UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA

Forensic evidence from the biochemical composition of human fingerprints

Bárbara Filipa Alves Morgado

Mestrado em Bioquímica
Bioquímica

Dissertação orientada por:
Professor Doutor Carlos Cordeiro

2020

Acknowledgments

Ao concluir este trabalho, gostaria de agradecer a todos os que, de alguma forma, contribuíram para a sua execução.

Primeiro, gostaria de agradecer aos meus familiares, principalmente aos meus pais, Carlos Morgado e Rosa Alves, por me terem dado e proporcionado a hipótese de ingressar num curso superior e por me terem dado o seu apoio incondicional ao longo deste percurso académico. Queria também deixar um agradecimento especial à minha irmã, Daniela Morgado, por me ter aturado nesta fase e pela sua ajuda imprescindível. E, como não poderia faltar, queria agradecer aos meus avós, Luís e Julieta Morgado, por todo o amor e carinho que me dão todos os dias.

Agradeço também à minha grande amiga e colega, Joana Calé, que me acompanhou durante todo este percurso académico, estando presente tanto nos momentos bons, como nos maus. Sem ti não conseguiria ter chegado onde cheguei hoje. Por todas as experiências partilhadas e por toda a cumplicidade que criámos, agradeço-te do fundo do meu coração.

Gostaria também de agradecer ao meu orientador, Dr. Carlos Cordeiro, por todo o seu apoio, paciência e incentivo. O seu entusiasmo por todo este mundo motivou-me a realizar esta dissertação. Obrigada por todas as experiências que me proporcionou durante este percurso. Não só me transmitiram mais conhecimentos científicos como me fizeram crescer como pessoa.

Gostaria de agradecer também ao Dr. Peter O'Connor que me acolheu no seu laboratório durante 2 meses, através do programa ERASMUS+. Durante este período transmitiu-me imensos conhecimentos científicos, mas acima de tudo queria agradecer-lhe pela amabilidade com que me recebeu. Esteve sempre pronto a ajudar, tanto em assuntos mais formais como em assuntos do quotidiano, por exemplo quando enumerou os melhores restaurantes e locais turísticos de Coventry. Foi sem dúvida uma grande experiência e tudo foi possível graças a si e ao Dr. Carlos Cordeiro.

Este trabalho foi apoiado pela Rede Portuguesa de Espectrometria de Massa (LISBOA-01-0145-FEDER-022125), e pelo Projeto EU_FT-ICR_MS, financiado pelo programa de investigação e inovação *Europe and Union's Horizon 2020* ao abrigo do *grant agreement* nr. 731077.

Table of Contents

Index of Figures

Index of Tables

Summary

Throughout the years, fingerprints have been used in criminal investigations to place a suspect at a crime scene. However, this method depends on the integrity of the fingermark, on the method used for its recovery and if the respective fingerprint is listed or not on the National/International Fingerprint Database. The ridge pattern does not use all available information, most notably, the chemical signature that is left behind. So, it could be extremely beneficial to invest in techniques that can retrieve basic donor information in fingerprints' chemical composition.

Until now, human fingerprints have been mainly analysed by low resolution mass spectrometry methods. There is one particular area of mass spectrometry that has led the study of fingerprint's chemical composition and it is Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI MS). With this technique, not only a fingerprint's image is obtained, which may be used in a biometric identification, but it's also possible to profile the fingerprint's chemical composition. However, this type of approach does not allow a bulk analysis of its chemical composition.

With that in mind, this work coupled Electrospray Ionization (ESI) with a high resolution mass spectrometry method such as FT-ICR MS, to reveal the chemical composition of human fingerprints. The compositional space of 10 fingerprints was analysed combining accurate mass measurements, database search and molecular formula determination, which was explored trough Principal Component Analysis (PCA), Kruskal-Wallis tests and other tools.

From the thirty eight compounds commonly found in all fingerprints that were analysed, twenty (mainly lipids), such as palmitic acid, oleic acid, stearic acid and linoleic acid, were present with similar intensities. This was the first step towards the definition of a chemical background for the composition of human fingerprints in relation to which everything else can be found. Understanding the source of these compounds will be important for the interpretation of their possible roles on fingertips and potential uses in forensic and other contexts.

Despite the main focus being endogenous compounds, we also found several exogenous com-pounds, for example, pantothenol which is present in personal care products, mainly shampoos. It was possible to verify that in fact this compound was part of the composition of the shampoo used by the volunteer in question. Therefore, the use of ESI together with FT-ICR MS has proven to be a promising analytical method for studying the chemical signature of fingerprints due to its unmatched sensitivity for detecting any type of molecules, extreme sensitivity, requiring only minute amounts of sample and the fact that it is an high-throughput method. Thus, it can be used not only from a forensic point of view, but also in other contexts, namely, clinical.

Keywords: Forensics, Fingerprint, Mass spectrometry, FT-ICR MS, Endogenous Compounds, Exogenous Compounds, Lipids, Metabolomics

Resumo

As impressões digitais resultam da transferência do material presente na ponta dos dedos, mais precisamente nas cristas, para uma superfície, após contato. Ao longo dos anos, as impressões digitais têm sido utilizadas, maioritariamente, em investigações criminais para colocar um suspeito no local do crime. Embora os casos em que as impressões digitais são usadas como evidência para condenar um suspeito superem todos os outros casos combinados, há limitações que precisam ser consideradas ao estudá-las. A análise do padrão de cristas depende da integridade da impressão digital latente, do método utilizado para recuperá-la e se a respetiva impressão digital consta ou não de alguma base de dados nacional ou internacional. Esta abordagem não utiliza toda a informação disponível, principalmente a assinatura química que é deixada para trás. Portanto, pode ser extremamente benéfico investir no desenvolvimento de métodos analíticos que permitam a recuperação de mais informações criando uma impressão digital química que possibilita a obtenção de pormenores como, por exemplo, a idade, o género, os hábitos de vida do indivíduo, que são relevantes do ponto de vista forense.

Variáveis ambientais (por exemplo, luz, humidade e temperatura), bem como a duração do contacto entre a ponta do dedo e a superfície (cujo impacto não é totalmente compreendido) são exemplos de fatores que podem afetar o padrão das cristas aquando a deposição, o que pode impossibilitar a identificação do dono da impressão digital. Por outro lado, também existem fatores que podem afetar a integridade das impressões digitais mesmo antes destas serem sequer depositadas, como é o caso de algumas doenças (por exemplo, adermatoglifia) ou como acontece com a toma de certos medicamentos (por exemplo, capecitabina e paclitaxel).

As impressões digitais são constituídas maioritariamente por compostos intrínsecos, no entanto também possuem compostos de origem extrínseca. Os componentes intrínsecos (endógenos) podem incluir metabolitos (lípidos, aminoácidos, etc.), vestígios de medicamentos e drogas. A maior parte destes compostos têm origem na epiderme (queratinócitos) e no suor (sebócitos). Com a informação dos compostos intrínsecos, é possível traçar o perfil do suspeito uma vez que estes variam consoante as diferentes características do dador, incluindo o género, idade, dieta, raça, medicação, estado psicológico, saúde e metabolismo. Por outro lado, os compostos extrínsecos (exógenos) são substâncias que aparecem nas pontas dos dedos após o contato, por exemplo, sangue, pó, maquiagem, vestígios de alimentos, produtos de beleza (cremes, champôs, etc.), explosivos, entre outros. Caso este tipo de compostos apareça em impressões digitais latentes, podemos obter informações sobre a atividade do suspeito antes do crime.

Desta forma, compreender a complexidade do conteúdo intrínseco e extrínseco das impressões digitais pode ser extremamente útil para possibilitar a diminuição do número de suspeitos numa investigação criminal e para permitir o aperfeiçoamento das técnicas atualmente usadas na recuperação de impressões digitais latentes em cenários forenses, através da identificação de potenciais moléculas-alvo.

A espectrometria de massa é uma poderosa técnica analítica que determina a razão massa/carga (m/z) de iões. Pode também ser utilizada em estudos quantitativos ou mesmo para obter mais informações sobre a estrutura molecular. A sua sensibilidade única, o seu alto rendimento, o seu baixo limite de deteção e a suas diversas aplicações tornam a espectrometria de massa um método analítico excecional quando comparado a outros.

Até à data, a composição química das impressões digitais humanas tem sido analisada, principalmente, por métodos de espectrometria de massa de baixa resolução, fornecendo, desta forma, um número limitado de compostos identificados. Estes métodos focam-se, sobretudo, na obtenção de uma imagem da impressão digital em que a distribuição dos diversos compostos químicos se encontra assinalada. Para este efeito, são utilizados métodos como, por exemplo, anticorpos conjugados com partículas magnéticas, espectroscopia no infravermelho por transformada de fourier com reflexão total atenuada (ATR-FT-IR), espectrometria de massa com ionização de dessorção por electrospray (DESI-MS) e espectrometria de massa de tempo de voo por ionização por dessorção assistida por laser a partir de superfícies (SALDI-TOF-MS), entre outros.

No entanto, existe uma área particular da espectrometria de massa que tem liderado o estudo da composição química das impressões digitais. Esta é a espectrometria de massa por ionização por dessorção com laser assistida por matriz (MALDI MS). Com esta técnica, não só é obtida uma imagem da impressão digital, que pode ser usada para uma identificação biométrica, como também é possível traçar o perfil da composição química da mesma. Por outro lado, este tipo de abordagem permite apenas uma análise direcionada para um certo tipo de compostos, não permite uma análise global da amostra.

Com esta informação em mente, este trabalho acoplou a ionização por electrospray (ESI) a um método de alta resolução como a espectrometria de massa de FT-ICR para desvendar a composição química de impressões digitais humanas.

A espectrometria de massa de ressonância ciclotrónica de ião com transformada de fourier (FT-ICR MS) é baseada no movimento clássico de iões sob a influência mútua de um campo elétrico e um campo magnético. Os iões são produzidos externamente à fonte de iões, são acumulados na ótica de iões e, por fim, são transferidos para a célula ICR. Esta célula está localizada dentro de uma câmara de alto vácuo, sob o efeito de um campo magnético unidirecional, espacialmente uniforme e homogéneo (proveniente do magneto supercondutor) e sob o efeito de um campo elétrico (normalmente, originado por duas placas condutoras opostas que não permitem que os iões sejam perdidos nesta direção). A determinação da frequência ciclotrónica de um ião num campo magnético de força conhecida leva diretamente à determinação da sua razão massa-carga.

Com este método, o espaço composicional químico de 10 impressões digitais humanas foi revelado através da combinação de medidas de massa precisas, pesquisas em bases de dados (como *Human Metabolome Database*, HMDB) e a determinação das fórmulas moleculares. Esta informação foi explorada por meio de uma Análise de Componentes Principais (PCA), testes de Kruskal-Wallis, entre outras ferramentas.

Dos trinta e oito compostos endógenos encontrados em todas as impressões digitais, vinte (principalmente lipídios), como ácido palmítico, ácido oleico, ácido esteárico e ácido linoleico, estavam presentes em intensidades semelhantes. Compreender a origem destes compostos será importante para a interpretação das suas possíveis funções na ponta dos dedos, assim como para potenciar a sua utilização em contextos forenses e não só. Neste sentido, este foi o primeiro passo para a definição de uma composição química standard para as impressões digitais humanas.

No entanto, apesar do foco principal ser a análise e identificação dos compostos endógenos, também conseguimos encontrar alguns compostos exógenos, como serve de exemplo o pantotenol. Este composto está presente em produtos de higiene pessoal, principalmente champôs. Foi possível constatar

que, de facto, este composto fazia parte da composição química do shampoo utilizado pelo voluntário em questão, no dia em que foi detetado. Portanto, o uso do ESI como método de ionização juntamente com a espectrometria de massa de ressonância ciclotrónica de ião com transformada de fourier, provou ser uma promissora abordagem analítica para conseguir determinar a assinatura química de impressões digitais humanas visto que possui uma sensibilidade inigualável para detetar qualquer tipo de moléculas, consegue detetar pequenas quantidades de amostra e, não menos importante, é um método com uma performance bastante rápida. Deste modo, esta abordagem pode vir a ser muito útil não apenas do ponto de vista das aplicações forenses, mas também noutras áreas, como por exemplo em contexto clínico.

Palavras chave: Forense, Impressão digital, Espectrometria de massa, Espectrometria de massa de ressonância ciclotrónica de ião com transformada de fourier, Compostos endógenos, Compostos exógenos, Lípidos, Metabolómica

Materials & Methods

Chemicals and reagents

Methanol (Merck, LC-MS hypergrade), acetonitrile (Merck, LC-MS grade), water (Type I), formic acid (Sigma Aldrich, MS grade), leucine enkephalin (LeuEnk) (YGGFL Sigma Aldrich) and glass slides which were microscope slides with ground edges frosted (50 pk, ECN 631-1553, LOT 032814).

Subjects

A total of 10 volunteers participated in this study to determine the chemical composition of human fingerprints, between October 2019 and February 2020. Subjects were recruited from the Fourier-Transform Ion-Cyclotron-Resonance and Structural Mass Spectrometry Laboratory (FT-ICR-MS-Lisboa) in the Faculty of Sciences, University of Lisbon. The volunteers were 4 males and 6 females with ages between 20 and 50 years. All individuals were asked to fill the Identification file (Capítulo 1Appendix A) and identities were anonymized. All provided written informed consent. No personal data or samples were kept beyond the duration of this study.

Chemical composition of Fingerprints

Volunteers were asked to press a glass slide with the index finger for approximately 10 seconds. The fingerprint was suspended with 20uL of extraction solution (1:1:1 Water : Acetonitrile : Methanol*, 1 mL* + formic acid (0,1%)) and a 1:100 dilution was prepared with the electrospray solution (1:1 Methanol : Water*, 1 mL* + formic acid (0,1%)). Formic acid was added to the samples to potentiate positive ionization. 0,2µL of LeuEnk (0,1 mg/mL) was added to all samples as internal standard for internal mass calibration. Thus, all collected mass spectra are accurate to at least 1 ppm.

Samples were loaded, positively ionized using an Electrospray Ion Source (ESI^+) and analysed with a 7 T Solarix XR FT-ICR Mass Spectrometer (FT-ICR MS, Brüker Daltonics).

Data were acquired in magnitude-mode with a 4MB transient, 50 scans, 0.02 seconds accumulation time, 0.8 seconds ToF, in a mass range between 200 and 1000 m/z. For each sample, per day, 3 spectra were acquired. This procedure was repeated 4 times on different days with freshly collected fingermarks.

Due to the COVID-19 pandemic during the period of this dissertation, it was not possible to find a way to safely collect samples. As a result, no more samples were gathered as of February 2020.

Data Analysis

The data were extracted using DataAnalysis 5.0 (Brüker Daltonics) with a signal to noise threshold of 6. The results were then analysed in MetaboScape 4.0 (Brüker Daltonics) to assigned mass lists to the corresponding compound using the *Human Metabolome Database* (HMDB). Spectral data was aligned within 1 ppm and the resulting bucket table was exported to Microsoft Excel (16.16.27 version) and the list manually analysed to determine the common ones within the annotated compounds. The following premises were considered:

- The compound is considered common if it is present in at least 8 out of 10 individuals;
- The compound is considered present in an individual, in one day, if it is at least in 2 replicates;
- If the number of days that a specific fingerprint was analysed is pair than a compound is considered present in that individual if it is in at least half the days;
- If the number of days is odd, then it needs to be present in more than half the days.

At the end, the mass deviation in ppm was calculated to all the compound determined to be common, applying the following equation:

$$Deviation\ (ppm) = \frac{m/z - m/z_{theoretical}}{m/z_{theoretical}} \times 10^6 \qquad \textit{(Equation 1)}$$

Being m/z the value of the peak in our sample, and $m/z_{theoretical}$ the value of m/z calculated through the chemical formula in the extension of DataAnalysis Compass IsotopePattern.

Principal Component Analysis (PCA)

To study the variability of all fingerprint samples, a PCA was performed using MetaboScape. The following parameters were used: no scaling algorithm, cross validation mode of 10% and minimum explained variance of 98

Statistical Analysis

To statistically analyse the results, Microsoft Excel was used. Considering that the size of each group is very small and are not equal, the one-way ANOVA's assumptions are difficult to establish. Instead, a non-parametric and more conservative test was performed – Kruskal-Wallis (Appendix B - ***Table B.1***).

As multiple tests have been done with the same data, one for each compound, a correction for multiple tests was required. In this case, the correction of Benjamini-Hochberg was performed (Appendix B - ***Table B.2***).

Chapter I : Introduction

Fingerprints

Fingerprints result from the transfer of material present in the raised portion of the friction ridge fingertip skin to a surface upon contact. [1] This pattern of ridges and structural units, *minutae*, is formed in the womb during the 3rd – 4th month of embryo development. [1][2] There, embryos are exposed to different pressures and conditions which will lead to fingerprints' unique pattern. This individual characteristic do not naturally change throughout one's life and that's why fingerprints have been used worldwide for biometric identification based on matching the pattern of ridges. [2] Some examples where fingerprint recognition is used are the following: criminal investigations, civilian investigations (e.g. border security), pre-employment screening and new biometric based security technologies (e.g., cell phones, workplace fingerprinting access devices on entry to facility/rooms, etc.). [3]

When the print is obtained, for example, in a crime scene the identity of the donor is unknow and in this case it's not called fingerprint but fingermark. In this situation a fingerprint is used as a control to place a suspect in that crime scene. [1]

I. Fingerprint Applications in Biometric Identification

On June 29, 1892, Francesca Rojas stabbed to death her two children (six-year-old boy and four-year-old girl) in their home in Necochea, Argentina. Francesca tried to blame the murders on their neighbour Velasquez, saying he threatened her when she rejected his sexual advances. Police arrested and questioned Velasquez using rather unconventional interrogation methods, in order to force a confession. However, he maintained his innocence. [4]

Until then, the only method of identification was the Bertillonage, named after its inventor, Alphonse Bertillon, and involved the recording of body measurements in more than 11 different places. [4]

When Rojas' house was examined a bloody thumb print was found on the bedroom door. Investigators requested Rojas' fingerprint and even though the understanding of forensic identification was rudimentary, they were able to determine that the fingermark on the door belonged to Rojas. [4]

Seemingly, Rojas had killed her own children because she wanted to increase her chances of marrying her boyfriend who apparently disliked her children, and then pegged the crime on Velasquez. She was sentenced to life imprisonment. [4]

It wasn't until 1893, that fingermarks were officially accepted as evidence to use in court to convict a suspect. Nevertheless, fingerprints were first accepted by U.S. courts as a reliable means of identification on December 21, 1911. [5]

From this moment on emerged the concept of Dactyloscopy. The word is derived from the two Greek words 'daktylos' meaning 'finger' and 'skopein' meaning 'to examine'. It deals with fingerprints as material evidence, especially with their analysis and comparison. [6] Since this date, fingermarks have

been considered one of the most important physical evidence used in criminal investigations and the most commonly used forensic evidence worldwide. [7] There are alternative biological specimens belonging to the suspect that may be retrieved but it's not always possible to obtain them. For example, it's possible to find DNA in some fingermarks but swabbing the mark is undesirable because the integrity of the evidence gets compromised and the amount of DNA that is actually obtained is limited and often insufficient to obtain any relevant information. [1]

For a fingermark to be enhanced effectively, there are 3 aspects that need to be taken in consideration when selecting the appropriate fingermark enhancement technique: the deposition surface, the environment that the mark was exposed to and the probable age. The 'Manual of Fingerprint Development Techniques' outlines a sequential guide for the enhancement of fingermarks found in different scenarios. [8]

II. Fingermarks Limitations

Although the cases where fingermarks are used as evidence outnumber all other forensic casework combined, there are limitations that need to be considered when studying fingermarks. [7] Not all fingermarks recovered from a crime scene can be used for suspect identification specially if the respective fingerprint is not listed on the National or International Fingerprint Database or if the fingermark integrity is damaged. [8]

i. Factors That May Alter The Pattern Of Ridges Upon Deposition

Environmental variables as well as the time of deposition (the impact of this factors is not fully understood yet) are examples of factors that may change the pattern of ridges upon deposition which makes identification of the donor impossible. [7] Also, there are physical and chemical factors that may damage fingermarks integrity prior to deposition.

ii. Factors That May Alter The Pattern Of Ridges Prior Deposition

<u>Diseases</u>

Several diseases impair the skin at the fingertip leading to a change in the ridge pattern. Acanthosis nigricans, eczema, dry and atrophic skin, scleroderma, celiac disease (temporary modification) and leprosy are some examples of diseases that may modify fingerprints. In patients with rickets and acromegaly, the pattern of ridges isn't affected, but the distance between the ridges is. [3]

Absent fingerprints, or adermatoglyphia, is a very rare medical condition. Affected persons have completely smooth fingertips, palms, toes and soles, but no other medical signs or symptoms. [9] A 2011 research indicates that this disease may be caused by a mutation in the gene SMARCAD1 near a key splicing site that prevents SMARCAD1 from being made correctly. [10]

Medication

Regarding medication, several patients have reported loss of finger ridges after a long-term external corticosteroids treatment for eczema, and after treatment with some chemotherapy agents such as capecitabine and paclitaxel. [3]

Paclitaxel is an anti-microtubule chemotherapy agent used in the treatment of ovarian, breast and lung cancers. In this case, the fingerprint loss may result from the hand-and-foot syndrome (HFS). It causes erythema on the palms and soles, dysesthesia and/or paresthesia. The symptoms can evolve to blistering, desquamation and ulceration. Nevertheless, fingerprint changes may be independent from HFS occurrence and without any correlation with dose and duration of the therapy. [3]

III. Chemical Composition of Fingermarks

With all the limitations the ridge pattern comes with, there has been a great interest in techniques that can still retrieve basic donor information from these kinds of fingermarks. So, it could be extremely beneficial to obtain additional information, besides the ridge pattern, most notably their chemical composition.

Fingerprints are composed of substances from the epidermis and the secretory glands in the dermis. [7] The compounds present can be from three types of secretory glands:

- Eccrine glands – they are found all over the body but are abundant on the palms. Eccrine secretions consist of water, inorganic salts (such as chloride, bromide, iodide, fluoride, and phosphate), and water soluble organic species (such as amino acids, proteins, fatty acids, and urea). [7][11]–[13] Their main function is thermoregulation during exposure to a hot environment or during physical exercise. [14]

- Sebaceous glands – they're found all over the body, except on the friction ridge surfaces of the hands and the feet. They are mainly abundant on the face, scalp and neck. Sebaceous secretions (sebum) are found on the fingertips due to people's habit to touch their hair and their face. [13] Sebum is a mixture of fatty acids, glycerides, long chain fatty acid esters, squalene, sterols (such as cholesterol), waxes and numerous lipid esters. [1][15][16] Functional attributes of sebum include the maintenance of a protective emollient lipid film, retardation of water evaporation, and shielding against bacterial and fungal infections. [17]

- Apocrine glands – are found primarily in axillae and anogenital areas of humans. [18] Apocrine secretions consist of proteins, carbohydrates, cholesterol and iron. [7] These glands respond to emotional stimuli such as pain, fear, sexual stimulation and anxiety. [19]

Apart from these compounds, fingerprints also possess intrinsic (endogenous) and extrinsic (exogenous) substances. Both components can vary between individuals (inter-individual variability), as well as from the same individual from day to day (intra-individual variability) or at different times on the same day. [20]

i. Intrinsic Compounds

Intrinsic components may include metabolites (lipids, amino acids, etc), traces of medication and drugs. [7] With the intel of intrinsic compounds it's possible to profile the suspect because they vary between donors due to different donor features including gender, age, diet, race, medication, psychological state, health and metabolism. [21]

Gender Differentiation

There have been some efforts to determine the differences in fingerprint composition due to gender differences and some experiments show that fatty acids tend to be present in higher concentrations in male donors (such as saturated C15, C16 and C17 acids) and sterols, sterol esters, amino acids (such as alanine, glycine and serine) tend to be more concentrated in female donors. However, none of these differences were considered statistically significant. [12][15][22]–[24]

Another study, which revealed 85% accuracy, demonstrated that there might be differences in small peptides, such as SSL-29, LEK-45 and DCD-1L, that may be used as potential biomarkers of gender. SSL-29 and LEK-45 appeared to be important in the molecular profile for the detection of the male sex, and DCD-1L for the female sex. [1]

Age Differentiation

Regarding age differentiation, fingerprint composition can alter significantly between birth, puberty and old age. [12] Prior to puberty fingerprints are composed of eccrine related compounds (cholesterol and cholesterol esters (epidermal lipids), straight chain fatty acids, long chain fatty esters), while fingerprints from more developed donors contain sebaceous components (branched fatty acids, squalene, wax esters, while the proportion of exogenous lipids, such as cholesterol, Δ9 fatty acids, and linoleic acid decreases). [21][25][26] With the onset of puberty, androgens will stimulate the sebaceous glands which will lead to an increase in glandular activity (increased synthesis and secretion of sebum) and glandular hyperplasia. As age begins to increase, the hormonal stimulation starts to diminish and the sebaceous gland activity decreases as well. [17] These differences in the rate of sebum secretion will affect fingerprint longevity as a child's fingerprint (composed of volatile compounds mostly) may completely disappear in 48 h, while an adult fingerprint can last a week. [27]

ii. Extrinsic Compounds

Extrinsic Compounds are substances that have contaminated the fingertips following contact, for example, blood, dirt, make-up, food contaminants, moisturisers, skin care products, explosives, drugs of abuse etc. [7][8][21][28]–[32] Accessing the extrinsic compounds found in fingermarks we can obtain information about the suspect's activity prior to the crime.

Understanding the complexity of the intrinsic and extrinsic content of fingerprints could be extremely useful in allowing the narrowing in the pool of suspects in a criminal investigation and in the further development of enhancement techniques currently used in forensic scenarios by identifying potential target molecules. Moreover, it paves the way for further application of fingermark chemical composition analysis beyond forensics.

IV. Factors That Affect Chemical Composition of Fingermarks

Although we can get more information through the chemical composition of the fingerprint than by just analysing the pattern of ridges, there are factors that influence fingerprint composition as well, like the deposition conditions, the substrate and the environment. These factors play an important role in 2 main stages: the deposition stage and the aging stage. [21]

i. Deposition Stage

In the deposition stage, fingerprint composition is affected by deposition conditions (such as the deposition action, angle, pressure, duration of the contact, dimension of the digit, the finger used), nature of the substrate (porosity, curvature and surface texture) and donor features (such as age, gender, race and diet). It is important to have in mind that if the donor washed his hand near the deposition time, the concentration of compounds present in his fingerprints will change. [7][15][21]

ii. Aging Stage

The aging stage initiates after deposition. In this stage fingerprint composition is influenced by the nature of the substrate (such as texture, physico-chemical structure and electrostatic forces), environmental conditions (such as light, temperature and humidity), the enhancement technique used and the time elapsed since deposition (the longer this stage is, the greater the degradation of fingerprint's components is). [21] At this stage, the substrate has a bigger impact in fingerprint composition than at the deposition stage because chemical reactions between fingerprint's components and the substrate may occur, for example metal corrosion by ionic salts. [33] This might lead to the disappearance of compounds of interest or the appearance of compounds that were not initially there which may give rise to false information. Nevertheless, these chemical changes offer the opportunity to relate time with chemical composition thus offering the potential for fingermark dating

Therefore, the final composition of the fingerprint is a combination of all the factors that may influence both deposition and aging stages. [7]

V. Analytical methods for fingerprint analysis

When analysing the compounds present in fingerprints there are some requirements the technique used must possess: has to be sensitive to a range of compounds and has to detect reproducible differences between donors. [34]

To date, there have been attempts to determine the chemical composition of fingerprints using a plethora of advanced analytical techniques, including: Gas Chromatography-Mass Spectrometry (GC-MS), Liquid Chromatography-Mass Spectrometry (LC-MS), Attenuated Total reflectance Fourier Transform Infrared Spectroscopy (ATR-FT-IR), Raman Spectroscopy, Desorption Electrospray Ionisation Mass spectrometry (DESI-MS), Gold Nanoparticle-Assisted Laser Desorption Ionisation Mass Spectrometry, Antibody-Magnetic Particle Conjugates, Secondary Ion Mass Spectrometry (SIMS), SDS-PAGE and immuno-detection, Direct Analysis in Real Time Mass Spectrometry (DART-MS), Surface-Assisted

Laser/Desorption Ionisation Time-of-Flight Mass Spectrometry (SALDI-TOF-MS) and Laser Desorption Ionization-Mass Spectrometry (LDI-MS). [34]

Each method has different capabilities in terms of sensitivity, selectivity and reproducibility. Previous studies demonstrated that techniques like MALDI and ATR-FT-IR are more sensitive to endogenous compounds in fingerprints. Whereas ToF-SIMS and MeV-SIMS are sensitive to exogenous compounds in fingerprints. [34]

Mass Spectrometry

The vast majority of analytical methods already applied to study the chemical composition of fingerprints were Mass Spectrometry (MS) based.

Mass spectrometry is a powerful analytical technique that measures the mass-to-charge ratio (m/z) of ions. Not only that, but it can be used in quantification studies or to obtain more information on molecular structure. Its unequalled sensitivity, high-throughput, low detection limits and diversity of applications make mass spectrometry an outstanding analytical method. [35]

Since 1912, when the first mass spectrometer was constructed, mass spectrometry continues to evolve rapidly. New instruments were developed, new hybrid instruments were crafted combining several analysers or other techniques (such as HPLC and GC), existing analysers were perfected and new ones were developed (such as the Orbitrap). [35] This led to the development of new applications. Some applications include: study of biochemical problems (such as proteome and metabolome), food control, reaction physics and kinetics, ion-molecule reactions, drug testing, and many others. Recently, a study was made in order to introduce this technique in operatory blocks. The aim was to distinguish healthy tissue from cancerous tissue using ambient ionization MS to obtain molecular information in real time, directly from tissue samples. [36] Another example of the application of MS in real-time medical diagnostics was the distinction made between cystic fibrosis patients and healthy patients through the PTR-ToF-MS analysis of their tidal breath. [37]

Mass spectrometers analyse ions in the gas phase to obtain the mass-to-charge ratio (m/z). [38] Therefore they all should be able to perform the following steps:

1. Produce ions of the compound in the ionization source (Electron Ionization, Chemical Ionization, Laser Desorption, MALDI, Atmospheric Pressure Ionization, ESI, etc.);
2. Separate the ions according to their mass to charge ratio in the mass analyser (Quadrupole, Ion Trap, Orbitrap, ToF, FT-ICR, etc);
3. If necessary, fragment the selected ions and analyse the fragments in a second analyser;
4. Detect the ions in proportion to their relative abundance;
5. Process the signal from the detector in order to produce a mass spectrum, which is a plot of ion abundance versus mass-to-charge ratio. [35]

I. Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI MS)

There is one particular area of mass spectrometry that has lead the study of fingerprint's chemical composition and that is Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI MS). [39] With this technique, not only a fingerprint's image can be obtained, which may be used in a biometric identification, but it's also possible to profile the fingerprint's chemical composition.

Since the initial application of this technique in a forensic investigation for the analysis of fingermarks [40], a range of applications have now been established, including the identification of condom lubricants that have been used to contaminate marks (highly relevant in a possible sexual assault case) [41][42], separation of overlapping fingermarks [43], the confirmation of blood residues in marks [44], detection of drugs and excreted metabolites [45], the detection of explosives [46], prediction of sex from statistical analysis of excreted peptides [1], and the development of forensically applicable workflows [47].

In MALDI the compound of interest is dissolved in an organic matrix. The analyte molecules are embedded throughout the matrix and are isolated from one another. Irradiation of the solid solution by an intense laser pulse (under vacuum) induces rapid heating of the solution by the accumulation of a large amount of energy in the condensed phase through excitation of the matrix molecules. This causes localized "desorption" or "ablation," depending on the laser energy. The transition from solid to gas may be smooth (desorption) or it may occur a "phase explosion". The condensed material ejected can then be captured on a cold plate and directly imaged or measured. Nevertheless, MALDI mechanisms are not completely understood yet. [48]

Normally, in positive ion mode, MALDI spectra include mostly monocharged molecular species by protonation. Therefore, MALDI spectra are simple. Compared with other ionization techniques, MALDI is relatively less sensitive to contaminations (such as by salts, buffers, detergents). [48]

However, this technique has some disadvantages such as:

- Low shot-to-shot reproducibility – Each laser shot ablates a few layers of the sample at the spot where the laser irradiates which will produce variations in the spectrum;
- Strong dependence on the sample homogeneity (preparation method) – the impact position on the surface of the deposit may lead to spectral variations and a low reproducibility;
- As the sample is embedded in the matrix, the spectrum will have peaks that correspond to the matrix, mainly below 1000 m/z, this is why MALDI is principally used to analyse proteins.
- When analysing small molecules, MALDI is generally coupled with TOF analysers which results in a far lower resolution and lower mass accuracy compared to FTMS instruments. [48]

II. Electrospray Ionization (ESI MS)

The other major ionization technique is Electrospray Ionization (ESI). ESI emerged as a good technique to produce intact ions from polar macromolecules present in complex mixtures, most notably, proteins. [48]

ESI is produced by applying a strong electric field, under atmospheric pressure, to a liquid passing through a capillary tube with a weak flux. There are three stages in the production of gas-phase ions from electrolyte ions (such as sodium acetate) in solution:

1. Production of charged droplets at the ES capillary tip – a potential difference is applied between the capillary and the counter-electrode to create an electric field which will induce a charge accumulation at the liquid surface at the end of the capillary. This will break and form highly charged droplets (bearing the electrolyte on the surface and carrying the sample within);

2. Shrinkage of the charged droplets by solvent evaporation and repeated droplet disintegrations leading ultimately to very small highly charged droplets capable of producing gas-phase ions – As solvent is stripped away in the vacuum (heated inert gas, most often nitrogen), the droplet size decreases, and sample-rich particles result, often bearing high numbers of charges;

3. The actual mechanism by which gas-phase ions are produced from the very small and highly charged droplets – ions may be clustered with solvent molecules/other additives so they are subjected to a thermal declustering in the heated capillary leading to the partial vacuum of the first chamber and collisional activation due to an electric potential difference imposed between the sampling capillary exit and the skimmer leading to the second, high-vacuum chamber that is the housing of the mass spectrometer. [48]

Nevertheless, ESI also has some limitations:

- Complexity of the spectrum – Contrary to the MALDI spectra, ESI is able to produce multiply charged ions from large molecules, potentially producing more crowded spectra; [48]
- ESI is sensitive to concentration rather than to total amount of sample – ESI current is limited by the electrochemical process that occurs at the probe tip (formation of ions); [35]
- ESI is very sensitive to contaminants, unlike MALDI. [48]

III. Mass Accuracy and Resolving Power

The key parameter in MS is mass accuracy which is the difference between the theoretical mass of an ion and its mass measured by the mass spectrometer. [38] The benefit of measuring a compound's mass with adequately high accuracy is the possibility of directly determine its elemental composition. Accurate mass thus acts as a powerful "filter" useful for confirming the identity of a compound or even identification of an unknown one. [49]

However, mass accuracy depends upon the mass Resolving Power (R) and Signal-to-Noise ratio (S/N). The resolving power it's the instrument ability to distinguish two adjacent peaks of equal intensity. An accurate mass measurement can only be performed if the peak of interest is well resolved and distinguished from others. [38] In this context, the resolution settings used in the analysis define a confidence interval (mass tolerance) within which we can trust our mass measurements. [49] The overlap of such co-eluting peaks of a very similar mass will skew the peak shape of our compound of interest, and the peak centroid will no longer correspond to its accurate mass. [49]

Many types of modern mass spectrometers (quadrupole, time-of-flight) offer the mass accuracies in the level of at best tens of ppm in a limited mass range, whereas FT-ICR and Orbitrap analyzers can achieve mass accuracies in sub-ppm, or even ppb (parts per billion) level in state-of-the-art instruments. [38] With these analyzers a high-resolution spectra is obtained from ions formed by practically every ionization method, [50] to perform tandem mass spectrometric measurements [51] and to examine ion chemistry [52] and photochemistry. [49] With such extraordinary performance, the ion's elemental composition can be revealed without tandem MS experiments. [38][49]

IV. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS)

Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) is a versatile instrument that can be adapted to a variety of analytical and physical chemistry measurements. Its versatility follows from the fact that it is an ion trapping instrument. Hence, quench, ion injection, excitation and detection are operated in the same space (in this case ICR cell), but spread out in time. To control the FT-ICR hardware a "pulse program" is used (***Figure 1.1***). [49]

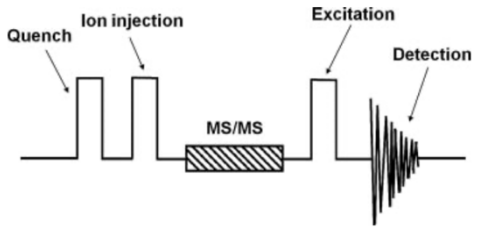

Figure 1.1. *An example of a simple pulse program that can be used in a FT-ICR MS experiment. [38]*

FT-ICR MS is based on the classical motion of ions under the simultaneous influence of electric and magnetic fields. [53] Ions are produced externally from an ion source, accumulated in the ion optics, and transferred into the ICR cell. [38] The ICR cell is located inside a high-vacuum chamber within a spatially uniform unidirectional and homogeneous magnetic field, from the superconducting magnet, and an electric field, typically from two opposed conductive trapping plates. [38][54]

Figure 1.2. *7 T Solarix XR FT-ICR Mass Spectrometer present in Fourier-Transform Ion-Cyclotron-Resonance and Structural Mass Spectrometry Laboratory (FT-ICR-MS-Lisboa).*

When a particle of mass m, charge q, and initial velocity v experiences a "Lorentz Force"(Magnetic force , $\vec{F}$) in a magnetic field, $\vec{B}$, causes it to follow a spiral path (***Equation 1.1***).

$$\vec{F} = q\left(\vec{v} \times \vec{B}\right) \qquad (\textbf{\textit{Equation 1.1}})$$

The projection of this path onto a plane perpendicular to the magnetic field is a circle (***Figure 1.3***). The frequency of this circular (or cyclotron) motion is given by ω_c, so that each ion of a given mass-to-charge ratio has a characteristic cyclotron frequency. [53]

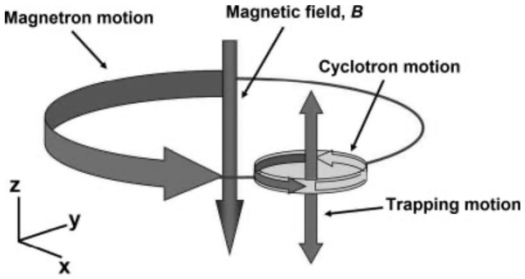

Figure 1.3. *Representation of the three motional modes of an ion: cyclotron rotation, magnetron motion and trapping motion. [38]*

Consequently, the determination of an ion's cyclotron frequency in a magnetic field of known strength leads directly to knowledge of its mass-to-charge ratio. [53] The fact that ion's cyclotron frequency is independent of its velocity (***Equation 1.3***) and so, independent of its kinetic energy, is one of the key reasons why FT-ICR is capable of ultra-high resolving power, as the performance of most other types of mass spectrometers is affected by the ion kinetic energy spread. [38]

$$v_c = \frac{1}{2\pi}\,\omega_c \qquad (\textbf{\textit{Equation 1.2}})$$

$$\omega_c = q\,\frac{B}{m} \qquad (\textbf{\textit{Equation 1.3}})$$

The magnetic field applied can only confine ions in the *xy*-plane of the cell, which means the ions are not bound in the z direction and can be lost from the observing region (***Figure 1.3***). For this reason, to stabilize the ion's motion along the magnetic field direction (*z* direction) an electrical trapping potential (1 V for trapping positive ions or -1 V for trapping negative ions) is applied to two plates perpendicular to the magnetic field (***Figure 1.4***). [38] So, when an ion approaches the trapping plate, it turns back towards the centre of the observing region. Therefore, an oscillatory motion, between the trapping plates, is superimposed on the circular cyclotron motion of any ion. [38][53]

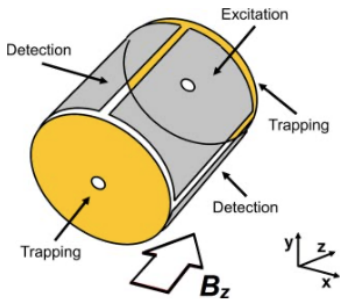

Figure 1.4. *Schematic representation of a closed cylindrical cell. In this example, the detection plates are positioned on the x-axis whereas the excitation plates are along the y-axis. The magnetic field (B_Z) axis is coaxial with the trapping axis (z-axis). The trapping electrodes are located at each end of the cell. [38]*

Ions need to be coherently excited to a larger radius for a signal to be detected. Hence, the magnetic field is kept constant, and the frequency of the external electric field is swept so that it passes through the cyclotron frequencies of the ions of interest. When the RF excitation frequency equals the cyclotron frequency of the ion, the ion will absorb energy from the electric field and spiral up into a larger cyclotron orbit where they'll remain until their energy is dissipated by collisions with residual gas (***Figure 1.5c***). Whereas, ions that are off-resonance do not absorb energy (mostly) and remain at the centre of the cell (***Figure 1.5a***). [53]

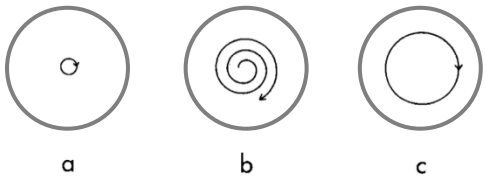

Figure 1.5. *Behaviour of an ion motion under: (a) the magnetic field only, (b) an oscillating electric field pulse, and (c) larger cyclotron orbit. [53]*

All ions of a given mass, which were originally moving with random phases due to their initial random velocities upon formation, are excited coherently and undergo motions as ion packets, so that a much more intense image current can be recorded by the detection plates. Each ion packet has a specific resonance frequency and every time the ions pass through the detection plates an analog 'image current' is generated. Ions' motion in the cell is periodic, and the signal recorded is a composite sum of N sinusoidal waves with different frequencies in the time domain. [38]

The superposition of image currents from all excited 'clumps' of different mass-to-charge ratio ions (thus possessing information about the cyclotron frequency of all ions excited during the frequency sweep portion of the experiment) is converted to a voltage, amplified by a preamplifier, digitized by an analog-to-digital converter, and stored in a computer. [53]

The mathematical process of Fourier transformation of this time-domain signal results in a frequency-domain spectrum, showing peaks of all ions in the sample at their cyclotron frequencies. This frequency spectrum can be converted easily into a mass spectrum if the magnetic field strength is known. [53]

Chapter II : Objectives

MALDI MS it's been the most used technique in the analysis of fingerprints. Never before, a high resolution method such as FT-ICR MS was coupled with ESI in order to unveil the chemical composition of human fingerprints. We opted to use ESI as an ionization method because it allows a bulk analysis of the chemical composition of fingerprints coupled to the extreme resolution and mass accuracy for direct injection and high throughput.

Therefore, this work will focus on revealing the common compositional space of human fingerprints through accurate mass measurements and fine isotopic structure, combining database search and molecular formula determination, which will be explored trough Principal Component Analysis (PCA), Kruskal-Wallis tests and other tools. Endogenous compounds are of particular interest since they held information about the fingerprint's donor.

General characterization of the samples

Through FT-ICR mass spectrometry the spectra of all fingerprints were obtained and ***Figure 3.1*** shows one of the them.

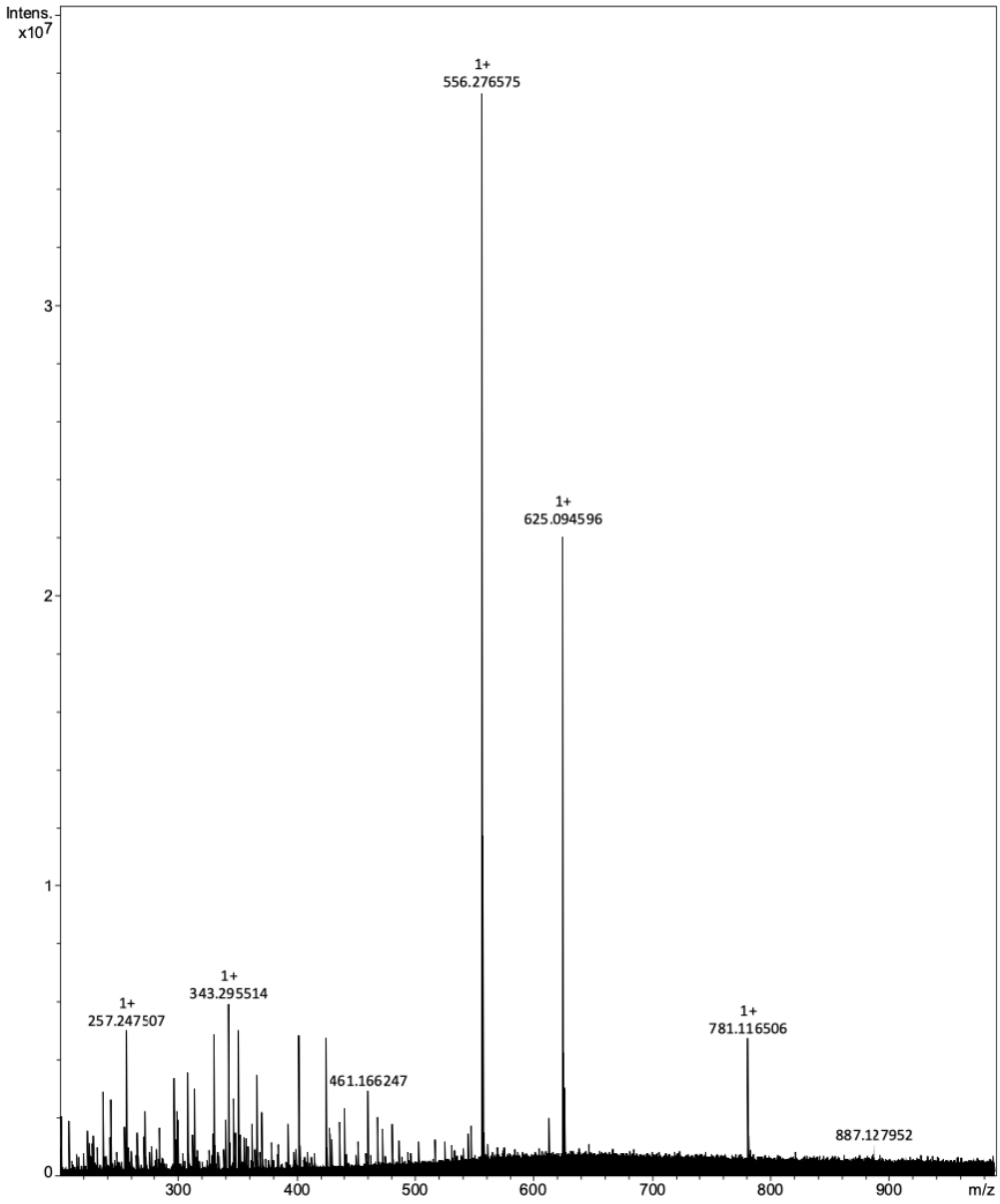

Figure 3.1. *A representative positive ion ESI FT-ICR MS spectrum of a human fingerprint with Leu-Enk (556.276575 m/z) as internal standard for mass spectrum calibration. The spectrum is from volunteer V10 on the 29th October 2019.*

All spectra obtained were very similar. The spectrum is characterized by the presence of small molecules mainly in the range of 200-700 m/z. In terms of relative intensities, the internal standard LeuEnk (peak 556.276412) is the most intense peak which means that the existing compounds are present in concentrations below 2 x 10^{-5} mg/mL.

The main goal of this study is to find similarities between all fingerprints recovered. So, in order to test the variability of the samples a Principal Component Analysis (PCA) was done with all the FT-ICR MS spectra data.

PCA is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set. [55] This reduction is accomplished through the identification of directions, called principal components, along which the variation in the data is maximal. Principal components are used to summarize the data without losing too much information, instead of showing the values for all variables. [56]

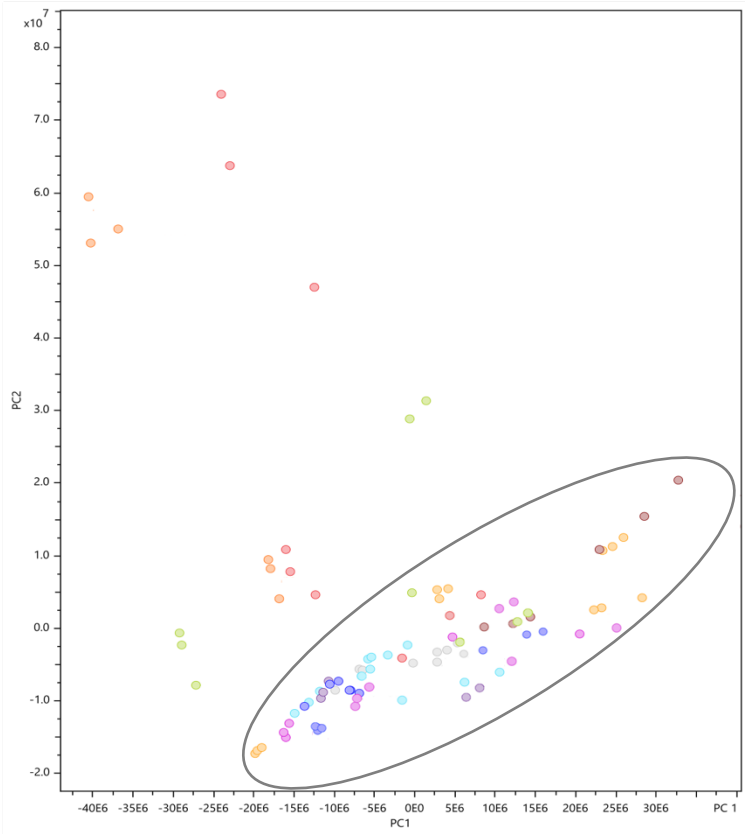

Figure 3.2. Principal Component Analysis (PCA). *PCA of all fingerprint samples, including replicates. The following parameters were used: no scaling algorithm, cross validation mode of 10% and minimum explained variance of 98. Each volunteer is indicated with different colours.*

As it is possible to observe in ***Figure 3.2***, the first principal component (**PC1**) is the direction along which the samples show the largest variation. The second principal component (**PC2**) is the direction uncorrelated to the first component along which the samples show the largest variation. The majority of the samples seem to cluster in a well-defined space which means that, despite the obvious variability between individuals, they may have somethings in common.

In order to estimate the main compound categories present in human fingerprint samples, a van Krevelen diagram was used. The van Krevelen diagram represents atomic H:C vs. O:C ratios of organic compounds, [57] and was created by Dirk Willem van Krevelen, in the 1950s, to evaluate the origin and the chemical evolution of kerogen and petroleum. [57][58] Since then, this representation has been used, beyond petrochemical applications, to study chemical reactions [57] (reaction trend lines reveal the

relationship between the compounds that involve the loss or gain of elements in a specific molar ratio [58]), as well as to examine the chemical characteristics of carboxyl-rich alicyclic molecules and oxidized black carbon in natural organic matter, but most commonly to define a sample as lipid-like, protein-like, etc. [57]

According to the O:C and H:C ratios, compounds occupy a specific region in the plot as represented in ***Figure 3.3 (Left)***.

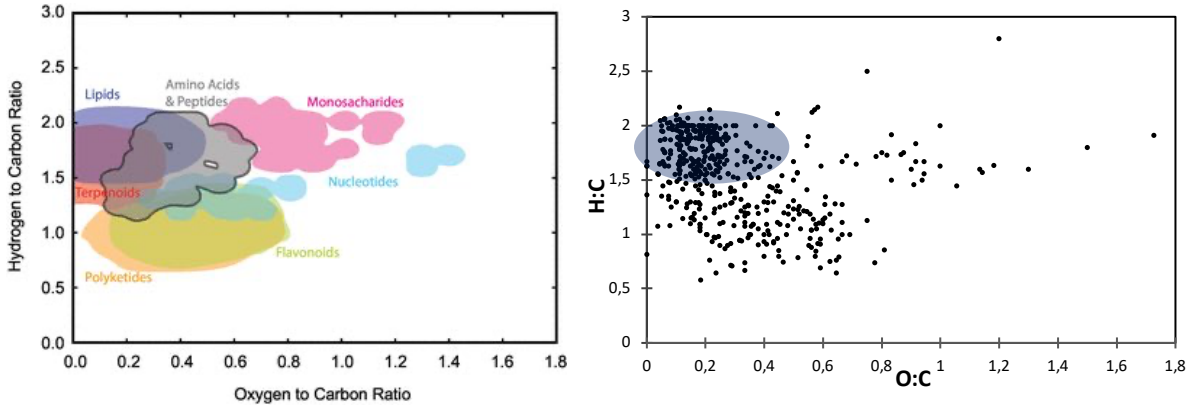

Figure 3.3. Van Krevelen diagrams. Left *– Representation of each compound category in different colours. The areas are an approximation based on previous compound classification of organic matter studies. [59]* ***Right*** *– Representation of all of HMDB assigned compounds in the samples. The blue region represents the main lipid region.*

Comparing both van Krevelen diagrams in ***Figure 3.3***, it's clear that the majority of compounds are concentrated in the lipid region. So it's expected that, in the compounds assigned with HMDB, the main class will turn out to be the lipids.

Nevertheless, there are other classes of compounds that might also be present, such as peptides, aminoacids and carbohydrates.

Common Compounds Identified – Endogenous Compounds

Through the coupling of FT-ICR mass spectrometry, MetaboScape tools as well as *Human Metabolome Database* (HMDB) search, 28.618 was the total number of compounds found in all samples collected. However, only 476 of them were annotated with HMDB and, as starting point, data analysis was done only with these annotated compounds.

All samples were analysed manually using Microsoft Excel and through the application of the defined premises (Materials & Methods – Data Analysis), 38 compounds were found in at least 8 out of 10 individuals and were defined as common. (***Table 3.1***)

Table 3.1. *Compounds identified as common (present in at least 8 individuals). The ratio m/z represents the monoisotopic peak for each compound; the ion is the compound form that was found; and the deviation (ppm) gives us the perception of how accurate is the m/z found.*

Name	Formula	m/z	Ion	Deviation (ppm)
Myristic acid	C14H28O2	229,21625	$[M+H^+]$	0,188

3-(5,6,6-Trimethylbicyclo[2.2.1]hept-1-yl)cyclohexanol	C16H28O	237,22131	$[M+H^+]$	0,076
3-Methylcyclopentadecanone	C16H30O	239,23697	$[M+H^+]$	0,117
Pentadecanoic acid	C15H30O2	243,23189	$[M+H^+]$	0,136
N-Undecanoylglycine	C13H25NO3	244,19075	$[M+H^+]$	0,123
Monoisobutyl phthalic acid	C12H14O4	245,07846	$[M+Na^+]$	0,122
Hypogeic acid	C16H30O2	255,23187	$[M+H^+]$	0,051
Palmitic acid	C16H32O2	257,24752	$[M+H^+]$	0,051
Cyclohexaneundecanoic acid	C17H32O2	269,24751	$[M+H^+]$	0,011
Heptadecanoic acid	C17H34O2	271,26316	$[M+H^+]$	0,011
Linoleic acid	C18H32O2	281,24751	$[M+H^+]$	0,011
Oleic acid	C18H34O2	283,26315	$[M+H^+]$	-0,025
Stearic acid	C18H36O2	285,27881	$[M+H^+]$	0,011
Lauroyl diethanolamide	C16H33NO3	288,25332	$[M+H^+]$	0,000
(9Z,11E,13E,15Z)-4-Oxo-9,11,13,15-octadecatetraenoic acid	C18H26O3	291,19546	$[M+H^+]$	-0,038
Cyclandelate	C17H24O3	299,16177	$[M+Na^+]$	0,017
3-Oxooctadecanoic acid	C18H34O3	299,25808	$[M+H^+]$	0,030
Sphingosine	C18H37NO2	300,28971	$[M+H^+]$	0,013
alpha-CEHC	C16H22O4	301,14103	$[M+Na^+]$	0,000
Hexaethylene glycol	C12H26O7	305,15706	$[M+Na^+]$	-0,046
3-Hydroxy-6,8-dimethoxy-7(11)-eremophilen-12,8-olide	C17H26O5	311,18530	$[M+H^+]$	0,000
N-Undecylbenzenesulfonic acid	C17H28O3S	313,18319	$[M+H^+]$	-0,006
2-Dodecylbenzenesulfonic acid	C18H30O3S	327,19883	$[M+H^+]$	-0,037
(2'E,4'Z,7'Z,8E)-Colnelenic acid	C18H28O3	331,16700	$[M+K^+]$	-0,009
Methylgingerol	C18H28O4	331,18797	$[M+Na^+]$	-0,030
MG(0:0/16:0/0:0)	C19H38O4	331,28428	$[M+H^+]$	-0,018
1,20-Eicosanediol	C20H42O2	337,30769	$[M+Na^+]$	-0,033
Polyoxyethylene (600) monoricinoleate	C21H40O3	341,30501	$[M+H^+]$	-0,035
Lactapiperanol D	C18H28O5	347,18288	$[M+Na^+]$	-0,043
Heptaethylene glycol	C14H30O8	349,18328	$[M+Na^+]$	-0,026
MG(0:0/16:1(9Z)/0:0)	C19H36O4	351,25057	$[M+Na^+]$	-0,029
MG(18:0/0:0/0:0)	C21H42O4	359,31558	$[M+H^+]$	-0,017
Melibiose	C12H22O11	365,10542	$[M+Na^+]$	-0,033
12-Ketodeoxycholic acid	C24H38O4	391,28427	$[M+H^+]$	-0,041
Octaethylene glycol	C16H34O9	393,20949	$[M+Na^+]$	-0,033
Squalene	C30H50	411,39851	$[M+H^+]$	-0,044
Stearoylcarnitine	C25H49NO4	428,37342	$[M+H^+]$	-0,037
Eplerenone	C24H30O6	453,16738	$[M+K^+]$	-0,038

Knowing that the compounds present in the fingerprints can be endogenous or exogenous [20], we can classify the compounds in ***Table 3.1*** as of endogenous origin since it would be extremely unlikely that 10 independent individuals would have the same exogenous compound present in their fingerprints.

Analysing the deviation in ppm (***Table 3.1***), the majority of the compounds have deviations below 1ppm, in absolute value. Although this is not sufficient to guarantee that indeed these selected peaks correspond to the compounds assigned by HMDB, it is a good prediction. Moreover, within this mass deviation and with information on isotopic distribution, the chemical formulas can be unequivocally assigned, providing an additional measure of confidence in the identified compounds.

In mass spectrometry, the most often used mass is the monoisotopic mass. This mass corresponds to the sum of the accurate masses of the principle isotope of each atom in the molecule, generally the lightest, more stable and the most abundant one (^{1}H, ^{12}C and ^{14}N). [60] Following this peak, comes an isotope peak ("+1" peak) with the next more abundant elements, like ^{13}C and ^{15}N. The heavy isotopes ^{18}O and ^{34}S lead to "+2" peaks. Other isotopes are rare to find. [60]

Since we are analysing small molecules, the monoisotopic peak is always the most intense one. [60] And given that we have access to most isotopic distributions, the charge of the correspondent molecules can be determined. Thus, in our samples the majority of species found are monocharged.

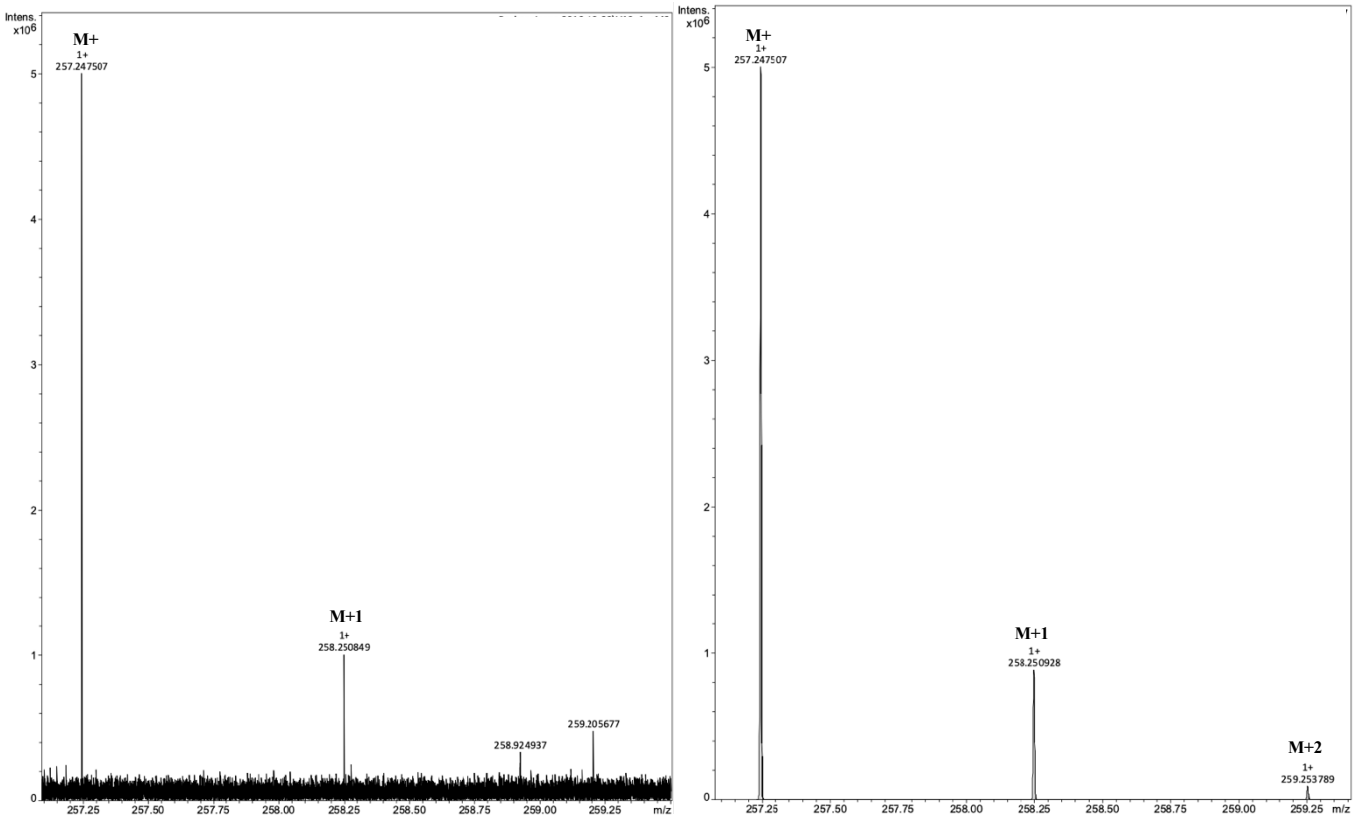

Figure 3.4. Isotopic distribution of Palmitic acid. Left *– Close-up at the m/z range 257-260 on FT-ICR MS spectrum from volunteer 10 on 29th October 2019 of C16H32O2 ([M+H+]).* ***Right*** *– Correspondent simulated isotope pattern using Smart Formula. The monoisotopic peak (**M+**) and the **M+1** peak are represented in both spectra. (Resolution 587 966, FWHM 0.00044)*

An example of an isotopic distribution present in ***Figure 3.1*** is the palmitic acid's distribution (assigned palmitic acid by HMDB), represented in ***Figure 3.4***. In this case, both monoisotopic peak ($\mathbf{M^+}$) as well as the **M+1 peak** are present. The reason why we probably don't see more peaks in the isotopic distribution is because the sample is very little concentrated and/or because the peaks may be below the

noise level. The fact that the theoretical spectrum is similar to the sample spectrum, gives us a very high degree of confidence in the chemical formula attributed to this peak.

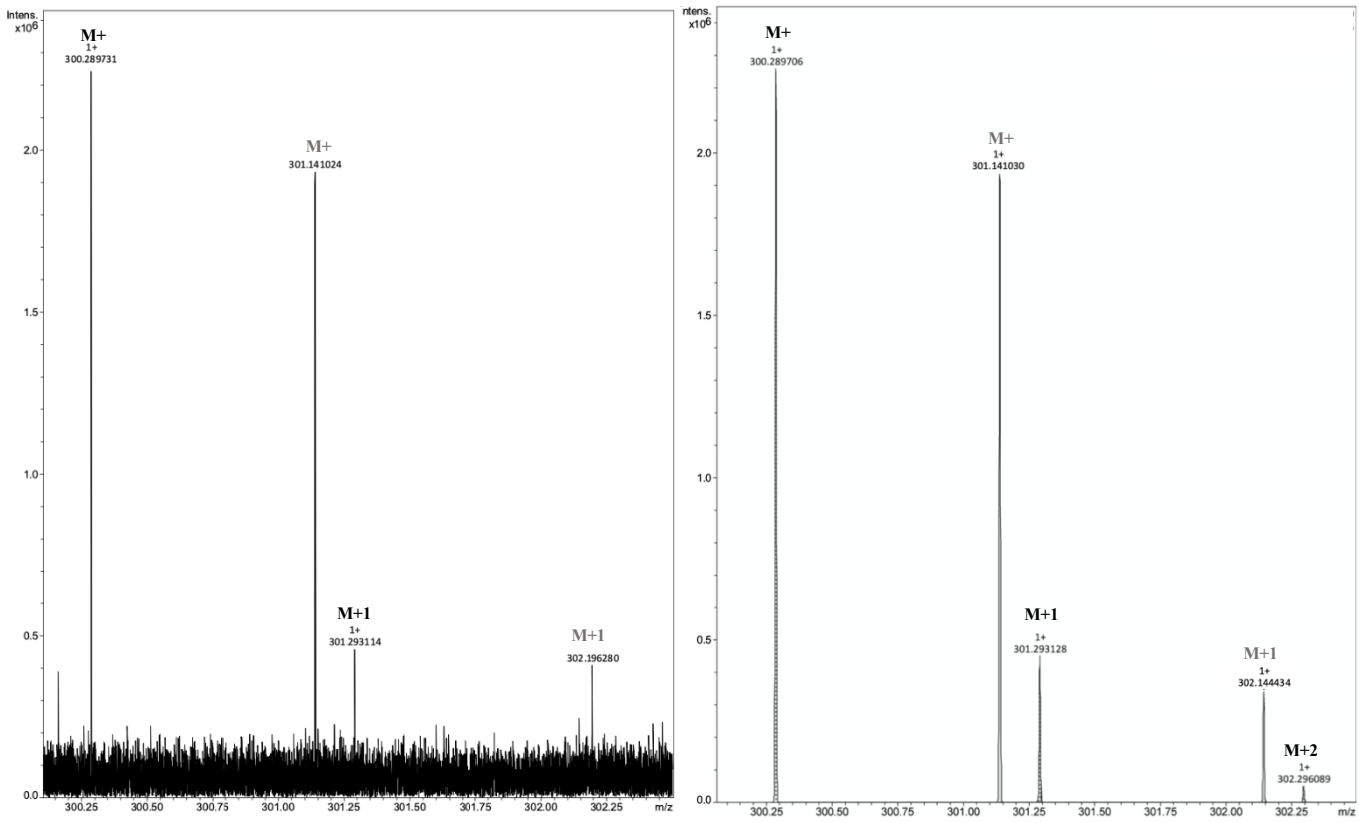

Figure 3.5. Isotopic distribution of sphingosine and alpha-CEHC. Left *– Close-up at the m/z range 300-303 on FT-ICR MS spectrum from volunteer 10 on 29th October 2019 of sphingosine (C18H37NO2, [M+H+]) and alpha-CEHC (C16H22O4, [M+Na+]).* ***Right*** *– Correspondent simulated isotope pattern using Smart formula. The monoisotopic peak (**M+**) and the **M+1** peak are represented in both ions. (Resolution 612 509 / 532 235, FWHM 0.00049 / 0.00057, respectively)*

Another example of isotopic distributions found are exemplified in ***Figure 3.5***. In this case, it's also possible to see, for both sphingosine and alpha-CEHC, 2 peaks of the isotopic distributions, $\mathbf{M}^+$ and **M+1**.

Some of the compounds assigned were recognised based solely in one peak. However, in most cases one cannot obtain much information about the chemical formula from just one peak. That is why the information attached to the isotopic distribution is important, because the total mass of an ion can be decomposed into the masses of its constituents, which is crucial for the chemical formula determination.

In the direction of understanding the kind of compounds found, a detailed search in the *Human Metabolome Database* (HMDB) was done for each compound listed on ***Table 3.1***. Though, there are two compounds assigned by HMDB whose identification is extremely unlikely to be correct – eplerenone and cyclandelate.

Eplerenone ($C_{24}H_{30}O_6$) is a steroidal antimineralocorticoid that is used as an adjunct in the management of chronic heart failure and high blood pressure, particularly for patients with resistant hypertension due to elevated aldosterone. It is an aldosterone receptor antagonist similar to spironolactone, has been shown to produce sustained increases in plasma renin and serum aldosterone, consistent with inhibition of the negative regulatory feedback of aldosterone on renin secretion. [61]

Cyclandelate ($C_{17}H_{24}O_3$) is only found in individuals that have used or taken this drug. It is a direct-acting smooth muscle relaxant used to dilate blood vessels. It produces peripheral vasodilation by a direct effect on vascular smooth muscle. Pharmacological action may be due to calcium-channel antagonism. [62]

Considering their pharmaceutical origin, it is difficult to believe that their identifications are correct. It would be extremely unlikely that all 10 volunteers had taken such medication. But both chemical formulas have isomers and so, it is possible that these compounds may correspond to another isomer instead. On the other hand, the possibility of other chemical formulas correspond to such peaks was also studied using Smart Formula. And, in fact, there were other chemical formulas that may justify a peak at 453.16738 m/z within the experimental mass deviation (assigned eplerenone by HMDB).

Nonetheless, these peaks have been found in the majority of the volunteers, so the correspondent compounds still need to be identified. To do so, further investigations using MS/MS are needed.

I. Class of Compounds

According to the results of the Van Krevelen diagram (***Figure 3.3***), the class that was predominant in our sample was the lipid class. However, our goal now is to focus on 38 of the 476 compounds assigned by HMDB. To find out which is the chemical category these 38 compounds belonged to, a schematic representation of all their classes was made (***Figure 3.6***).

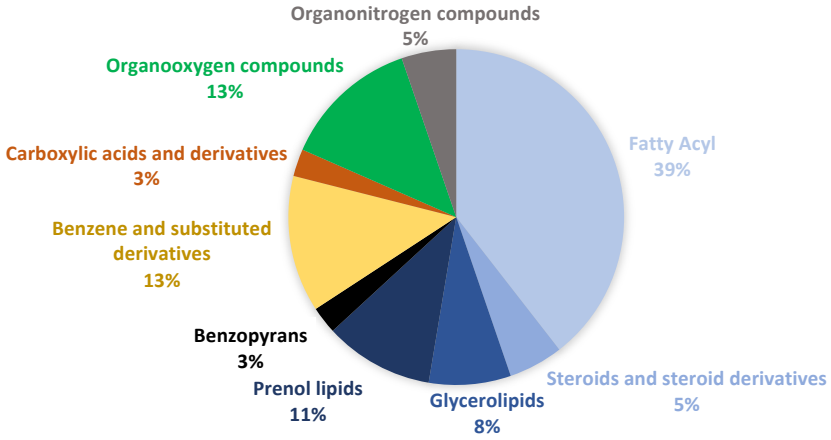

Figure 3.6. Compounds' Classes. *Representation of all compound classes of the 38 common ones. For each category, emphasized by different colours, the percentage of compounds present in that class is marked.*

The pie graph in ***Figure 3.6*** shows that about 63% of the compounds presented in ***Table 3.1*** are lipids or lipid-like molecules (**Fatty acyls**, **Steroids and steroid derivatives**, **Glycerolipids** and **Prenol lipids**). This information is in agreement with the predictions done with the help of the Van Krevelen diagram (***Figure 3.3***) for all compounds in the sample.

Lipids are a heterogeneous group of compounds, mainly composed of hydrocarbon chains characterised by their solubility in nonpolar solvents and insolubility in water. [63] According to ***Figure 3.6***, we have:

- **Fatty Acyls** – are a diverse group of molecules synthesized by chain elongation of an acetyl-CoA primer with malonyl-CoA groups that may contain a cyclic functionality and/or are substituted with heteroatoms. Can either be unsaturated (one or more double bonds) or saturated. Frequently, they naturally occur with an even number of carbon atoms and are unbranched. [63]
  E.g. palmitic acid, oleic acid, heptadecanoic acid, 3-oxooctadecanoic acid, etc.

- **Steroids** are a kind of hormone that is typically recognized by their tetracyclic skeleton, composed of three fused six-membered and one five-membered ring. [63]
  E.g. 12-ketodeoxycholic acid, eplerenone

- **Glycerolipids** have at least one hydrophobic chain linked to a glycerol backbone in an ester or ether linkage. [64]
  E.g. MG(0:0/16:0/0:0), MG(18:0/0:0/0:0) and MG(0:0/16:1(9Z)/0:0).

- **Prenol lipids** are synthesized by the successive addition of the five-carbon-unit precursors isopentenyl diphosphate and dimethylallyl diphosphate, and are classified according to number of these terpene units. [63]
  E.g. lactapiperanol D, 3-(5,6,6-trimethylbicyclo[2.2.1]hept-1-yl)cyclohexanol, squalene, etc.

Apart from lipids we have other type of compounds present, however due to the fact that they are rare/uncommon, they are less studied compared to lipids and, therefore, their characterization becomes more incomplete.

II. Biological Location

Considering that we already know the type of compounds that we have in this set, it is important to understand where they may be in the organism, their biological location.

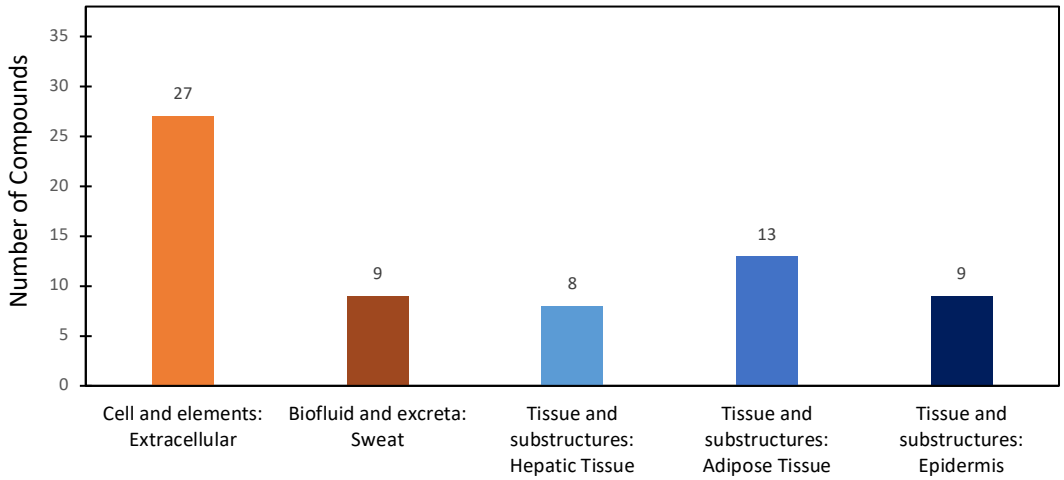

Figure 3.7. Biological Location. *Illustration of the biological locations of the compounds in this study. There are more possible locations for the compounds, however, these were the ones that were relevant. For each location represented, the number of compounds that may be in that location is marked.*

The compounds in ***Table 3.1*** have many possible biological locations, nonetheless, only a few were shown in ***Figure 3.7***.

Like one may observe in ***Figure 3.7***, 27 out of 38 compounds exist extracellularly. Since we are analysing molecules present in human fingerprints this makes sense.

As it was clear in ***Figure 3.6***, the majority of the compounds are lipids or lipid-like molecules and so it was expected that they appear in both the hepatic (8 compounds) and adipose tissue (13 compounds) since these are the places where they are produced and stored, respectively (***Figure 3.7***).

Fingerprints are composed by the molecules present in the skin (epidermis). These compounds can come from the epidermis itself or from sweat. [7] However, from the 38 endogenous compounds found, only a few were already detected and/or quantified in the skin or in sweat, in previous reports (***Figure 3.7***). Some examples of these compounds are palmitic acid, stearic acid, oleic acid and linoleic acid.

Although these compounds may have several biological locations, it is important to note that the actual way they travel from their biological location to the epidermis remains unknown. [65]

III. Source

Although most compounds present in fingerprints are classified as endogenous, this does not necessarily mean that they are produced endogenously. So, they may have other sources.

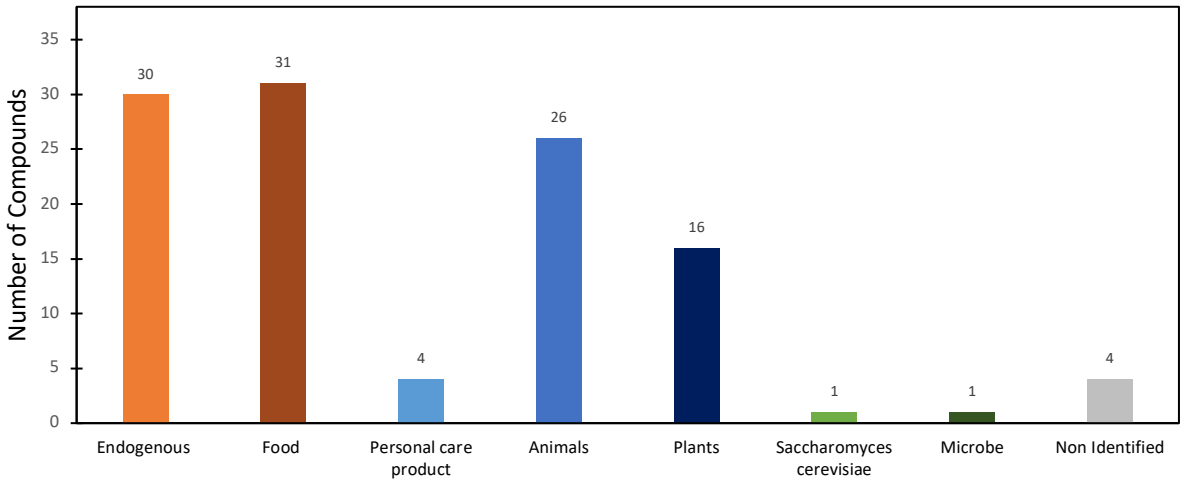

Figure 3.8. Source. *Representation of the possible compounds' sources. For each source, the number of compounds that may have that source is marked.*

Though there are 38 compounds, only 34 of them have known sources. For 4 compounds, the source has not been identified yet (***Figure 3.8***).

There are four main compound sources: food (31 compounds), endogenous sources (30 compounds) and animal and plants products (26 and 16 compounds) (***Figure 3.8***). However, the majority of compounds in the animal and plant categories enter our system in the form of food.

The majority of these compounds can be produced by the human body and so they have an endogenous source. For example, monoacylglycerols (MG(0:0/16:0/0:0), MG(18:0/0:0/0:0) and MG(0:0/16:1(9Z)/0:0)) are formed biochemically via release of a fatty acid from diacylglycerol by diacylglycerol lipase or hormone sensitive lipase. [66]–[68]

However, humans can also obtain these compounds through food, mainly from animal or plant origin. Melibiose is found in fermented molasses, brown sugar, or honey. [69] Methylgingerol can be found in herbs and spices. [70] Many compounds are found in plant oils, for example palmitic acid (palm oil), oleic acid(mainly in olive oil), myristic acid and squalene. [71]–[74] Another major source for these compounds are dairy fats (milk and meat). Examples of compounds that have this has a source too are: monoacylglycerols and myristic acid. [66]–[68][73]

Heptadecanoic acid and pentadecanoic acid ("odd"-chain fatty acids) are examples of compounds that, although they can be produced endogenously, for example, from gut-derived propionic acid [75], they are mainly obtained through exogenous sources like dairy fats (e.g. milk fat and ruminant meat). [66][67] On the other hand, linoleic acid can only be obtained through diet (plant glycosides). [78]

Nevertheless, there are 2 compounds that despite having other origins, can also originate from *Saccharomyces cerevisiae* and microbe (***Figure 3.8***). Melibiose is a sugar produced and metabolized only by enteric and lactic acid bacteria and other microbes and squalene may also come from yeast lipids. [69][74]

IV. Biological Role

It is also interesting to see what kind of functions these compounds play. However, we must bear in mind that the functions listed are functions that have already been studied in previous articles. It does not necessarily mean that these compounds perform these same functions on the skin (fingertips). On the other hand, we also have to consider that the classification made by HMDB is not restrictive. There is a chance that these compounds have other functions on the skin that have not yet been reported.

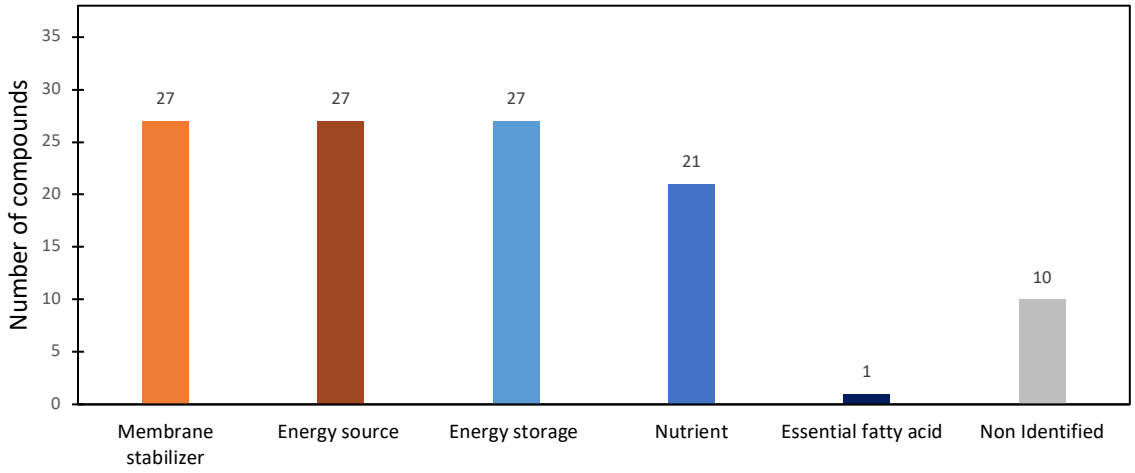

Figure 3.9. Biological Role. *Representation of the biological roles the compounds may perform. For each category, the number of compounds that do have that biological role is marked.*

As in the previous section, only 28 compounds have identified functions. Compounds that have non-identified functions do not belong to the lipid class and so the majority of the compounds analysed, in this case, are lipids.

Lipids yield high energy and play crucial roles in the normal function of the cell. Not only do lipids serve as highly reduced storage forms of energy, but they also play an intimate role in the structure of all biological membrane, providing hydrophobic barriers that permit partitioning of the aqueous contents of the cell and subcellular structures. Additionally, they provide insulation to the body, can act as hormones, may participate in signalling pathways, amongst other functions. [63][69][70]

According to ***Figure 3.9***, the majority of the compounds have the 3 following biological roles: membrane stabilizer, energy source and energy storage. All of which are some of the main functions of lipids as mentioned.

Linoleic acid is an essential fatty acid (***Figure 3.9***) because it cannot be synthesized in mammalian tissues and must be obtained from the diet. It is used in the biosynthesis of prostaglandins (via arachidonic acid) and cell membranes. [78]

V. Biological Processes

Even though we already have listed some of their possible functions, it is also relevant to understand what pathways these compounds intervene in.

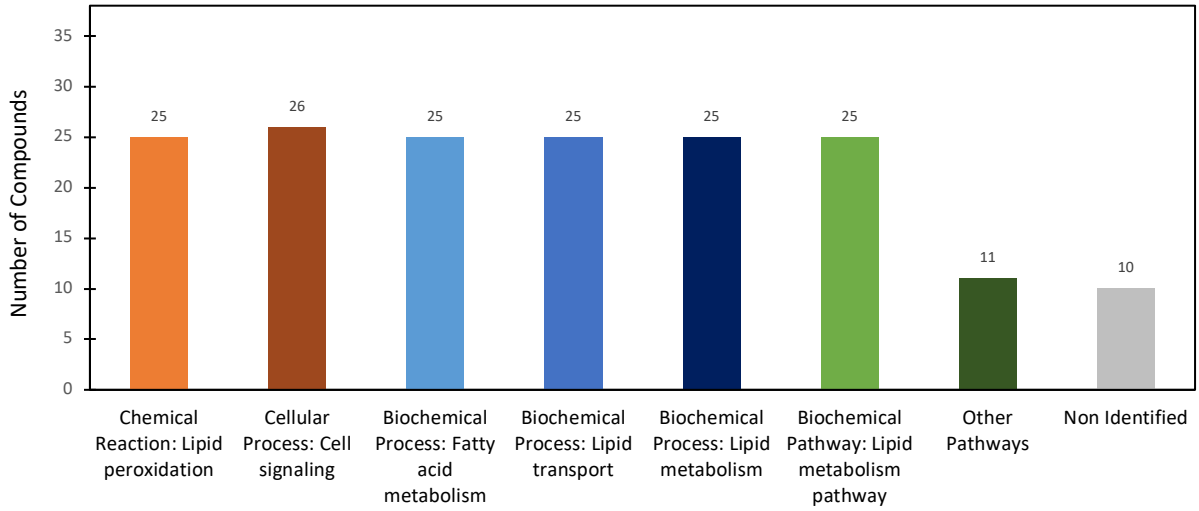

Figure 3.10. Biological Process. *Representation of the biological processes the compounds may be involved in. For each category, the number of compounds that participated in that biological process is marked.*

As in the previous section, only 28 compounds have identified biological processes, most of which are lipids.

Analysing ***Figure 3.10***, one can observe that the majority of the identified processes involve lipids: lipid peroxidation (25 compounds), fatty acid metabolism (25 compounds), lipid transport (25 compounds) and lipid metabolism (25 compounds).

Though these compounds participate in other pathways (11 compounds), like, for example, alpha linolenic acid and linoleic acid metabolism and mitochondrial beta-oxidation of long chain saturated fatty acids, the main pathway they are involved in is lipid metabolism (25 compounds) (***Figure 3.10***).

VI. Industrial Applications of Endogenous Compounds

Even though what is relevant for this study is knowing the characteristics of the compounds, what they go through before they reach the fingerprints as well as how they get there, it can also help to understand what are the possible industrial applications of all these compounds. After all, it can also justify some of the compound sources described above (**Source**).

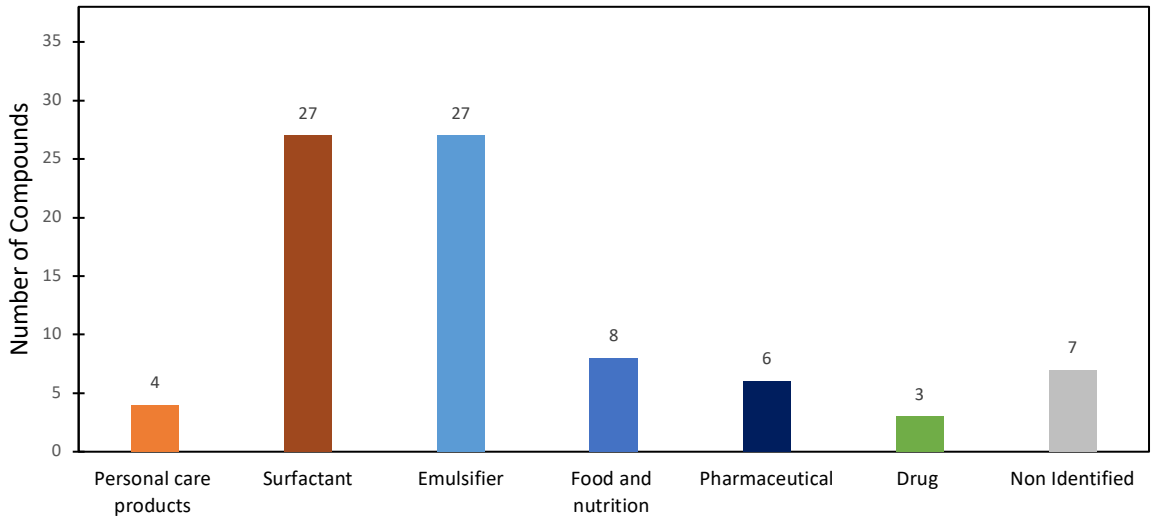

Figure 3.11. Industrial Applications. *Representation of the industrial applications the compounds may possess. For each category, the number of compounds that do have that biological role is marked.*

In accordance to the analysis of ***Figure 3.8***, the identified compounds can be used in personal care products (4 compounds) and in food and nutrition (8 compounds) (***Figure 3.11***). On the other hand, these compounds may also be employed in surfactants (27 compounds), emulsifiers (27 compounds), pharmaceuticals (6 compounds) and drugs (3 compounds).

As it was mentioned already, two of the common compounds found (eplerone and cyclandelate) were used for medical issues. [61][62] And in this case, they both are used to produce drugs as well as pharmaceuticals. It's important to have in mind that the definition of these classes was done by HMDB and the differences between classes may not be clear.

Interestingly, the other compound used in drug manufacturing is linoleic acid. It was found that the administration of low daily doses of this compound along with calcium during the third trimester of pregnancy reduced the incidence of preeclampsia significantly in women at high risk, possibly by correcting the PGE2 levels. [81]

VII. Statistical Analysis

With the aim of determining which compounds present in ***Table 3.1*** have similar intensities, a statistical analysis was performed. Since our sample sizes are very small and unequal, the one-way ANOVA's assumptions were difficult to insure. [82] Instead a Kruskal-Wallis test was performed. [83]

Kruskal-Wallis test is a non-parametric test in which no assumptions are made about the type of underlying distribution, and no population parameters are estimated. Each group sample has to have at least 5 elements and each group must be independent from each other. [83]

Since the same data were used in multiple tests, one for each compound, a correction for multiple tests was required. [83] A Benjamini-Hochberg's correction was applied to the p-values obtained by the Kruskal-Wallis test.

Table 3.2. Kruskal-Wallis Test. *The compounds represented in this table result from the Kruskal-Wallis test performed on all human fingerprints (including replicates), and for which the null hypothesis (H0) wasn't rejected with an α=0,05. The degrees of freedom were 9. P-values represented were obtained with Kruskal-Wallis test (Appendix B -* ***Table B.1****); alphas were calculated with the Benjamini-Hochberg correction (38 tests, FDR=0,05)(Appendix B -* ***Table B.2****).*

Name	p-value	alpha
3-Methylcyclopentadecanone	0,02636834	0,02500000
Cyclohexaneundecanoic acid	0,0378247	0,02631579
Pentadecanoic acid	0,06464665	0,02763158
Linoleic acid	0,08294001	0,02894737
Heptadecanoic acid	0,08387415	0,03026316
Myristic acid	0,08960404	0,03157895
N-Undecanoylglycine	0,09228112	0,03289474
12-Ketodeoxycholic acid	0,16712249	0,03421053
Oleic acid	0,19296415	0,03552632
Lactapiperanol D	0,21074484	0,03684211
2-Dodecylbenzenesulfonic acid	0,23771711	0,03815789
3-Hydroxy-6,8-dimethoxy-7(11)-eremophilen-12,8-olide	0,26326445	0,03947368
(9Z,11E,13E,15Z)-4-Oxo-9,11,13,15-octadecatetraenoic acid	0,31551758	0,04078947
Stearic acid	0,3486732	0,04210526
Palmitic acid	0,39199221	0,04342105
Polyoxyethylene (600) monoricin	0,46799697	0,04473684
Methylgingerol	0,49289019	0,04605263
MG(18:0/0:0/0:0)	0,5050603	0,04736842
(2'E,4'Z,7'Z,8E)-Colnelenic acid	0,59473112	0,04868421
MG(0:0/16:0/0:0)	0,77850059	0,05000000

From the twenty compounds present in ***Table 3.2***, only 4 of them don't belong to the class of lipid and lipid-like molecules – 3-methylcyclopentadecanone, N-undecanoylglycine, 2-dodecylbenzenesulfonic acid and methylgingerol.

Despite all the variability inherent to human beings, we not only found 38 compounds present in all individuals (***Table 3.1***), but also 20 of them appear with similar intensities (***Table 3.2***). Keeping in mind the main goal of this project, this is the first step towards the definition of a chemical background for human fingerprints' composition.

Exogenous Compounds

Considering that fingerprint have both endogenous and exogenous compounds, it would also be interesting to see if applying this method, we would be able to access exogenous compound. For this, the sample used was the same, however, instead of looking for compounds present in all individuals, we looked for compounds that were only present in one, on a given day.

Table 3.3. Exogenous Compounds. *Compounds only found in one volunteer. The ratio m/z represents the monoisotopic peak for each compound; the ion is the compound form that was found; and the deviation (ppm) gives us the perception of how accurate is the m/z found.*

Name	Formula	m/z	Ion	Deviation (ppm)
Pantothenol	C9H19NO4	206,13875	$[M+H^+]$	0,315
Tangeritin	C20H20O7	373,12817	$[M+H^+]$	-0,024
Hexamethylquercetagetin	C21H22O8	403,13873	$[M+H^+]$	-0,035
Isosakuranin	C22H24O10	449,14420	$[M+H^+]$	-0,051
Limonin	C26H30O8	471,20133	$[M+H^+]$	-0,030
(7alpha,10beta)-1(10->19)-Abeo-7-acetoxyisoobacun-3,10-olide	C28H34O9	515,22752	$[M+H^+]$	-0,076
Naringin	C27H32O14	581,18643	$[M+H^+]$	-0,089
Hesperidin	C28H34O15	611,19705	$[M+H^+]$	0,005

As with the compounds above, these were the identifications **(*Table 3.3*)** made with the HMDB and considering that the deviation values are all less than 1 ppm, all the interpretation was made considering the assignment is correct. Further analysis will require a MS/MS approach.

Pantothenol is used in cosmetics, mainly in shampoos and hair conditioners. [84] With that in mind we try to know which shampoo was used by this volunteer, at that day, and in fact this compound is one of those present in its composition. This information supports the fact that there might be compounds in fingerprints that may come from personal care products (***Figure 3.8***).

All the other compounds present in ***Table 3.3*** are all common in citrus, specifically sweet oranges. [85]–[91] And actually, this individual ate an orange at one of the sampling days.

Although this information might not be of extreme importance in a criminal investigation, it can be used to narrow the pool of suspects as the suspect is probably not allergic to sweet oranges and possibly washed his hair that day.

These results demonstrate that this type of analysis can go far beyond endogenous compounds and can be very useful in forensic analysis. And a fundamental condition is that a common constant chemical background is identified.

Chapter IV : Discussion

Fingerprints' chemical composition has been the focus of several studies, however, not only have these studies focused on finding specific molecules (for example, finding drugs, age and gender biomarkers, etc.), but other methods have also been used. [1][6][30][31][44][45] This study coupled electrospray ionization with FT-ICR MS to carry out a bulk analysis of human fingerprints, with the aim of finding common compounds in all individuals.

By analysing the results with MetaboScape and HMDB, a total of 28 618 compounds were found but only 476 were assigned possible identifications. To facilitate interpretation, only the annotated compounds were analysed. Still, it is important to have in mind that further studies will be required for the identification of the other 28 142 compounds.

In order to examine the variability of the samples, including replicates, a PCA (***Figure 3.2***) was performed. Considering that most samples were grouped in a well-defined space, this is already an indication that there may be something common to all samples. And, in fact, through manual analysis of all 476 compounds, there were 38 endogenous compounds found in almost every fingerprint sample, every day (***Table 3.1***). These results are quite surprising given that we are studying human beings that, by themselves, already have a great inherent diversity as well as a considerable day to day variation in fingerprint chemical composition from the same individual but even so it was possible to find common compounds.

It is important to have in mind that the compounds were assigned by the HMDB and this identification may not be definitive. There is not only the possibility of isomers, but the peaks can correspond to other chemical formulas. The only way to ensure that the identification is correct is by applying MS/MS. Given the spectral complexity, conventional MS/MS methods based on single ion-isolation at the time are not feasible in a reasonable time scale. Moreover, it's not always possible to isolate single ions in the gas phase, most notably isobaric ones so the resulting MS/MS spectrum may not be very useful. Instead, the method to be used in the future is two-dimensional (2D) FT-ICR MS, an all ion-fragmentation method where MS/MS product ions are specifically assigned based on phase. 2D MS enables easy access to the fragmentation patterns of all compounds present in a complex sample because precursor ion's isolation is not required. This approach has proven to be a versatile method for small molecules, and in the field of lipidomics, polymer analysis, bottom-up and top-down proteomics. [92]–[97] And so, since we are studying mainly small molecules with MS/MS it's not always possible to obtain enough fragments to identify the molecule and that is why the MS/MS approach would serve better for targeted and quantitative analysis. [97]

Nonetheless, in this study we were able to obtain deviation values below 1 ppm which is very good since no calibration was done (neither internal nor external). And so, the interpretation of the results was done assuming the assignment of compounds was correct.

The majority of common endogenous compounds found in fingerprints are lipids or lipid-like molecules. (***Figure 3.3, Figure 3.6***). As previously mentioned, although 38 compounds have been identified with HMDB (including 2 that we know are incorrectly identified), not all of these compounds have been deeply studied yet, probably due to the fact that they are rare (mainly non-lipids ones). However, the majority of them was studied but not in terms of understanding their presence in the skin or in sweat.

With that being said, this discussion will focus on compounds that have already been identified in this scenario.

Considering that we are analysing fingerprints' samples, the compounds found come from all the material that might be present on the skin of the fingertips. The skin is the largest organ of the body and it is primarily regarded as a barrier tissue for it provides the interface between the external environment and the host, protecting individuals from ultraviolet, microbial, chemical, and physical insults. [65][98]

The epidermis, the outer layer of the skin, is covered by an hydrophobic lipid matrix constituted by long-chain fatty acids, squalene, wax monoesters, triglycerides, ceramides and small amounts of cholesterol and cholesterol esters. [65][99] These lipids may already be part of the epidermis, but many of them come from sweat (***Figure 3.7***). The compounds found that had already been identified in these biological locations are: palmitic acid, oleic acid, stearic acid, linoleic acid, myristic acid, sphingosine, squalene, pentadecanoic acid, MG(0:0/16:1(9Z)/0:0), MG(0:0/16:0/0:0) and MG(18:0/0:0/0:0).

Since we are discussing lipids which are primarily produced in the liver and stored in the adipose tissue, these two biological locations may be possible origins for the compounds found in fingerprints. (***Figure 3.7***) Nevertheless, how lipids are transported into the epidermis, regardless of their biological location, it is still unknown. Recent studies have shown that lipoprotein receptors (LDL and scavenger receptor class B type 1 receptors) present in keratinocytes and fatty acid transporters (fatty acid transport proteins 1, 3, 4, and 6 and CD36) expressed in the epidermis may play a role in the transport of lipids into the epidermis, required for permeability barrier formation. [65]

Although skin lipids may originate from food intake, they are mainly endogenously produced (***Figure 3.8***), except for compounds such as linoleic acid which is an essential fatty acid and can only be obtained through diet. These lipids are essentially produced by sebocytes (sebaceous glands) and keratinocytes (primary type of cell found in the epidermis). [65] Sebocyte-derived lipids synthesis depends on local flux through the *de novo* lipogenesis (DNL) pathway within the sebocyte. [80], [100] Keratinocyte-derived lipid precursors are synthesized by keratinocytes and are secreted in the form of lamellar bodies, that eventually release lipid precursors and acid lipases into extracellular space of epidermis, followed by the lipid precursors are catalysed by the acid lipases to produce skin lipids. [80][99]

Both keratinocytes and sebocytes express the full complement of DNL enzymes. Acetyl–coenzyme A (CoA) carboxylase (ACC), which catalyses the conversion of acetyl-CoA to malonyl-CoA, plays a key role in regulating lipid metabolism. ACC is the rate-limiting step in the *de novo* synthesis of fatty acids (DNL) and also regulates oxidation of long-chain fatty acids. [100] Knowing this, it is easier to understand why lipid metabolism, lipid transport and fatty acid metabolism are pathways in which these compounds participate in (***Figure 3.10***), even in the skin.

Though sebaceous glands and keratinocytes are capable of DNL, the importance of this pathway in humans for sebum biosynthesis relative to the utilization of circulating lipids derived from external sources, including diet, adipose tissue, or hepatic tissue (liver), is unknown. [98][100]

Apart from sebocytes and keratinocytes, there is another important source of skin lipids that should not be ignored – skin microbiome. Skin is colonized by numerous microorganisms (bacterial species are the most common), most of which produce lipids to the skin during their metabolism, such as short-chain fatty acids (SCFA). [80] Some bacterial microbiota of the skin produces lipases capable of hydrolysing

sebaceous triglycerides and produce fatty acids. However, the relative significance of bacterial triglyceride hydrolysis compared with hydrolysis by epidermal acid lipase remains unknown. [99]

Lipids are essential metabolites that have many crucial functions in the skin. It is known that skin status is dramatically influenced by skin lipids that do much more than maintaining the permeability barrier. They are considered to be important multifunctional mediators to influence the process of epidermal metabolism, inflammation, cell proliferation and differentiation. [80][98][101]

Squalene is only secreted by sebaceous glands and has a strong moisturizing and antioxidation effect owing to its special chemical structure. [80] This molecule is capable of neutralizing the formation of UV irradiation-induced reactive oxygen species in the skin having a potential role in the process of protection against sunburn and UV radiation. [80] This explains why lipid peroxidation (***Figure 3.10***) is one of the chemical reactions in which several compounds from our sample participate in. On the other hand, peroxidated squalene along with unsaturated free fatty acids, have been reported to be comedogenic and therefore responsible for pathological skin conditions. [102]

In human epidermis, Peroxisome Proliferator-Activated Receptors (PPARs) are a crucial set of transcription factors regulated by skin lipids that control lipid metabolism, skin barrier permeability, inflammation, cell proliferation and differentiation. There are three isoforms (α, β/δ and γ), but PPARγ is the main functional isoform in the skin. [80] They are activated by a wide variety of ligands that are derived from the metabolism of fatty acids such as palmitic acid, oleic acid, linoleic acid and arachidonic acid. [80] It is known, nowadays, that unsaturated fatty acids (e.g. linoleic acid) are the major ligands to PPARγ. When activated, PPARs heterodimerize with the retinoid X receptor (RXR) in the presence of activators, and bind to a PPAR response element (PPRE) in the promoters of target genes, including a network of genes involved in cell proliferation, differentiation and inflammatory responses. [80]

When unsaturated fatty acids bind to PPARγ, they regulate lipid metabolism and skin barrier permeability by increasing the production of cholesterol and ceramides in keratinocytes, promoting sebum accumulation, inhibiting keratinocyte cell proliferation and promoting epidermal terminal differentiation. [80][102] Additionally, these receptors play an important role in anti-inflammatory situations, for example, by decreasing proinflammatory genes expression in a ligand-dependent manner by antagonizing the activities of transcription factors such as members of the NF-κB and AP-1 families, but also inhibit Langerhans cell functions and decrease expression of adhesion molecules. [80]

The interrelationship between skin lipids and residential microbial skin flora is essential. On one hand, skin microbiome is selected through a combination of skin lipids. On the other hand, alterations in residential microbial skin flora influence skin status because these microorganisms will metabolize skin lipids, producing “active” fatty acids that can stimulate some signal network, such as PPARγ. This causes relevant alterations of the skin lipid pattern and affects skin condition such as skin moisturization, barrier restoration potential, and inflammation. [80][99] For example, a variety of naturally occurring lipids, such as fatty alcohols, free fatty acids (both saturated and unsaturated), and monoglycerides, encourage the growth of benign bacteria such as *Staphylococcus epidermidis, S. pyogenes*, and *C. albicans* while exhibiting a potent antimicrobial activity against enveloped viruses, gram-positive (e.g. *Staphylococcus aureus*), gram-negative bacteria, and fungi. [65][98]

In response to different inflammatory stimuli, e.g. against *P. acnes*, sebum lipids exert antimicrobial effects, induce the cytokine expression of sebocytes and keratinocytes and modulate both keratinocyte and macrophages functions by altering their gene and protein expression levels. [99][103] Macrophages,

the key phagocytes of the immune system, sense and transform signals and stimuli towards the initiation of inflammation and activation of the adaptive immune system. M. Lovászi et al. showed that oleic acid and linoleic acid promoted monocyte differentiation into alternatively activated macrophages, without the presence of IL-4/IL-13. [103] These express all the characteristic markers such as CD206, CD209 and FXIII-A, and are central to the maintenance of the tissue environment by producing extracellular matrix components and contributing to tissue remodelling. [99][103] Moreover, in the presence of *P. acnes*, palmitic acid, stearic acid and oleic acid augmented the secretion of interleukin (IL)-1β, whereas oleic acid had a selective effect of inducing IL-1β but downregulating IL-6 and tumour necrosis factor-α (TNF-α) secretion. [103] Additionally, linoleic acid also had an anti-inflammatory effect in *P. acnes*-activated macrophages, inhibiting the secretion of IL-1β, IL-6 and TNF-α. [103]

With all this information in mind, we can see that skin lipids, whether they originate from the microbiome or from sweat, are not just membrane stabilizers in the skin. They can provide a direct readout of cellular metabolic status and this information has a great value in applications that go way beyond forensic applications, for example in medical diagnostics. Gibson, L. E., and R. E. Cooke demonstrated sweat chloride levels are elevated in cystic fibrosis patients. [104] Considering that most fingerprint's molecules come from sweat, this change might be detected in fingerprints. However, this detection will not happen directly since the chlorine's mass is below the m/z range. But in an indirect way it may be possible since there will be greater signal suppression as well as greater abundance of sodiated species.

That justifies why the chemical pattern of all compounds that appear on fingerprints and remain approximately constant is of utmost importance. All the compounds present in ***Table 3.2*** are potential candidates for defining the chemical background for human fingerprints. In this sense, further studies, preferably with individuals of all age groups, are required not only to establish the chemical pattern of human fingerprints, but also to understand what the functions of these compounds are at the fingertips, and what kind of information we can extract from their presence/absence and possible changes in their quantity.

As mentioned in the introduction section, fingerprints have both endogenous and exogenous compounds and although this study focused mainly in endogenous compounds, it was also important to determine if this straight-forward approach could detect exogenous compounds too. In fact, it was possible to detect them as it is evident in ***Table 3.3***. Therefore, the use of FT-ICR MS coupled with ESI proved to be a promising method in the analysis of fingerprints, not only from a forensic perspective but also a clinical one.

Chapter V : Conclusion

Fingerprints result from the transfer of material present in the raised portion of the friction ridge skin to a surface upon contact. [1] The molecules left behind can be either endogenous or exogenous molecules. [20] Since our study focused on the compounds that were present in all individuals, we were looking for endogenous compounds.

In this sense, of 38 compounds found in all samples, 20 had similar intensities and were mainly lipids and lipid-like molecules (except for 3-methylcyclopentadecanone, N-undecanoylglycine, 2-dodecylbenzenesulfonic acid and methyl-gingerol). According to the literature, skin lipids present in fingerprints have three main sources: epidermis (keratinocytes), sweat (sebocytes), and the skin microbiome. [65][80][99][105] However, considering that the analysis of the compounds was done with HMDB and considering that the exact contribution of the skin microbiome is not yet known, in our samples we were only able to identify two of these sources: epidermis and sweat.

These lipids (palmitic acid, oleic acid, stearic acid, linoleic acid, pentadecanoic acid, myristic acid, MG(0:0/16:0/0:0) and MG(18:0/0:0/0:0)) not only have a role in maintaining the permeability barrier, but can also play important roles in the process of epidermal metabolism, inflammation, cell proliferation and differentiation.

Lipid composition can provide a direct readout of cellular metabolic status and the change of lipid metabolism is one important factor in various disease pathogenesis, such as metabolic diseases, several cancers, neurodegenerative diseases, and skin diseases. [80][92] And so, the analysis of fingerprints may be a potentially valuable and non-invasive diagnostic method when, for example, a disease, alters the lipid composition of fingertips. However, other than lipids, sweat also contains a wide variety of other small molecules that can be influenced by disease states and environmental exposures, which can also appear in fingerprints. For example, the eccrine sweat glands in the skin contribute to the elimination of water-soluble toxins such as heavy metals and drugs, that are believed to emerge in sweat by either transdermal migration or passive diffusion from blood, and can then, possibly, be detected in fingerprints. [101][102]

On the other hand, it is also important to understand the contribution of the skin microbiome in the composition of skin lipids, in future experiences. This can be a hard challenge since the skin microbiome varies across individuals. [80]

In conclusion, coupling ESI with FT-ICR MS to study the chemical composition of fingerprints proved to be a promising analytical method for future analyses, due to its unequalled sensitivity, high-throughput and low detection limits. With this kind of analytical method, we can obtain much more information than we could ever get just by looking at the ridges pattern.

Although this was a preliminary study and with a limited number of subjects, it allowed us to recognize, for the first time, that it might be possible to determine the chemical background for human fingerprints, which can have countless applications, being the most relevant one the chance to implement new non-invasive diagnostic methods.

In future reports it is necessary to perform MS/MS, most likely through 2D MS to confirm the identity of the compounds found, include samples from individuals of all age groups and, essentially, comprehend the source of these compounds so that we can understand why they appear on fingerprints.

References

[1] L. S. Ferguson *et al.*, "Direct detection of peptides and small proteins in fingermarks and determination of sex by MALDI mass spectrometry profiling," *Analyst*, vol. 137, no. 20, pp. 4686–4692, 2012.

[2] M. I. Szynkowska, K. Czerski, J. Rogowski, T. Paryjczak, and A. Parczewski, "ToF-SIMS application in the visualization and analysis of fingerprints after contact with amphetamine drugs," *Forensic Sci. Int.*, vol. 184, no. 1–3, pp. 2008–2010, 2009.

[3] P. Azadeh *et al.*, "Fingerprint changes among cancer patients treated with paclitaxel," *J. Cancer Res. Clin. Oncol.*, vol. 143, no. 4, pp. 693–701, 2017.

[4] History.com Editors, "A bloody fingerprint elicits a mother's evil tale in Argentina," 1892. [Online]. Available: https://www.history.com/this-day-in-history/a-bloody-fingerprint-elicits-a-mothers-evil-tale-in-argentina. [Accessed: 20-Feb-2020].

[5] "Crime Scene Forensics, LLC, Concord, NC; www.crimescene-forensics.com." [Online]. Available: www.crimescene-forensics.com. [Accessed: 20-Feb-2020].

[6] M. I. Szynkowska, K. Czerski, J. Rogowski, T. Paryjczaka, and A. Parczewski, "Detection of exogenous contaminants of fingerprints using ToF-SIMS," *Surf. Interface Anal.*, vol. 42, no. 5, pp. 393–397, 2010.

[7] S. Cadd, M. Islam, P. Manson, and S. Bleay, "Fingerprint composition and aging: A literature review," *Sci. Justice*, vol. 55, no. 4, pp. 219–238, 2015.

[8] R. Bradshaw, S. Bleay, R. Wolstenholme, M. R. Clench, and S. Francese, "Towards the integration of matrix assisted laser desorption ionisation mass spectrometry imaging into the current fingermark examination workflow," *Forensic Sci. Int.*, vol. 232, no. 1–3, pp. 111–124, 2013.

[9] B. Burger, D. Fuchs, E. Sprecher, and P. Itin, "The immigration delay disease: Adermatoglyphia-inherited absence of epidermal ridges," *J. Am. Acad. Dermatol.*, vol. 64, no. 5, pp. 974–980, 2011.

[10] N. Villacorta, "The Mystery of the Missing Fingerprints," 2011.

[11] A. M. Knowles and J. Phys, "Aspects of physicochemical methods for the detection of latent fingerprints," *J. Phys. E.*, vol. 11, pp. 713–721, 1978.

[12] R. S. Ramotowski, *Composition of Latent Print Residue, in Advances in Fingerprint Technology*. CRC Press, Boca Raton, London, New York, Washington D.C., 2001.

[13] O. Keisar *et al.*, "Measuring the water content in freshly-deposited fingermarks," *Forensic Sci. Int.*, vol. 294, pp. 204–210, 2019.

[14] K. Sato, W. H. Kang, K. Saga, and K. T. Sato, "Biology of sweat glands and their disorders. I. Normal sweat gland function," *J. Am. Acad. Dermatol.*, vol. 20, no. 4, pp. 537–563, 1989.

[15] R. Croxton, M. Baron, D. Butler, T. Kent, and V. Sears, "Variation in amino acid and lipid composition in latent fingerprints," *Forensic Sci. Int.*, vol. 199, pp. 93–102, 2010.

[16] J. F. W. M. G. H. Bourne, "International Review of Cytology," *Acad. Press. New York*, vol. 1, pp. 265 – 304, 1952.

[17] P. E. Pochi, J. S. Strauss, and D. T. Downing, "Age-related changes in sebaceous gland activity," *J. Invest. Dermatol.*, vol. 73, no. 1, pp. 108–111, 1979.

[18] K. Saga, "Structure and function of human sweat glands studied with histochemistry and cytochemistry," *Prog. Histochem. Cytochem*, vol. 37, pp. 323–386, 2002.

[19] K. Biel, A. Wilke, L. Martin, and S. S. Terstegen, "A short history of sweat gland biology," *Int. J. Cosmet. Sci.*, vol. 29, no. 3, 2007.

[20] S. Michalski, R. Shaler, and F. Dorman, "The evaluation of fatty acid ratios in latent fingermarks by gas chromatography/mass spectrometry (GC/MS) analysis," vol. 58, pp. S215–S220, 2013.

[21] A. Girod, R. Ramotowski, and C. Weyermann, "Composition of fingermark residue: a qualitative and quantitative review," *Forensic Sci. Int.*, vol. 223, no. 1, pp. 10–24, 2012.

[22] K. Asano, C. Bayne, K. Horsman, and M. Buchanan, "Chemical composition of fingerprints for gender determination," *J. Forensic Sci.*, vol. 47, no. 4, pp. 805–807, 2002.

[23] D. Penn *et al.*, "Individual and gender fingerprints in human body odour," *J. R. Soc. Interface*, vol. 4, pp. 331–340, 2007.

[24] M. de Puit, M. Ismail, and X. Xu, "LCMS analysis of fingerprints, the amino acid profile of 20 donors," *J. Forensic Sci.*, vol. 59, no. 2, pp. 367–370, 2013.

[25] G. Mong, S. Walter, T. Cantu, and R. Ramotowski, "The chemistry of latent prints from children and adults," *Fingerpr. Whorld*, vol. 27, no. 104, pp. 66–69, 2001.

[26] P. Ramasastry, D. T. Downing, P. E. Pochi, and J. S. Strauss, "Chemical composition of human skin surface lipids from birth to puberty," *J. Investig. Dermatol.*, vol. 54, no. 2, pp. 139–144, 1970.

[27] M. Buchanan, K. Asano, and A. Bohanon, "Chemical characterization of fingerprints from adults and children," *SPIE (International Soc. Opt. Eng. Forensic Evid. Anal. Crime Scene Investig.*, vol. 2941, pp. 89–95, 1997.

[28] H. Tang, W. Lu, C. Che, and K. Ng, "Gold nanoparticles and imaging mass spectrometry: double imaging of latent fingerprints," *Anal. Chem.*, vol. 82, no. 5, pp. 1589–1593, 2010.

[29] S. Jickells, "Fingerprinting: into the future," *J. Dyn. Syst. Meas. Control.*, vol. 41, no. 8, pp. 243–247, 2008.

[30] M. West and M. Went, "The spectroscopic detection of drugs of abuse in fingerprints after development with powders and recovery with adhesive lifters," *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, vol. 71, pp. 1984–1988, 2009.

[31] P. Ronnie, S. Walker, M. Tahtouh, and B. Reedy, “Detection of illicit substances in fingerprints by infrared spectral imaging,” *Anal. Bioanal. Chem.*, vol. 394, pp. 2039–2048, 2009.

[32] B. Yamashita and M. French, *Fingerprint Sourcebook*. NCJ 225320, US Department of Justice, Washington, 2011.

[33] J. Bond, “Visualization of latent fingerprint corrosion of metallic surfaces,” *J. Forensic Sci.*, vol. 53, no. 4, pp. 812–822, 2008.

[34] M. J. Bailey *et al.*, “Chemical characterization of latent fingerprints by matrix-assisted laser desorption ionization, time-of-flight secondary ion mass spectrometry, mega electron volt secondary mass spectrometry, gas chromatography/mass spectrometry, X-ray photoelectron spec,” *Anal. Chem.*, vol. 84, no. 20, pp. 8514–8523, 2012.

[35] E. Hoffmann and V. Stroobant, *Mass spectrometry: Principles and Applications*, 3rd editio., vol. 1040, no. 2. Wiley, 2016.

[36] J. Zhang, W. Yu, J. Suliburk, and L. S Eberlin, “Will Ambient Ionization Mass Spectrometry Become an Integral Technology in the Operating Room of the Future?,” *Clin. Chem.*, vol. 62, no. 9, pp. 1172–1174, 2016.

[37] I. R. White *et al.*, “Real-time multi-marker measurement of organic compounds in human breath: Towards fingerprinting breath,” *J. Breath Res.*, vol. 7, no. 1, 2013.

[38] P. B. O’Connor and Y. Qi, “DATA PROCESSING IN FOURIER TRANSFORM ION CYCLOTRON RESONANCE MASS SPECTROMETRY,” *WHO Libr. Cat. Data*, no. i, pp. 221–235, 2011.

[39] S. Francese, “Techniques for fingermark analysis using MALDI MS - a practical overview,” in *Advances in MALDI and laser induced soft ionisation mass spectrometry*, R. Cramer, Ed. Springer, Cham, 2015.

[40] R. Wolstenholme, R. Bradshaw, M. Clench, and S. Francese, “Study of latent fingermarks by matrix-assisted laser desorption/ionisation mass spectrometry imaging of endogenous lipids.,” *Rapid Commun Mass Spectrom*, vol. 23, pp. 3031–3039, 2009.

[41] R. Bradshaw *et al.*, “Spectroscopic Imaging based approach for condom identifi- cation in condom contaminated fingermarks.,” *Analyst*, vol. 138, no. 9, pp. 2546–2557, 2013.

[42] R. Bradshaw, R. R, Wolstenholme, R. Blackledge, M. Clench, L. Ferguson, and S. Francese, “A novel matrix-assisted laser desorption/ionisation mass spectrometry imaging based methodology for the identification of sexual assault suspects.,” *Rapid Commun Mass Spectrom*, vol. 25, pp. 415–422, 2011.

[43] R. Bradshaw, W. Rao, R. Wolstenholme, M. Clench, S. Bleay, and S. Francese, “Separation of overlapping fingermarks by matrix assisted laser desorption ionisation mass spectrometry imaging.,” *Forensic Sci Int*, vol. 222, no. 1–3, pp. 318–326, 2012.

[44] R. Bradshaw, S. Bleay, M. Clench, and S. Francese, “Direct detection of blood in fingermarks by MALDI MS profiling and Imaging.,” *Sci Justice*, vol. 54, pp. 110–117, 2014.

[45] G. Groeneveld, M. De Puit, S. Bleay, R. Bradshaw, and S. Francese, “Detection and mapping of illicit drugs and their metabolites in fingermarks by MALDI MS and compatibility with forensic techniques.,” *Sci Rep*, vol. 5, p. 11716, 2015.

[46] K. Kaplan-Sandquist, M. LeBeau, and M. Miller, “Chemical analysis of pharmaceuticals and explosives in fingermarks using matrix-assisted laser desorption ionization/time-of-flight mass spectrometry.,” *Forensic Sci Int*, vol. 235, pp. 68–77, 2014.

[47] L. Ferguson, R. Bradshaw, R. Wolstenholme, M. Clench, and S. Francese, “Two-step matrix application for the enhancement and imaging of latent fingermarks.,” *Anal Chem*, vol. 83, no. 14, pp. 5585–5591, 2011.

[48] R. B. Cole, Ed., *Electrospray and MALDI Mass Spectrometry*, Second Edi. John Wiley & Sons, Inc., 2010.

[49] M. Scigelova, M. Hornshaw, A. Giannakopulos, and A. Makarov, “Fourier transform mass spectrometry,” *Mol. Cell. Proteomics*, vol. 10, no. 7, pp. 1–19, 2011.

[50] S. C. Beu, M. W. Senko, J. P. Quinn, F. M. Wampler, and F. W. McLafferty, “Fourier-transform electrospray instrumentation for tandem high-resolution mass spectrometry of large molecules,” *J. Am. Soc. Mass Spectrom.*, vol. 4, pp. 557–565, 1993.

[51] S. G Penn, M. T. Cancilla, and C. B. Lebrilla, “Collision-induced dissociation of branched oligosaccharide ions with analysis and calculation of relative dissociation thresholds,” *Anal. Chem.*, vol. 68, p. 2331, 1996.

[52] C. Q. Jiao, D. R. A. Ranatunga, and B. S. Freiser, “Reactions of Nbn+ (n = 1−15) with Propene and 1-Butene: Product Distribution as a Function of Cluster Size,” *J. Phys. Chem.*, vol. 100, no. 12, pp. 4755–4759, 1996.

[53] G. Baykut and J. R. Eyler, “Fourier transform ion cyclotron resonance mass spectrometry,” *Trends Anal. Chem.*, vol. 5, no. 2, pp. 44–49, 1986.

[54] J. Amster, “Fourier transform mass spectrometry,” *J. MASS Spectrom.*, vol. 31, no. 12, pp. 1325–1337, 1996.

[55] M. Ringnér, “What is principal component analysis?,” *Nat. Biotechnol.*, vol. 26, no. 3, pp. 303–304, 2008.

[56] S. Maitra and J. Yan, “Principle Component Analysis and Partial Least Squares : Two Dimension Reduction Techniques for Regression,” *Not A Pap.*, pp. 79–90, 2008.

[57] A. Rivas-Ubach, Y. Liu, T. S. Bianchi, N. Tolić, C. Jansson, and L. Paša-Tolić, “Moving beyond the van Krevelen Diagram: A New Stoichiometric Approach for Compound Classification in Organisms,” *Anal. Chem.*, vol. 90, no. 10, pp. 6152–6160, 2018.

[58] T. H. Trinh and Y. Uemura, “A theoretical equation presenting slope in van Krevelen diagram for biomass pyrolysis,” *Platform - A Journal of Engineering*, vol. 3, no. April. pp. 56–64, 2019.

[59] S. A. Brockman, E. V. Roden, and A. D. Hegeman, “Van Krevelen diagram visualization of high

resolution-mass spectrometry metabolomics data with OpenVanKrevelen," *Metabolomics 14*, vol. 48, 2018.

[60] R. Smith, *Understanding mass spectra: a basic approach*, 2nd editio. Wiley, 2004.

[61] "Metabocard for Eplerenone (HMDB0014838)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0014838#links. [Accessed: 26-Mar-2020].

[62] "Metabocard for Cyclandelate (HMDB0015586)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0015586. [Accessed: 26-Mar-2020].

[63] E. Fahy, D. Cotter, M. Sud, and S. Subramaniam, "Lipid classification, structures and tools," *Changes*, vol. 29, no. 11, pp. 997–1003, 2012.

[64] R. Bittman, "Glycerolipids: Chemistry.," *Encycl. Biophys. Springer, Berlin, Heidelb.*, 2013.

[65] K. R. Feingold, "The outer frontier: The importance of lipid metabolism in the skin," *J. Lipid Res.*, vol. 50, no. SUPPL., pp. 417–422, 2009.

[66] "Metabocard for MG(0:0/16:0/0:0) (HMDB0011533)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0011533. [Accessed: 26-Mar-2020].

[67] "Metabocard for MG(18:0/0:0/0:0) (HMDB0011131)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0011131. [Accessed: 26-Mar-2020].

[68] "Metabocard for MG(0:0/16:1(9Z)/0:0) (HMDB0011534)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0011534. [Accessed: 26-Mar-2020].

[69] "Metabocard for Melibiose (HMDB0000048)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0000048. [Accessed: 26-Mar-2020].

[70] "Metabocard for Methylgingerol (HMDB0029852)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0029852. [Accessed: 26-Mar-2020].

[71] "Metabocard for Palmitic acid (HMDB0000220)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0000220. [Accessed: 26-Mar-2020].

[72] "Metabocard for Oleic acid (HMDB0000207)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0000207. [Accessed: 26-Mar-2020].

[73] "Metabocard for Myristic acid (HMDB0000806)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0000806. [Accessed: 26-Mar-2020].

[74] "Metabocard for Squalene (HMDB0000256)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0000256. [Accessed: 26-Mar-2020].

[75] M. Pfeuffer and A. Jaudszus, "Pentadecanoic and Heptadecanoic Acids: Multifaceted Odd-Chain Fatty Acids," *Adv Nutr.*, vol. 7, no. 4, pp. 730–4, 2016.

[76] "Metabocard for Heptadecanoic acid (HMDB0002259)." [Online]. Available: https://hmdb.ca/metabolites/HMDB0002259. [Accessed: 26-Mar-2020].

[77] “Metabocard for Pentadecanoic acid (HMDB0000826).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0000826. [Accessed: 26-Mar-2020].

[78] “Metabocard for Linoleic acid (HMDB0000673).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0000673. [Accessed: 26-Mar-2020].

[79] S. Aryal, “Lipids - definition, properties, structure, types, examples, functions,” 2018. .

[80] Y. Jia, Y. Gan, C. He, Z. Chen, and C. Zhou, “The mechanism of skin lipids influencing skin status,” *J. Dermatol. Sci.*, vol. 89, no. 2, pp. 112–119, 2018.

[81] J. Herrera, M. Arevalo-Herrera, and S. Herrera, “Prevention of preeclampsia by linoleic acid and calcium supplementation: a randomized controlled trial,” *Obs. Gynecol.*, vol. 91, no. 4, pp. 585–90, 1998.

[82] M. G. Larson, “Analysis of variance,” *Circulation*, vol. 117, no. 1, pp. 115–121, 2008.

[83] C. Zaiontz, “Kruskal-Wallis Test.” [Online]. Available: https://www.real-statistics.com/one-way-analysis-of-variance-anova/kruskal-wallis-test/. [Accessed: 02-Oct-2020].

[84] “Metabocard for Panthenol (HMDB0004231).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0004231. [Accessed: 26-Mar-2020].

[85] “Metabocard for Tangeritin (HMDB0030539).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0030539. [Accessed: 26-Mar-2020].

[86] “Metabocard for Hexamethylquercetagetin (HMDB0029308).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0029308. [Accessed: 26-Mar-2020].

[87] “Metabocard for Isosakuranin (HMDB0029481).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0029481. [Accessed: 26-Mar-2020].

[88] “Metabocard for Limonin (HMDB0035921).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0035921. [Accessed: 26-Mar-2020].

[89] “Metabocard for (7alpha,10beta)-1(10->19)-Abeo-7-acetoxyisoobacun-3,10-olide (HMDB0035079).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0035079. [Accessed: 26-Mar-2020].

[90] “Metabocard for Naringin (HMDB0002927).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0002927. [Accessed: 26-Mar-2020].

[91] “Metabocard for Hesperidin (HMDB0003265).” [Online]. Available: https://hmdb.ca/metabolites/HMDB0003265. [Accessed: 26-Mar-2020].

[92] H. J. Simon *et al.*, “Uncoiling collagen: A multidimensional mass spectrometry study,” *Analyst*, vol. 141, no. 1, pp. 157–165, 2016.

[93] F. Floris *et al.*, “2D FT-ICR MS of Calmodulin: A Top-Down and Bottom-Up Approach,” *J. Am. Soc. Mass Spectrom.*, vol. 27, no. 9, pp. 1531–1538, 2016.

[94] M. A. van Agthoven *et al.*, "Differentiating Fragmentation Pathways of Cholesterol by Two-Dimensional Fourier Transform Ion Cyclotron Resonance Mass Spectrometry," *J. Am. Soc. Mass Spectrom.*, vol. 26, no. 12, pp. 2105–2114, 2015.

[95] M. A. van Agthoven and P. B. O'Connor, "Two-dimensional mass spectrometry in a linear ion trap, an in silico model," *Rapid Commun. Mass Spectrom.*, vol. 31, no. 8, pp. 674–684, 2017.

[96] M. A. van Agthoven *et al.*, "Two-Dimensional Mass Spectrometry for Proteomics, a Comparative Study with Cytochrome c," *Anal. Chem.2*, vol. 88, no. 8, pp. 4409–4417, 2016.

[97] M. A. van Agthoven, Y. P. Y. Lam, P. B. O'Connor, C. Rolando, and M. A. Delsuc, "Two-dimensional mass spectrometry: new perspectives for tandem mass spectrometry," *Eur. Biophys. J.*, no. Biemann 2002, 2019.

[98] S. N Dumas and J. M Ntambi, "A Discussion on the Relationship between Skin Lipid Metabolism and Whole-Body Glucose and Lipid Metabolism: Systematic Review," *Physiol. Behav.*, vol. 176, no. 10, pp. 139–148, 2017.

[99] V. Kruse, D. Neess, and N. J. Færgeman, "The Significance of Epidermal Lipid Metabolism in Whole-Body Physiology," *Trends Endocrinol. Metab.*, vol. 28, no. 9, pp. 669–683, 2017.

[100] W. P. Esler *et al.*, "Human sebum requires de novo lipogenesis, which is increased in acne vulgaris and suppressed by acetyl-CoA carboxylase inhibition," *Sci. Transl. Med.*, vol. 11, no. 492, pp. 1–14, 2019.

[101] K. Agrawal, L. A. Hassoun, N. Foolad, T. L. Pedersen, R. K. Sivamani, and J. W. Newman, "Sweat lipid mediator profiling: A noninvasive approach for cutaneous research," *J. Lipid Res.*, vol. 58, no. 1, pp. 188–195, 2017.

[102] M. Lovászi, A. Szegedi, C. C. Zouboulis, and D. Törőcsik, "Sebaceous-immunobiology is orchestrated by sebum lipids," *Dermatoendocrinol.*, vol. 9, no. 1, p. e1375636, 2017.

[103] M. Lovászi *et al.*, "Sebum lipids influence macrophage polarization and activation," *Br. J. Dermatol.*, vol. 177, no. 6, pp. 1671–1682, 2017.

[104] L. E. Gibson and R. E. Cooke, "A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocarpine by iontophoresis.," *Pediatrics*, vol. 23, pp. 545–549, 1959.

[105] E. A. Grice *et al.*, "Topographical and Temporal Diversity of the Human Skin Microbiome," vol. 324, no. 5931, pp. 1190–1192, 2010.

Appendix A – Identification file

S. R.

MINISTÉRIO DA JUSTIÇA

POLÍCIA JUDICIÁRIA

DIRECTORIA NACIONAL

Laboratório de Polícia Científica
Faculdade de Ciências da Universidade de Lisboa

Projecto FINGERPRINTS - Ficha de Identificação

V1

Nome: _____

Data de nascimento:_____/_____/_____

Género:_____

Raça:_____

Profissão:_____

Local de trabalho:_____

Medicação:_____

Doenças/estado de saúde:_____

Contacto:_____

_____

_____

V1- 1

Data da recolha : Hora:

Procedimento de recolha:

Observações:

_____

V1- 2

Data da recolha : Hora:

Procedimento de recolha:

Observações:

V1

V1- 3
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 4
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 5
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 6
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1

V1- 7
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 8
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 9
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 10
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1

V1- 11
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 12
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 13
Data da recolha : Hora:
Procedimento de recolha:

Observações:

V1- 14
Data da recolha : Hora:
Procedimento de recolha:

Observações:

Appendix B – Statistical Analysis

Table B.1. Kruskal-Wallis Test. *Kruskal-Wallis Test of all the compounds present in* ***Table 3.1****. Compound intensities were calculated using the natural-logarithmic transformation.*

Compound	3-(5,6,6-Trimethylbicyclo[2.2.1]hept-1-yl)cyclohexanol	3-Methylcyclopentadecanone	Monoisobutyl phthalic acid	Lauroyl diethanolamide	Cyclohexaneundecanoic acid	3-Oxooctadecanoic acid	N-Undecylbenzenesulfonic acid	Hexaethylene glycol
Count	93	93	93	93	93	93	93	93
r^2/n	224159,0625	218166,9722	225903,9167	225265,75	218310,2222	227654,5556	220377,0278	235474,0833
H-stat	25,69946809	17,47422405	28,09460078	27,21859986	17,67086098	30,49767406	20,50793106	41,23141158
H-ties	25,78699968	18,86468613	28,1197744	27,22266161	17,77909215	30,54781072	20,50808406	41,23141158
df	9	9	9	9	9	9	9	9
p-value	0,002213469	0,026368339	0,000910885	0,00128491	0,0378247	0,000353808	0,015023069	4,53942E-06
alpha	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05
Sig.	yes	yes	yes	yes	yes	yes	yes	yes

Compound	Cyclandelate	Sphingosine	alpha-CEHC	1,20-Eicosanediol	Heptaethylene glycol	MG (0:0/16:1(9Z)/0:0)	Melibiose	Octaethylene glycol
Count	93	93	93	93	93	93	93	93
r^2/n	219924,8333	241806,75	231592,3889	230927,9722	240142,5833	221556,2222	242398,75	246954,75
H-stat	19,88721116	49,92415923	35,90307329	34,99103943	47,63978495	22,12659193	50,73678792	56,99073439
H-ties	19,88721116	51,17485929	35,90307329	37,55125006	47,63978495	22,89432484	50,73678792	56,99073439
df	9	9	9	9	9	9	9	9
p-value	0,018621367	6,47181E-08	4,12477E-05	2,09769E-05	2,982E-07	0,006438061	7,82752E-08	5,07419E-09
alpha	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05
Sig.	yes	yes	yes	yes	yes	yes	yes	yes

Compound	Hypogeic acid	Stearoylcarnitine	Squalene	Eplerenone	Methylgingerol	2-Dodecylbenzenesulfonic acid	Pentadecanoic acid	N-Undecanoylglycine
Count	93	93	93	93	93	93	93	93
r^2/n	222835,5278	249626,3611	222902,3958	220170,3889	211567,5833	213876,5556	217172,1667	216329,0556
H-stat	23,88267368	60,65801113	23,97446237	20,22428125	8,415351178	11,58483947	16,10867078	14,95134599
H-ties	23,88267368	60,75772989	24,64631362	20,22956335	8,415351178	11,58518519	16,10867078	14,95134599
df	9	9	9	9	9	9	9	9
p-value	0,004490707	9,57636E-10	0,003388374	0,016547348	0,492890189	0,237717112	0,064646646	0,092281118
alpha	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05
Sig.	yes	yes	yes	yes	no	no	no	no

Compound	Palmitic acid	Heptadecanoic acid	Linoleic acid	Oleic acid	Stearic acid	(9Z,11E,13E,15Z)-4-Oxo-9,11,13,15-octadecatetraenoic acid	(2'E,4'Z,7'Z,8E)-Colnelenic acid
Count	93	93	93	93	93	93	93
r^2/n	212361,6389	216558,5556	216555,0278	214452,4722	212738,3333	212568,7222	210670,1944
H-stat	9,505338214	15,26637688	15,26153436	12,37539083	10,0224205	9,789598109	7,183520171
H-ties	9,505338214	15,26637688	15,3030903	12,37539083	10,0224205	10,44708051	7,407894647
df	9	9	9	9	9	9	9
p-value	0,391992209	0,083874154	0,082940014	0,192964152	0,348673195	0,315517578	0,594731119
alpha	0,05	0,05	0,05	0,05	0,05	0,05	0,05
Sig.	no	no	no	no	no	no	no

Compound	MG (0:0/16:0/0:0)	Polyoxyethylene (600) monoricinoleate	Lactapiperanol D	MG (18:0/0:0/0:0)	12-Ketodeoxycholic acid	Myristic acid	3-Hydroxy-6,8-dimethoxy-7(11)-eremophilen-12,8-olide
Count	93	93	93	93	93	93	93
r^2/n	209496,9444	210939,9722	213660,0278	211097,1944	214835,5417	216400,0556	213562,8889
H-stat	5,573019141	7,553839701	11,28761534	7,769656066	12,90122398	15,04880653	11,15427438
H-ties	5,607152984	8,672927064	12,04570657	8,291476481	12,90132022	15,04880653	11,18464659
df	9	9	9	9	9	9	9
p-value	0,778500592	0,46799697	0,210744838	0,505060299	0,167122493	0,08960404	0,263264452
alpha	0,05	0,05	0,05	0,05	0,05	0,05	0,05
Sig.	no	no	no	no	no	no	no

Table B.2. Benjamini-Hochberg Correction. *Benjamini-Hochberg Correction applied to the p-values obtained by the Kruskal-Wallis test (**Table B.1**). 38 tests were performed with an FDR of 0.05.*

test	*p-value*	*rank*	*alpha*	*sig*
Stearoylcarnitine	9,5764E-10	1	0,00131579	yes
Octaethylene glycol	5,0742E-09	2	0,00263158	yes
Sphingosine	6,4718E-08	3	0,00394737	yes
Melibiose	7,8275E-08	4	0,00526316	yes
Heptaethylene glycol	2,982E-07	5	0,00657895	yes
Hexaethylene glycol	4,5394E-06	6	0,00789474	yes
1,20-Eicosanediol	2,0977E-05	7	0,00921053	yes
alpha-CEHC	4,1248E-05	8	0,01052632	yes
3-Oxooctadecanoic acid	0,00035381	9	0,01184211	yes
Monoisobutyl phthalic acid	0,00091088	10	0,01315789	yes
Lauroyl diethanolamide	0,00128491	11	0,01447368	yes
3-(5,6,6-Trimethylbicyclo[2.2.1]hept-1-yl)cyclohexanol	0,00221347	12	0,01578947	yes
Squalene	0,00338837	13	0,01710526	yes
Hypogeic acid	0,00449071	14	0,01842105	yes
MG(0:0/16:1(9Z)/0:0)	0,00643806	15	0,01973684	yes
N-Undecylbenzenesulfonic acid	0,01502307	16	0,02105263	yes
Eplerenone	0,01654735	17	0,02236842	yes
Cyclandelate	0,01862137	18	0,02368421	yes
3-Methylcyclopentadecanone	0,02636834	19	0,02500000	no
Cyclohexaneundecanoic acid	0,0378247	20	0,02631579	no
Pentadecanoic acid	0,06464665	21	0,02763158	no
Linoleic acid	0,08294001	22	0,02894737	no
Heptadecanoic acid	0,08387415	23	0,03026316	no
Myristic acid	0,08960404	24	0,03157895	no
N-Undecanoylglycine	0,09228112	25	0,03289474	no
12-Ketodeoxycholic acid	0,16712249	26	0,03421053	no
Oleic acid	0,19296415	27	0,03552632	no
Lactapiperanol D	0,21074484	28	0,03684211	no
2-Dodecylbenzenesulfonic acid	0,23771711	29	0,03815789	no
3-Hydroxy-6,8-dimethoxy-7(11)-eremophilen-12,8-olide	0,26326445	30	0,03947368	no
(9Z,11E,13E,15Z)-4-Oxo-9,11,13,15-octadecatetraen-oic acid	0,31551758	31	0,04078947	no
Stearic acid	0,3486732	32	0,04210526	no
Palmitic acid	0,39199221	33	0,04342105	no
Polyoxyethylene (600) monoricin	0,46799697	34	0,04473684	no
Methylgingerol	0,49289019	35	0,04605263	no
MG(18:0/0:0/0:0)	0,5050603	36	0,04736842	no
(2'E,4'Z,7'Z,8E)-Colnelenic acid	0,59473112	37	0,04868421	no

MG(0:0/16:0/0:0)	0,77850059	38	0,05000000	no