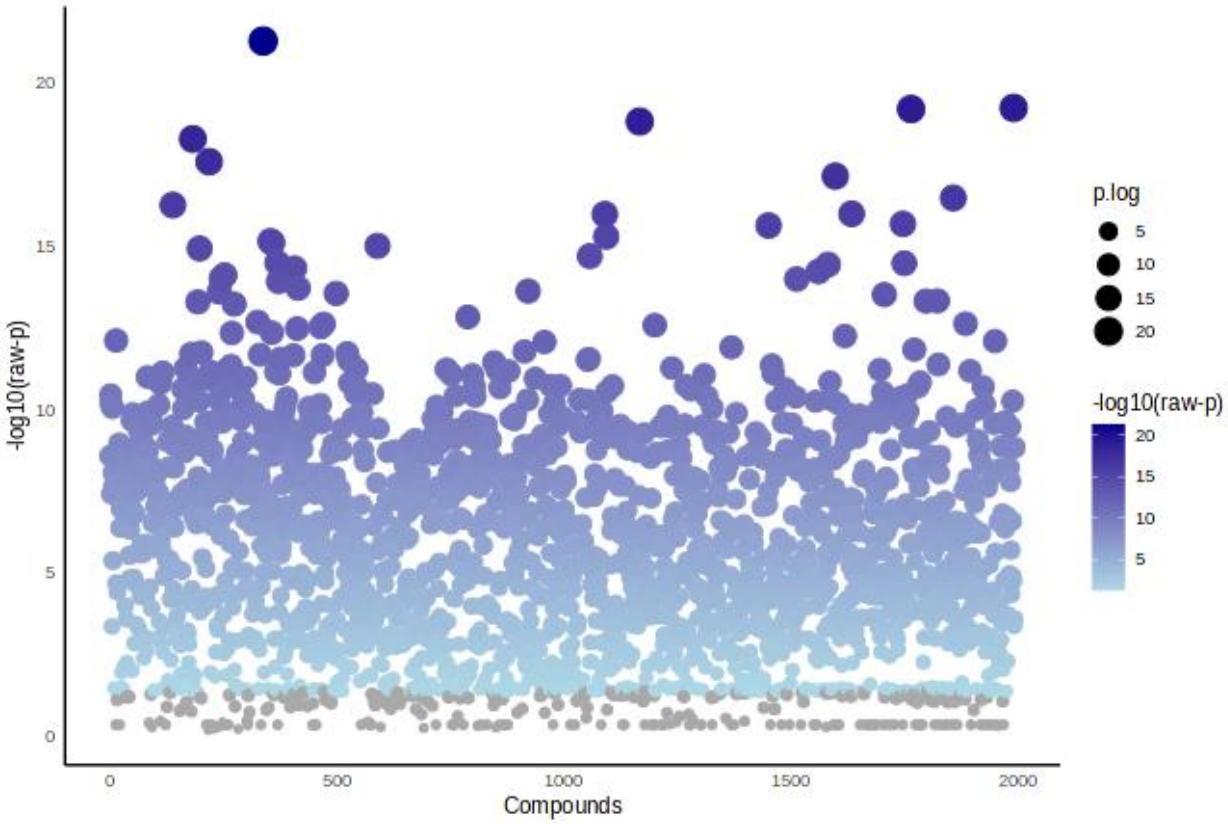
	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD (top 1000)
2	414.261	588020	8.3344e-2	26.079	1.661e-23	EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7
3	1095.805	50247	1.8288e-2	20.738	1.8224e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
4	446.112	36635	8.8756e-2	20.052	5.8963e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
5	447.109	32633	1.5827e-2	19.801	7.8859e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
6	976.871	11615	2.7683e-1	17.558	1.1034e-1	EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7
7	932.848	9166.5	9.0394e-1	17.044	2.8777e-1	EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7
8	948.849	8964	1.0107e-1	16.995	2.8777e-1	EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7
9	1154.882	7850.6	1.9613e-1	16.707	4.6512e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
10	1047.761	7743.7	2.1004e-1	16.678	4.6512e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
11	121.099	7427.7	2.5867e-1	16.587	5.1552e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
12	550.333	7236.9	2.9459e-1	16.531	5.3374e-1	EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7
13	244.101	6790.5	4.0496e-1	16.393	6.7256e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
14	856.491	5672.6	9.95e-17	16.002	1.5254e-1	EF4 - EF3; EF4 - EF5; EF4 - EF6; EF4 - EF7
15	1187.894	4970.4	1.9259e-1	15.715	2.7416e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
16	111.043	4230.4	4.3097e-1	15.366	5.3064e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
17	1134.861	4205.3	4.44e-16	15.353	5.3064e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7
18	880.54	4189.1	4.5263e-1	15.344	5.3064e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
19	272.285	3926.3	6.2566e-1	15.204	6.9275e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6
20	1196.798	3760.1	7.7656e-1	15.11	8.0482e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6



European Print 3 Data

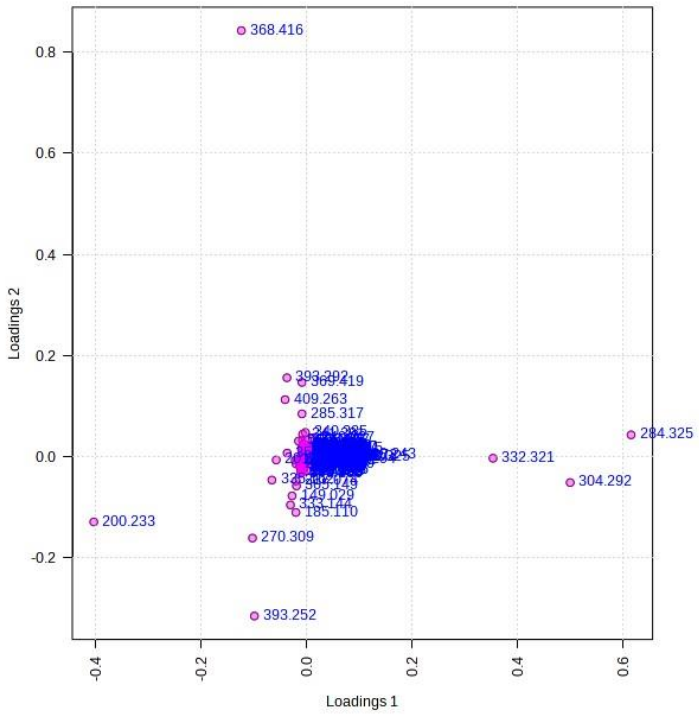



Scores Plot
EF3
EF4
EF5
EF6
EF7
PC 2 (19 %)
PC 1 (65.7 %)
1000
500
0
-500
-1000
-1000
0
1000
2000

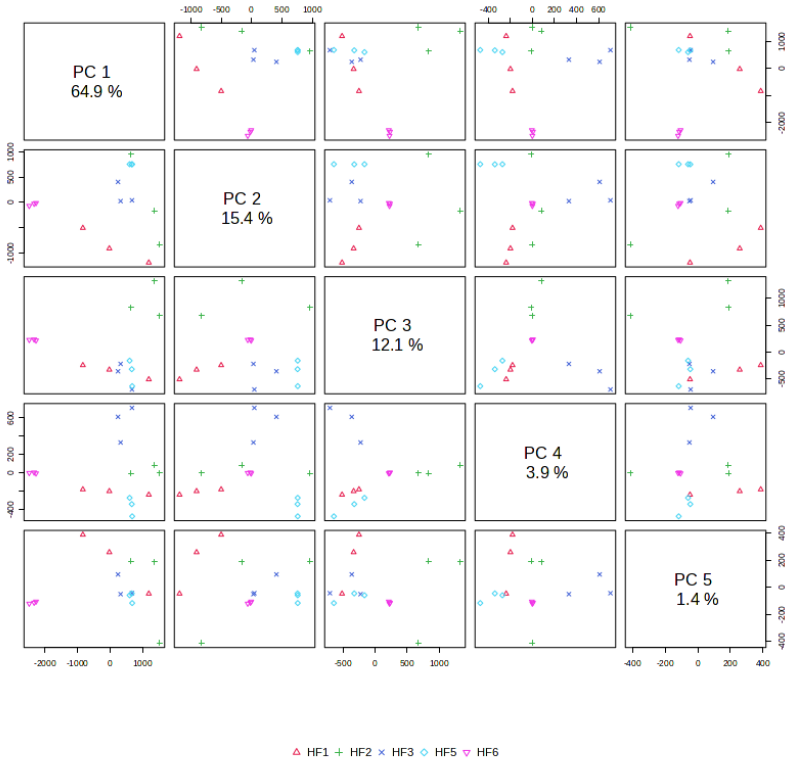

-log10(raw-p)
Compounds
0
500
1000
1500
2000
0
5
10
15
20
p.log
5
10
15
20
-log10(raw-p)
20
15
10
5

A1 m/z

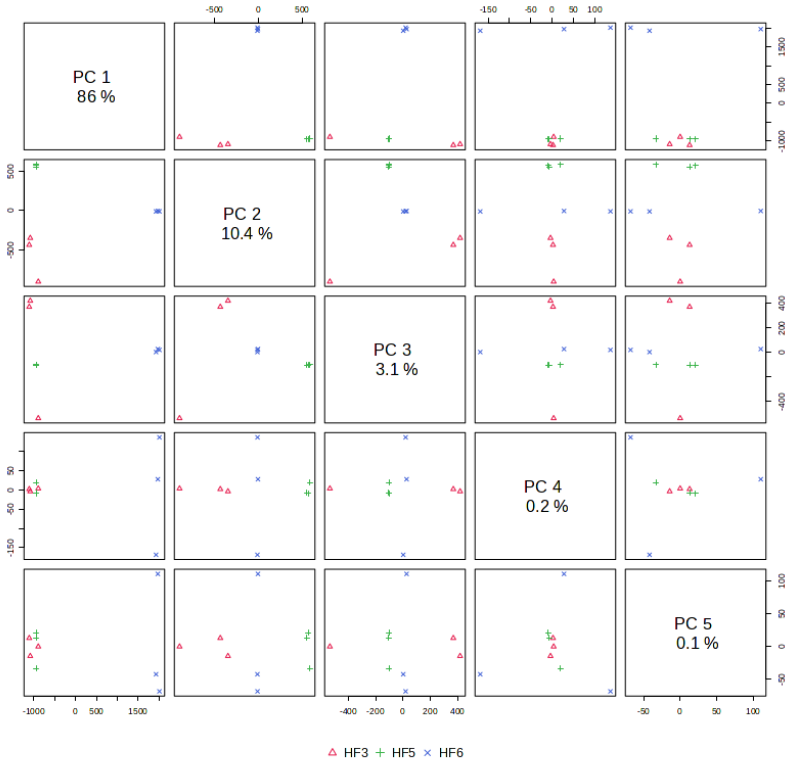
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD (top 1000)													
2	422.192	65007	5.046e-22	21.297	1.0067e-1	EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7													
3	1198.924	25292	5.6579e-2	19.247	4.0771e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7													
4	1069.773	24889	6.131e-20	19.212	4.0771e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7													
5	810.434	20884	1.4741e-1	18.831	7.3519e-1	EF4 - EF3; EF4 - EF5; EF4 - EF6; EF4 - EF7													
6	318.152	16313	5.067e-19	18.295	2.0217e-1	EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7													
7	341.391	11838	2.5174e-1	17.599	8.3705e-1	EF4 - EF3; EF6 - EF3; EF7 - EF3; EF4 - EF5; EF6 - EF4; EF7 - EF4; EF6 - EF5; EF7 - EF5; EF7 - EF6													
8	978.848	9655.7	6.9707e-1	17.157	1.9866e-1	EF6 - EF3; EF6 - EF4; EF6 - EF5; EF6 - EF7													
9	1122.853	7097.9	3.2457e-1	16.489	8.094e-15	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7													
10	286.325	6419.3	5.3633e-1	16.271	1.1889e-1	EF6 - EF3; EF7 - EF3; EF6 - EF4; EF7 - EF4; EF6 - EF5; EF7 - EF5; EF7 - EF6													
11	996.69	5674.2	9.9363e-1	16.003	1.8652e-1	EF7 - EF3; EF7 - EF4; EF7 - EF5; EF7 - EF6													
12	777.202	5635.2	1.0284e-1	15.988	1.8652e-1	EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7													
13	1057.726	4919.3	2.0279e-1	15.693	3.3713e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7													
14	919.654	4810.2	2.2685e-1	15.644	3.4812e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7													
15	778.2	4112.1	4.9657e-1	15.304	7.0762e-1	EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7													
16	431.241	3855.1	6.8558e-1	15.164	9.1182e-1	EF5 - EF3; EF6 - EF3; EF5 - EF4; EF6 - EF4; EF5 - EF6; EF5 - EF7; EF6 - EF7													
17	433.201	3775	7.6135e-1	15.118	9.4931e-1	EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7													
18	581.573	3608.9	9.533e-16	15.021	1.1187e-1	EF3 - EF4; EF3 - EF5; EF3 - EF6; EF3 - EF7													
19	329.18	3476.3	1.1494e-1	14.94	1.2739e-1	EF5 - EF3; EF5 - EF4; EF5 - EF6; EF5 - EF7													
20	764.431	3122	1.9664e-1	14.706	2.0647e-1	EF4 - EF3; EF4 - EF5; EF4 - EF6; EF4 - EF7													

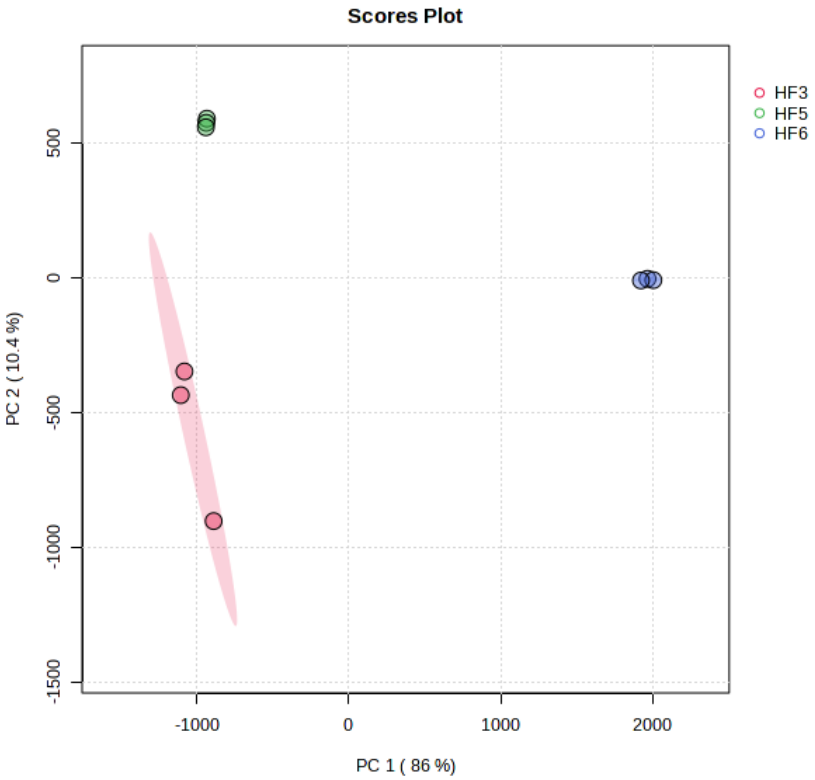


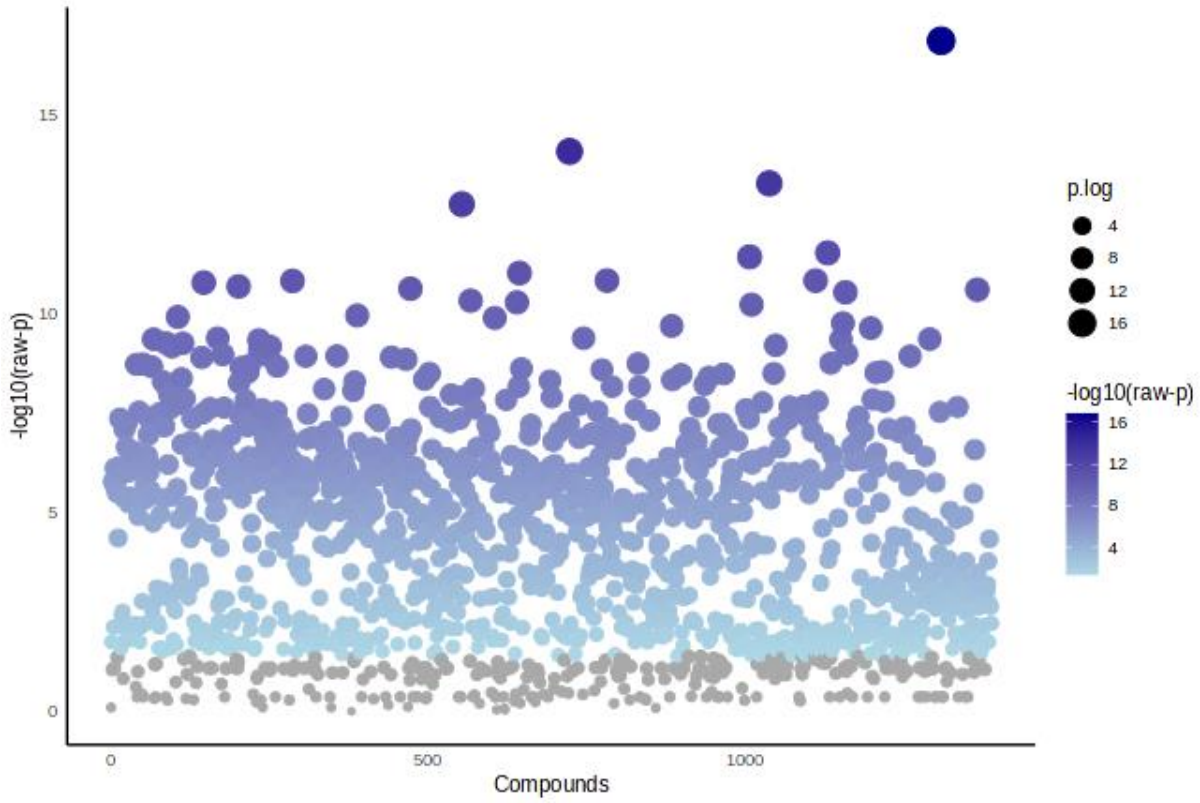
Hispanic Print 1 Data



Hispanic Print 2 Data

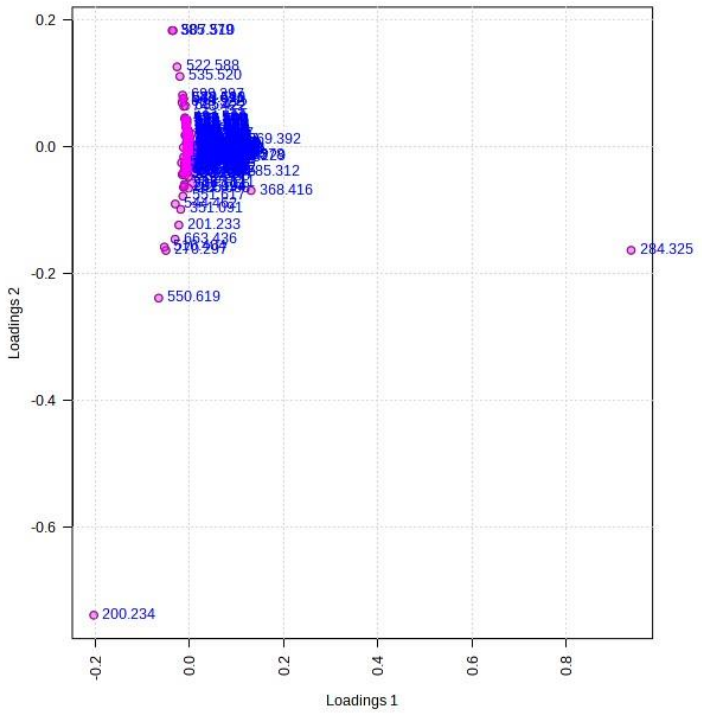




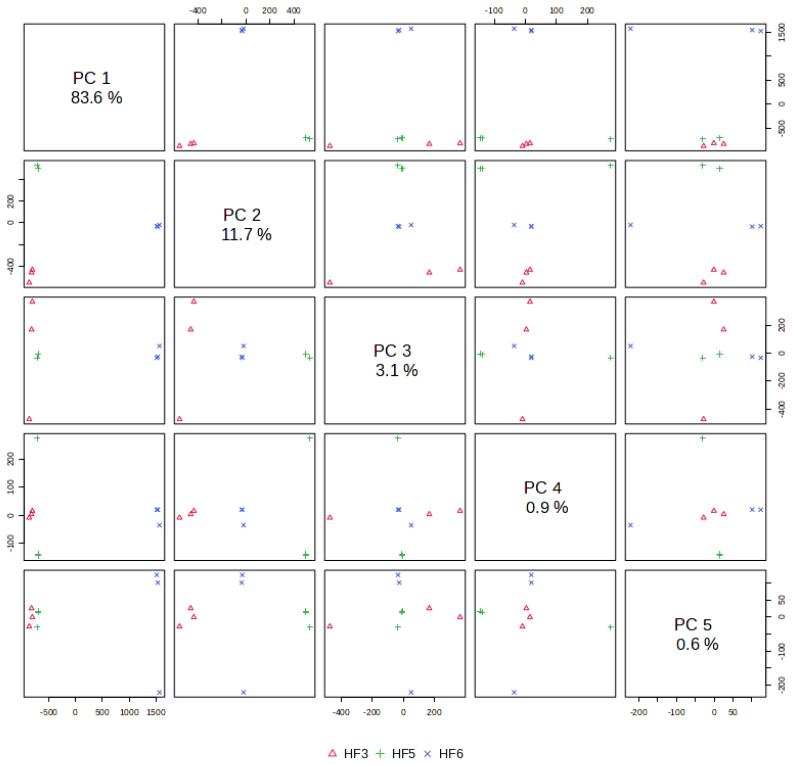


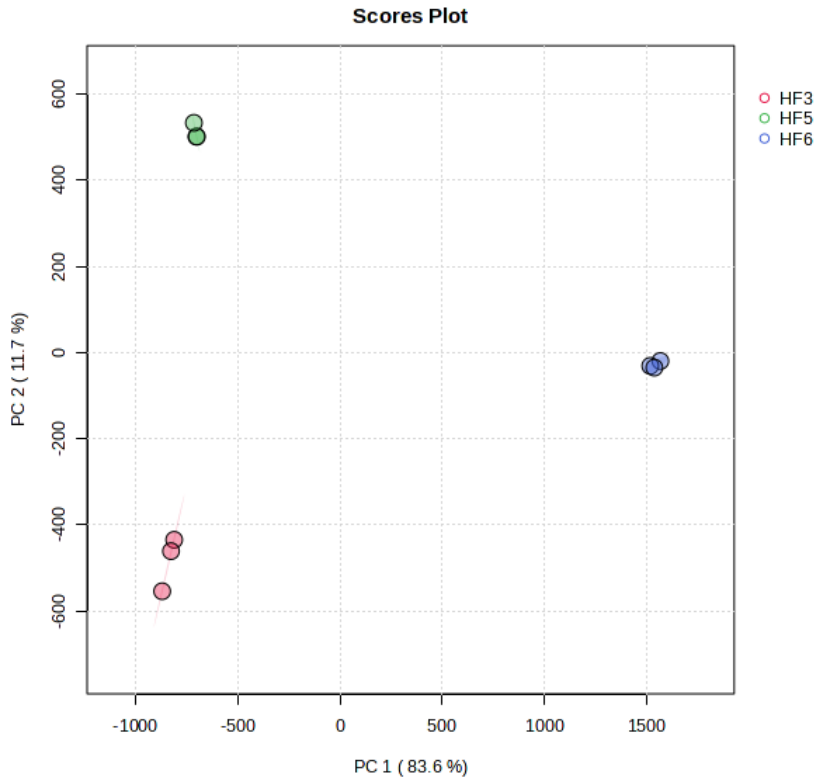
A1 m/z

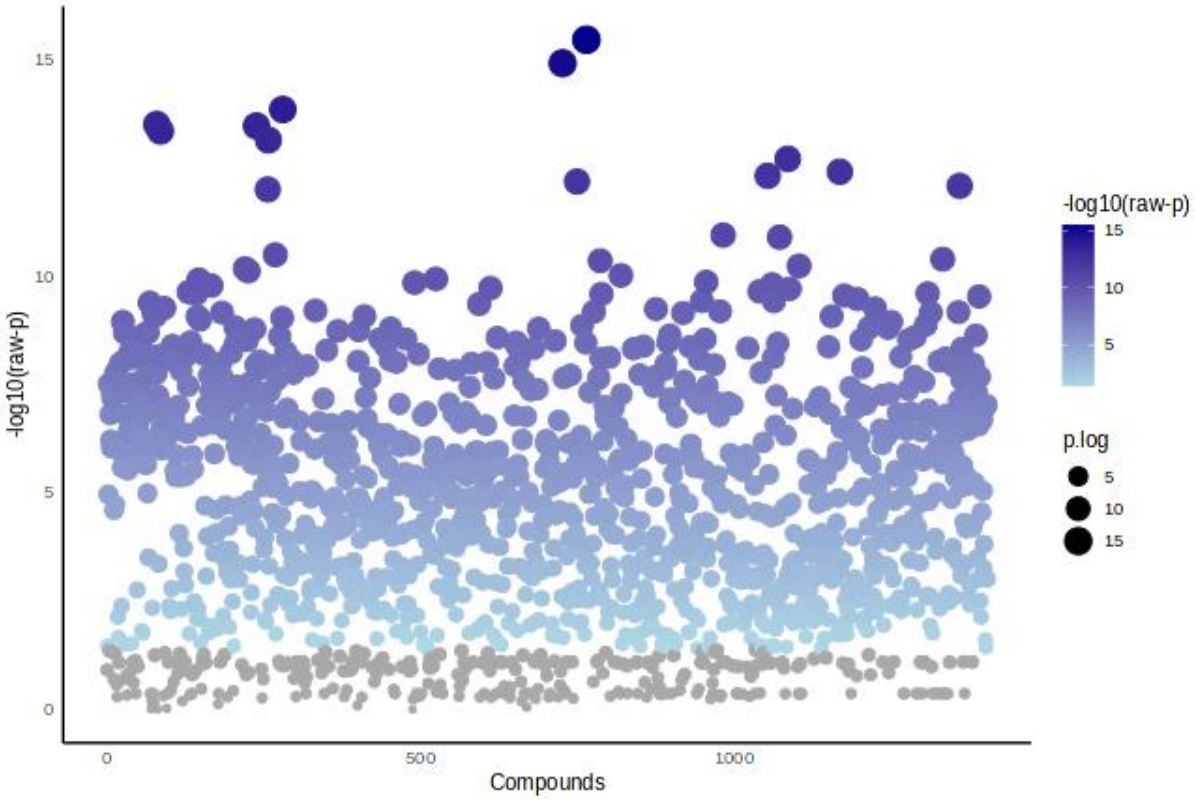
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD (top 1000)													
2	1105.982	1247100	1.3921e-1	16.856	1.9336e-1	HF6 - HF3; HF6 - HF5													
3	719.581	147750	8.3706e-1	14.077	5.8134e-1	HF5 - HF3; HF5 - HF6													
4	899.603	79816	5.3095e-1	13.275	2.4583e-1	HF5 - HF3; HF5 - HF6													
5	639.407	53451	1.7678e-1	12.753	6.1386e-1	HF5 - HF3; HF5 - HF6													
6	962.832	20901	2.9558e-1	11.529	8.2112e-1	HF6 - HF3; HF6 - HF5													
7	878.805	19349	3.7253e-1	11.429	8.624e-10	HF6 - HF3; HF6 - HF5													
8	682.493	14175	9.4734e-1	11.023	1.8798e-0	HF5 - HF3; HF5 - HF6													
9	746.598	12253	1.4668e-1	10.834	2.0788e-0	HF3 - HF5; HF3 - HF6; HF6 - HF5													
10	948.846	12238	1.4721e-1	10.832	2.0788e-0	HF6 - HF3; HF6 - HF5													
11	487.515	12166	1.4982e-1	10.824	2.0788e-0	HF6 - HF3; HF6 - HF5													
12	349.181	11790	1.6463e-1	10.783	2.0788e-0	HF6 - HF3; HF6 - HF5													
13	401.373	10954	2.0523e-1	10.688	2.3755e-0	HF6 - HF3; HF6 - HF5													
14	603.307	10478	2.3447e-1	10.63	2.4642e-0	HF6 - HF3; HF6 - HF5													
15	1170.022	10279	2.4837e-1	10.605	2.4642e-0	HF5 - HF3; HF5 - HF6													
16	980.841	9771.5	2.8913e-1	10.539	2.6773e-0	HF6 - HF3; HF6 - HF5													
17	647.33	8317.4	4.6873e-1	10.329	4.0692e-0	HF6 - HF3; HF6 - HF5													
18	681.303	8050.1	5.1699e-1	10.287	4.2241e-0	HF6 - HF3; HF6 - HF5													
19	879.807	7697.2	5.9136e-1	10.228	4.5633e-0	HF6 - HF3; HF6 - HF5													
20	557.425	6252.8	1.1029e-1	9.9575	8.0626e-0	HF5 - HF3; HF5 - HF6													



Hispanic Print 3 Data

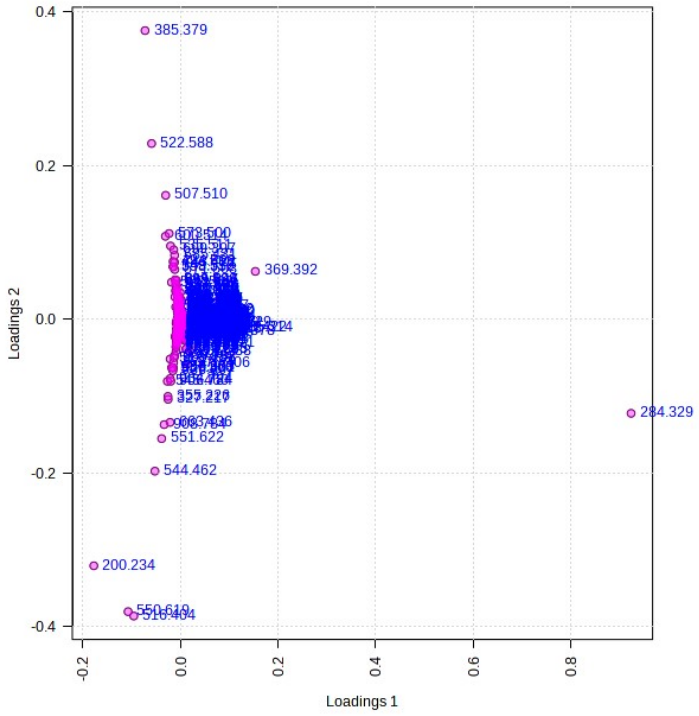



Scores Plot
HF3
HF5
HF6
PC 2 (11.7 %)
PC 1 (83.6 %)
600
400
200
0
-200
-400
-600
-1000
-500
0
500
1000
1500


-log10(raw-p)
Compounds
-log10(raw-p)
15
10
5
p.log
5
10
15
0
500
1000

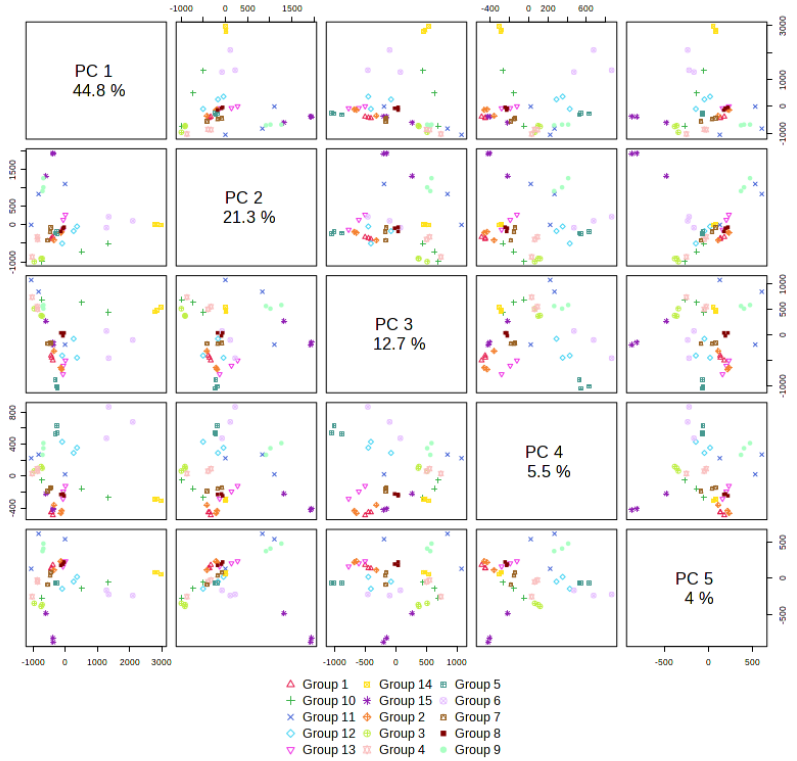
A1 m/z

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	m/z	f.value	p.value	#NAME?	FDR	Fisher's LSD (top 1000)													
2	733.571	425070	3.5154e-1	15.454	4.9356e-1	HF5 - HF3; HF6 - HF3; HF5 - HF6													
3	716.958	280410	1.2246e-1	14.912	8.5965e-1	HF5 - HF3; HF5 - HF6													
4	479.277	123610	1.4294e-1	13.845	6.6898e-1	HF5 - HF3; HF5 - HF6													
5	277.101	94894	3.1594e-1	13.5	9.6024e-1	HF6 - HF3; HF6 - HF5													
6	432.378	92423	3.4197e-1	13.466	9.6024e-1	HF6 - HF3; HF6 - HF5													
7	284.329	84286	4.5087e-1	13.346	1.055e-11	HF6 - HF3; HF6 - HF5													
8	450.365	71996	7.2341e-1	13.141	1.451e-11	HF5 - HF3; HF5 - HF6													
9	920.534	51574	1.9679e-1	12.706	3.4536e-1	HF5 - HF3; HF5 - HF6													
10	976.852	41073	3.8958e-1	12.409	6.0775e-1	HF6 - HF3; HF6 - HF5													
11	899.603	38227	4.8324e-1	12.316	6.7847e-1	HF5 - HF3; HF5 - HF6													
12	727.452	34494	6.5768e-1	12.182	8.3944e-1	HF6 - HF3; HF6 - HF5													
13	1142.859	31968	8.2622e-1	12.083	9.6668e-1	HF3 - HF5; HF3 - HF6													
14	449.363	30008	9.9893e-1	12	1.0788e-1	HF5 - HF3; HF6 - HF3; HF5 - HF6													
15	854.575	13376	1.1273e-1	10.948	1.1305e-0	HF5 - HF3; HF5 - HF6													
16	911.789	12895	1.2582e-1	10.9	1.1777e-0	HF6 - HF3; HF6 - HF5													
17	464.377	9441.7	3.2048e-1	10.494	2.8122e-0	HF5 - HF3; HF5 - HF6													
18	1115.804	8734.8	4.0472e-1	10.393	3.3425e-0	HF3 - HF5; HF3 - HF6													
19	744.44	8493.7	4.4016e-1	10.356	3.4333e-0	HF5 - HF3; HF5 - HF6													
20	930.798	7749	5.7958e-1	10.237	4.2828e-0	HF6 - HF3; HF6 - HF5													

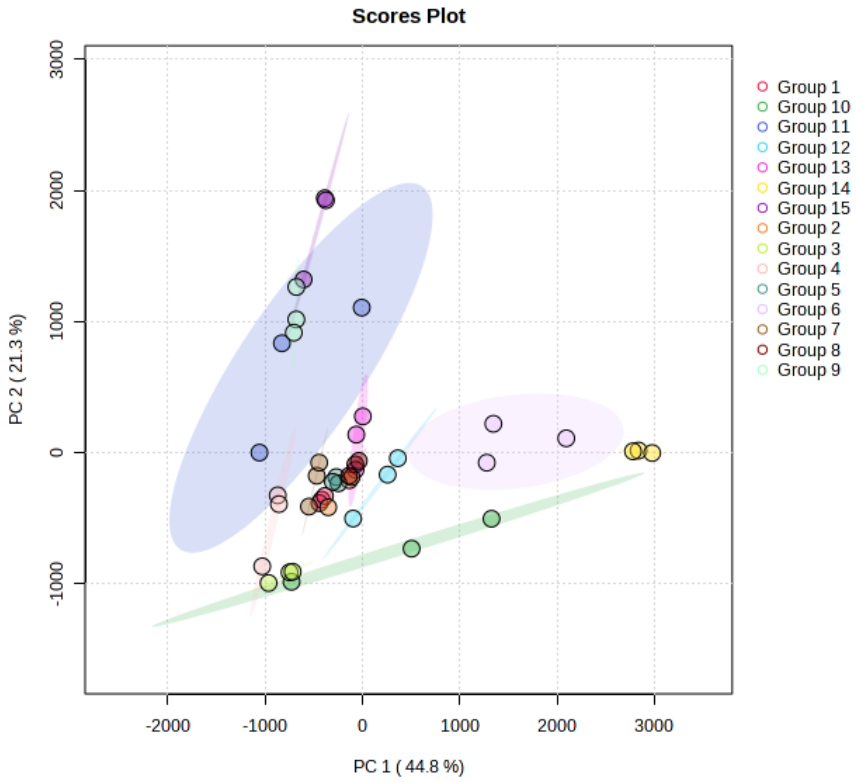


Different Geographical and Cultural Ancestry Data

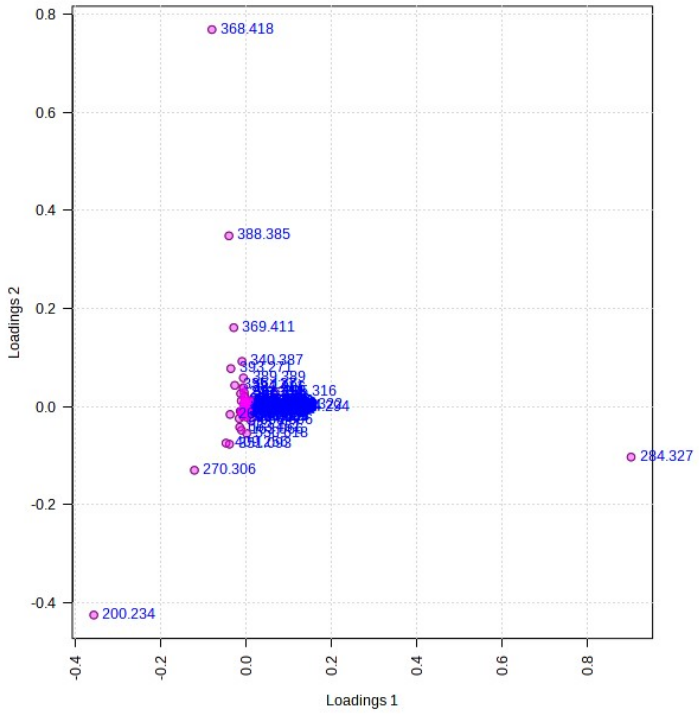
Print 1 Data



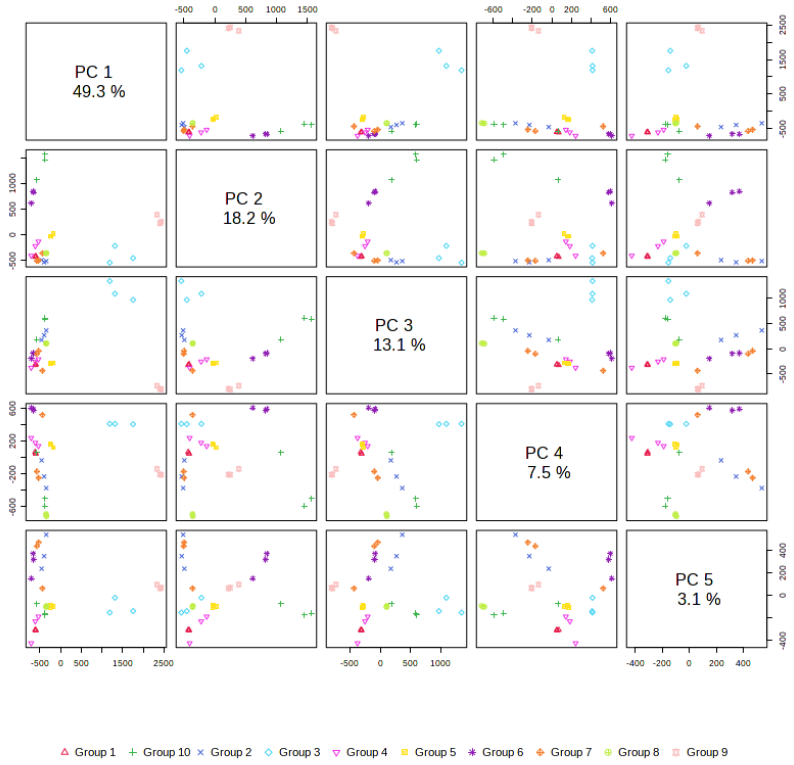
Key:
Group 1 – A1
Group 10 – H1
Group 11 – H2
Group 12 – H3
Group 13 – H5
Group 14 – H6
Group 15 – ME1
Group 2 – B1
Group 3 – E1
Group 4 – E2
Group 5 –E3
Group 6 – E4
Group 7 – E5
Group 8 – E6
Group 9 – E7



Key:
Group 1 – A1
Group 10 – H1
Group 11 – H2
Group 12 – H3
Group 13 – H5
Group 14 – H6
Group 15 – ME1
Group 2 – B1
Group 3 – E1
Group 4 – E2
Group 5 –E3
Group 6 – E4
Group 7 – E5
Group 8 – E6
Group 9 – E7



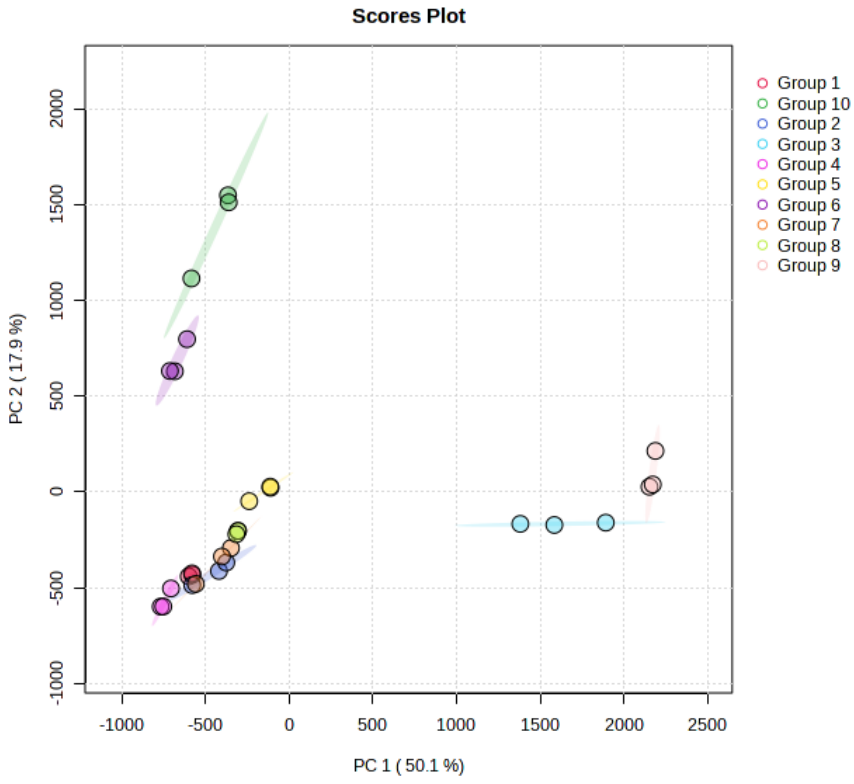
Print 2 Data



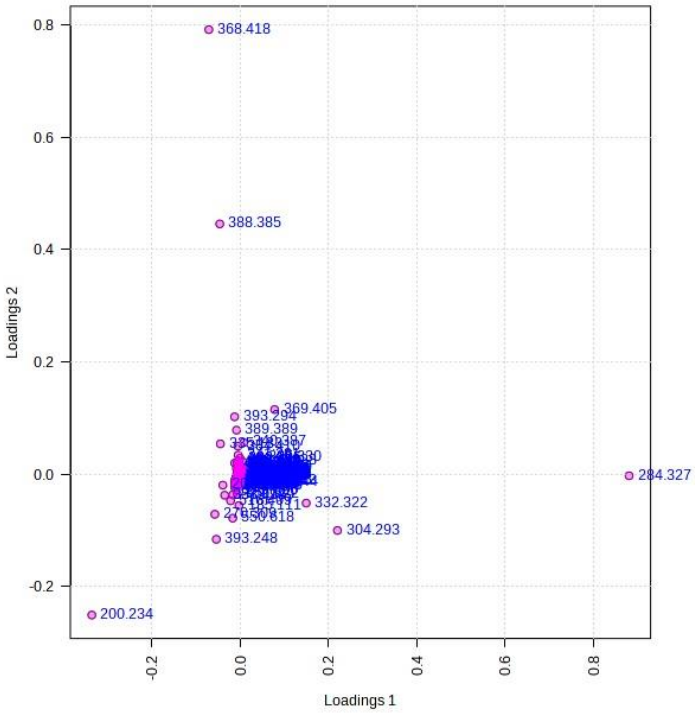
Key:
Group 1 – A1
Group 10 – ME1
Group 2 – E3
Group 3 – E4
Group 4 – E5
Group 5 – E6
Group 6 – E7
Group 7 – H3
Group 8 – H5
Group 9 – H6

Print 3 Data

Key:
Group 1 – A1
Group 10 – ME1
Group 2 – E3
Group 3 – E4
Group 4 – E5
Group 5 – E6
Group 6 – E7
Group 7 – H3
Group 8 – H5
Group 9 – H6



Key:
Group 1 – A1
Group 10 – ME1
Group 2 – E3
Group 3 – E4
Group 4 – E5
Group 5 – E6
Group 6 – E7
Group 7 – H3
Group 8 – H5
Group 9 – H6


368.418
388.385
369.405
393.294
389.389
284.327
332.322
304.293
393.248
200.234
Loadings 1
Loadings 2
0.8
0.6
0.4
0.2
0.0
-0.2
-0.2
0.0
0.2
0.4
0.6
0.8

Participant Demographic Survey

Participant Demographic Survey

It is imperative that you answer each question as accurately as possible. This information is necessary for the completion of this research project. As a reminder, you may withdraw from this study at any time. Your participation, however, is greatly appreciated.

1. Please circle the ancestry you **most** identify with:
   a) Hispanic Descendant
   b) Asian Descendant
   c) Native American Descendant
   d) African Descendant
   e) European Descendant
   f) Other

2. Please circle your **biological** sex:
   a) Male
   b) Female

3. How did you hear about this study?
   a) Email blast
   b) Professor
   c) Classmate
   d) Other (Please explain):

4. Were you offered extra credit for participating in this study?
   a) No
   b) Yes (Please list professor and class):

Thank you for your participation

For office use only

Participant Sample code:

Eligibility: