Vanessa Sofia Santos Ventura

IDENTIFICATION OF BIOMARKERS FOR SCHIZOPHRENIA

Mestrado em Química

Departamento de Química

FCTUC

Setembro/2015

UNIVERSIDADE DE COIMBRA

Vanessa Sofia Santos Ventura

IDENTIFICATION OF BIOMARKERS FOR SCHIZOPHRENIA

Dissertação apresentada para provas de Mestrado em Química, Área de especialização em Controlo de Qualidade e Ambiente

Professor Doutor Alberto António Caria Canelas Pais

Doutora Sandra Cristina da Cruz Nunes

Doutor Bruno José Fernandes Oliveira Manadas

Setembro de 2015

Universidade de Coimbra

Agradecimentos

Ao concluir mais uma etapa da minha vida gostaria de começar por agradecer a quem a tornou possível.

Ao Professor Dr. Alberto A. C. Canelas Pais por ter aceite este projecto e por toda a ajuda.

À Dra. Sandra Nunes por me ter dado a oportunidade de realizar este projecto ao seu lado, pela preocupação constante, dedicação, ajuda e todo o apoio recebido ao longo do ano.

Ao Dr. Bruno Manadas pela oportunidade de entrar neste projecto, por me ter recebido no laboratório e por me permitir evoluir nos meus conhecimentos.

À Vera por toda a paciência nos momentos de desespero, pela dedicação, pelo apoio e pelo conhecimento transmitido.

Um obrigado também às minhas colegas de laboratório. À Catia e à Sandra por toda a ajuda, especialmente nos comandos de procura, foram preciosos. À Mariana pela boa disposição e simpatia recebida desde o início. À Carolina, ao Nuno, à Margarida e à Joana pelos inúmeros momentos de lazer, pelas gargalhadas e sobretudo pela amizade, um obrigado muito especial. E ainda, à Tânia Firmino, do Departamento de Química, por toda a ajuda e conhecimentos transmitidos nesta etapa final do projecto.

À Ana e a Mariana que mesmo distante, nunca desistiram de mim e permaneceram sempre ao meu lado. À Maria, ao Ricardo e ao Serrano pelos momentos partilhados ao longo de seis anos, porque estiveram do início ao fim e terminam esta etapa ao meu lado. Aos meus afilhados, Bernardo e Kika, porque mesmo ausente também nunca esqueceram a madrinha. Um obrigado a todos por estarem sempre ao meu lado.

À família por estar sempre presente e por ser sempre essencial nos momentos difíceis. Mas especialmente, à minha Mãe por todo o amor e dedicação, por fazer o possível e o impossível por mim, por ser o meu exemplo de vida e por me permitir chegar até aqui, se não fosse por ela não estaria hoje a concluir esta etapa da minha vida.

E por último, ao Diogo por tudo o que nos trouxe até aqui, mas especialmente pelo último ano que o passou ao meu lado, por tê-lo partilhado comigo, pela infinita paciência que teve para me aturar, que nem sempre foi fácil mas sobretudo pela peça essencial que é na minha vida para me ajudar a nunca desistir. Obrigado, és o melhor do mundo.

“The only way to do great work is to love what you do”

Steve Jobs

Abstract

Schizophrenia is a severe and chronic mental disorder that leads to loss of contact with reality. The treatment is based on the continuous administration of drugs in order to reduce the risk of new episodes. Being a brain pathology, the goal of the present work is the identification, in the prefrontal cortex, of molecules that may be potential biomarkers, which would be very helpful in the treatment of schizophrenia.

Prefrontal cortex samples from 16 mice (4 of control, 4 administered with haloperidol, 4 with citalopram and 4 with clozapine) were fractionated in two parts: soluble and membrane. Each one was further separated in proteins and metabolites. The membrane metabolites samples (studied in the present work) were prepared and analyzed resorting to HPLC-MS in reversed phase. Data treatment was made resorting to Principal Component Analysis, which allowed the direct visualization of data structure and also provided the relative positioning of the samples. The most discriminating peaks and a potential biomarker of schizophrenia were identified.

Resumo

Considerada como uma das doenças mais incapacitantes, a esquizofrenia caracteriza-se como uma perturbação mental grave e crónica que leva à perda de contacto com a realidade. A doença apresenta vários sintomas, que dependem da fase da doença. O tratamento farmacológico é administrado continuamente de forma a reduzir o risco de novos episódios. Uma vez que a manifestação da patologia ocorre a nível cerebral, pretende-se com este trabalho identificar, em amostras de córtex pré-frontal, moléculas que possam ser reconhecidas como possíveis biomarcadores. Esta identificação seria certamente um forte contributo para o tratamento clínico da esquizofrenia.

Amostras de córtex provenientes de 16 ratos (4 controlo, 4 administrados com haloperidol, 4 com citalopram e 4 com clozapina) foram fraccionadas em duas partes: solúvel e membranar. Cada uma destas foi depois separada em proteínas e metabolitos. As amostras de metabolitos de fase membranar (estudados no presente trabalho) foram preparadas e analisadas por HPLC em fase reversa acoplada a espectrometria de massa.

A simplificação dos dados recolhidos foi efectuada recorrendo a métodos quimiométricos clássicos, nomeadamente a análise de componentes principais, que permitiu a visualização da estrutura dos dados e do posicionamento relativo das amostras. Foram identificados, em cada caso, os picos mais discriminantes, o que permitiu a identificação de um possível biomarcador para a esquizofrenia.

Abbreviations

ACN – Acetonitrile

AD – Antidepressants

AP - Antipsychotics

APCI – Atmospheric-pressure chemical ionization

APPI - Atmospheric pressure photoionization

CA – Catecholamine

CI - Citalopram

CID – Collisional induced dissociation

CL - Clozapine

CT – Control

DA – Dopamine

DSM-IV - Diagnostic and Statistical Manual of Mental Disorders IV

ESI – Electrospray Ionization

FA – Formic acid

FGAs - First-Generation antipsychotics

GABA - Gamma-Aminobutyric Acid

HA - Haloperidol

HILIC – Hydrophilic interaction liquid chromatography

HPLC - High performance liquid chromatography

I.S. – Internal Standard

ICD-10 – 10^{th} International Classification of Diseases

ICR - Ion cyclotron resonance

IDA – Information Dependent Acquisition

LC-MS - Liquid Chromatography - Mass Spectrometry

LIT - Linear ion trap

m/z – Mass-Charge

MEOH - Methanol

MF - Membrane fraction

MS –Mass Spectrometry

MS/MS – Tandem mass spectrometry

MVA – Multivariate Data Analysis

NMR – Nuclear Magnetic Resonance

NPC – Normal Phase Chromatography

PC's – Principal components

PCA – Principal components analysis

PFC – Pre-frontal Cortex

Q – Quadrupole

q2 – second quadrupole or collision cell (CID)

QIT - Quadrupole ion trap

Q-TOF – Quadrupole Time-of-Flight

RT – Retention time

RPC – Reversed Phase Chromatography

SCZ – Schizophrenia

SF – Soluble fraction

SGAs – Second-Generation antipsychotics

SNRIs – Norepinephrine Reuptake Inhibitors

SPE – Solid phase extraction

SPSS - Predictive analytics software and solutions

SSRIs – Selective Serotonin Reuptake Inhibitors

TEAB – Tetraethylammonium bromide

TOF - Time-of-Flight

WHO - World Health Organization

Table of contents

1. Introduction

1.1. Schizophrenia

In 1911 Eugen Bleuler introduced the term "schizophrenia" as a description of this mental illness replacing Kraepelin's term *dementia praecox* [1]. Since its demarcation and labeling as dementia praecox by Kraepelin (1887) and schizophrenia by Eugen Bleuler (1911), both definitions and scope have varied [2]. The process of splitting in schizophrenia according to Bleuler is the same as splitting of psychic connections in hysteria that in an extreme version can lead to the emergence of alter personalities and typical amnesia. In his Textbook of psychiatry he wrote (Bleuler, 1924): ''It is not alone in hysteria that one finds an arrangement of different personalities one succeeding the other. Through similar mechanism schizophrenia produces different personalities existing side by side.'' [1].

Schizophrenia (SCZ) is a complex psychiatric disorder with a heterogeneous clinical phenotype [3] which affects about 1% of the population and the understanding of its etiology remains incomplete [4]. It is characterized by an admixture of positive, negative, cognitive, mood, and motor symptoms whose severity varies across patients and through the course of the illness [2]. Presently, SCZ is not considered a single disorder but a group of conditions with manifestations common to other psychiatric and non-psychiatric disorders [4]. This disease is caused by the additive and interactive effects of genetic and non-genetic factors. Several genes have been identified as affecting brain development, plasticity and function, which may increase vulnerability to environmental stressors such as social adversity and illicit drug abuse.

Antipsychotic (AP) medication is the mainstay of pharmacological treatment [5]. The current medical treatment of SCZ consists of drugs acting at multiple receptors, but how the modulation of each of these receptor targets contributes to the anti-psychotic effects is still poorly understood [6]. Currently, available AP drugs modify neurotransmission primarily in dopaminergic pathways. The long-term effectiveness of antipsychotics is limited by a combination of inadequate efficacy and poor tolerability [5].

For clinical purposes, a biomarker[1] which might be detected by non-invasive methods would be desirable. To this end, several authors have identified possible biomarkers from

[1] A 'biomarker' is a biologic characteristic objectively measured and evaluated as an indicator of normal or pathogenic processes; or of response to a treatment or challenge (Group, 2001).

peripheral blood, including neurotransmitter metabolite levels, stress hormones, markers of immune response and fatty acids [7].

Drug discovery for SCZ is hampered by poor success rates and the lack of success is mainly due to the complexity of the disease characterized by a diverse symptomatology that is impossible to treat with a targeted approach. The available agents are efficient for psychosis but do not adequately address other core domains of schizophrenia psychopathology, namely negative symptoms and cognitive impairment. To proceed in the research, the industry has followed new paths and applies new methods to discover innovative medicines for SCZ. Undoubtedly, the discovery and development of novel anti-schizophrenic drugs will benefit from the study of the effects on biomarkers and from the correlation between these effects and the changes in patients behavior [6].

Schizophrenia is one of most debilitating mental illnesses with chronic psychotic symptoms and presents a wide range of symptoms affecting most of the domains of brain function [3, 8]. Core symptoms of schizophrenia were traditionally divided into two groups: positive symptoms and negative symptoms [3].Positive symptoms refers to a cluster of symptoms that are abnormal by their presence [8]. These involve impaired reality testing and include delusions, hallucinations, and other reality distortions. Several kinds of delusions can occur and they can have varying degrees of persistence and systematization, and influence the individual's functioning to different extents. Negative symptoms reflect the absence of certain normal behaviors and emotions; they include flat affect, apathetic social withdrawal and poverty of speech. Negative symptoms involve a blunting or loss of a range of affective and cognitive functions. These include impairments in affective experience and expression, abulia (loss of motivation), alogia (poverty of speech), anhedonia (inability to experience pleasure), avolition (lack of initiative), apathy (lack of interest), and reduced social drive [2].

These categories can be complemented with an additional group of cognitive dysfunction that include symptoms such as impaired attention, information processing, learning and memory [3].

The Course of SCZ

The course of schizophrenia can be divided into premorbid, prodromal, first-episode, and chronic phases [9]. The course of SCZ is typically characterized by psychotic exacerbations or relapses alternating with periods of partial remissions [3]. In premorbid phase patients

often have a subtle and nonspecific cognitive, motor and/or social dysfunction [5]. In prodromal phase they have a gradual onset of symptoms, misperceptions, over-valued beliefs, ideas of reference, prior to the onset of psychotic symptoms [9]. The first psychotic episode indicates the formal beginning of SCZ, and finally, there is a stable phase, when psychotic symptoms are less prominent, and negative symptoms and cognitive deficits more predominant [2].

1.1.1. Epidemiology and Pathophysiology

The World Health Organization (WHO) consider SCZ as the seventh greatest cause of disability worldwide being related both to genetic risk factors and environmental ones [5, 8, 10].

This disease does not affect only mental health. Patients with a diagnosis of SCZ die in average 12–15 years before the average population, and this difference have increased in recent decades [11]. Although variable degrees of recovery occur [12], complete cure is unusual and in average an affected person lives with SCZ approximately 30 years [13]. The annual incidence of this disorder is 1–7/10,000 worldwide, being 2–6/10,000 in Europe, with an individual risk is 0.36–1.87% [5, 8]. Some deaths are suicides, but the main reasons for mortality are physical causes, resulting from decreased access to medical care and increased frequency of routine risk factors, such as poor diet, little exercise, obesity, and smoking [11].

According to current knowledge, SCZ is a complex disease caused by multiple etiological factors like urbanicity, male gender and environmental factors such as community demands. The available resources and treatment significantly alter the course of the disease and provides compelling evidence supporting a role for social factors in its etiology; the specific risk mediating factor social or biological, however, remains to be elucidated [7, 13].

The pathogenesis of this illness is still largely unknown [7] and in the absence of a biological marker, the current diagnosis of SCZ and its treatment are mainly based on clinical questionnaires. Thus, it is not surprising that the response rate is unsatisfactory, in particular after multiple treatment attempts, and relapse is common for patients who discontinue medication [14].

The US-based 4th Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV) [15] and the 10th International Classification of Diseases (ICD-10) [16] are currently used to

diagnose SCZ. As referred above, the absence of an objective test for diagnosis has boosted the studies aiming at the identification of a biomarker that could be detected by non-invasive methods [5, 7, 17]. The potential utility of neuronal biomarkers such as neurotransmitter metabolite levels, stress hormones, markers of immune response and fatty acids is a topic that, recently, has attracted much interest [17-20].

Currently, resorting to modern neuroimaging techniques, the study of the pathophysiological changes of SCZ is possible [5]. Structural brain imaging has shown a subtle, almost universal, decrease in grey matter, enlargement of ventricles, and focal alteration of white matter tracts [21-23], indicating that the hippocampus [24] and cortex [25] have a central role in the neuropathology and pathophysiology of SCZ. These observations support the idea that SCZ is associated with altered brain function. Altered neurotransmission is a key pathophysiological mechanism underlying the expression of schizophrenic symptoms and for decades, pathophysiological studies relating to schizophrenia were focused on disturbances of dopaminergic and glutamatergic though GABAergic, serotonergic, cholinergic and opioid transmitter systems have also been implicated [4, 8].

Dopamin hypothesis

Abnormalities in dopaminergic systems are thought to be the bases for some neuropsychiatric disorders such as Parkinson's disease and SCZ [26]. Dopamine is a catecholamine (CA) neurotransmitter that regulates functional network activities in various regions of the brain [27] and participates in the regulation of motor functions and of cognitive processes such as learning and memory [26].

GABAergic and Glutamatergic hypothesis

In 2009, Howes and Kapur postulated a version of DA hypothesis that focuses on the modulating effect of other neurotransmitters such as glutamate and gamma-aminobutyric acid (GABA) in the DA system [7]. The contribution of this acid, GABA, in cognitive function has received increasing attention in diseases such as SCZ [28-30].

Postmortem studies revealed the existence of a relationship between cognition and some alterations in GABA receptors. The inhibition of GABAergic neurotransmission plays a major source since it mediates several cognitive operations including information processing and memory [31].

The glutamate hypothesis in SCZ focuses on disturbances in brain glutamatergic pathways and impairment in signaling at glutamate receptors [32, 33]. It is an alternative or complementary theory to the dopamine hypothesis [34, 35].

Serotonin hypothesis

According to Allman, the serotonergic participates in our emotions being associated with the regulation of anxiety, stress, and mood [36, 37].

Cholinergic hypothesis

The role of the cholinergic system in SCZ remains controversial. Several researches revealed that pharmacological manipulation in cholinergic system allow to observe changes in SCZ, suggesting that it has influence on positive and negative symptoms [38].

1.2. Therapeutic approaches in schizophrenia

There are a wide range of available antipsychotic drugs that differ in chemical structure and receptor profile. In spite of that, all AP drugs modify dopaminergic transmission in the brain [39]. These drugs are effective for delusions and hallucinations but less efficient for disabling cognitive and motivational impairments [5], however their main benefit is to reduce the risk of "relapse" [9].

Antipsychotic drugs can be divided into two categories, the first and second generation of AP. The First–Generation antipsychotics (FGAs), also called typical or conventional AP, started with the introduction of Chlorpromazine, in 1950s and prompted the development of several other antipsychotic drugs. These opened the new era of modern AP [40]. In treatment of SCZ, these drugs are used as a primary choice and many studies have concluded that they can be applied to treat patients with psychotic symptoms [41]. So far, the typical AP significantly decrease the positive symptoms, 60-70% [39]. The most known example of FGAs is Haloperidol (HA) which is widely used in treatment of SCZ and also in the treatment of delirium and other situations that include the control of the symptoms of acute psychosis, among others [42-44].

Haloperidol was firstly synthesized in 1958, but it was only introduced in the market in 1967, in the United States [43]. Initially the drug was used in the field of anesthesia and its

introduction as AP only came after the research of Ayd and Settle [45], in which the beneficial effects on hallucinations, delusions, aggression, impulsiveness and states of excitement was demonstrated. This AP is still known by its high capacity of dopamine blockade in comparison with other existing antipsychotics and due to its smaller doses in comparison with other drugs. According to these features it is considered as one of the most successful drugs to achieve antipsychotic effects [46].

Another type of AP is the second-generation (SGAs) also called atypical drugs. The use of this AP drugs started in 1970s, with the development of Clozapine (CL), and has increased recently [47, 48]. In spite of some controversy in the topic, some authors argued that the SGAs have a higher efficiency, with fewer side effects and less negative symptoms than de FGAs [49-51].

Clozapine was, as stated above, the first atypical drug or SGAs to be implemented in clinical practice [42]. Its implementation occurred almost half a century ago, and in spite of some mishaps [52], this drug return to market, in 1980, and revolutionized the world of antipsychotics. In 1990 Kane and his colleagues compared the effect of CL and chlorpromazine, and concluded that 70% of the patients responded positively to the treatment with CL and only 5% to the chlorpromazine. After this reference study, it was recognized that clozapine had a different spectrum profile of therapeutic effects that have not been recognized in other antipsychotics. Since then, other new second-generation antipsychotics, such as Olanzapine, Quetiapine and Risperidone, for example, have been developed [53-56]. The success of CL was mainly due to the improvements in both positive and negative symptoms of SCZ. Unlike other drugs, the use of CL is not associated with the risk of movement disorders [47]. It also features a strong affinity with dopamine and serotonin systems showing a locking effect [42] which contributes to its efficacy and increased therapeutic action [55].

Another treatment option in SCZ is the use of antidepressants (AD). Presently, the use of AD is the most used treatment, with almost 233 million prescriptions in 2007 [57].

Until the 1970s, AD were divided into two categories: tricyclic antidepressants and monoamine oxidase inhibitors [58]. Currently the various classes of antidepressants commonly used are the selective serotonin reuptake inhibitors (SSRIs), selective serotonin and norepinephrine reuptake inhibitors (SNRIs) and atypical antidepressants. Selective serotonin reuptake inhibitors are the most frequently used for the treatment of severe depression and

obsessive compulsive disorder, among others. The SNRIs behave similarly to SSRIs. Both groups are considered agents usually well tolerated and safe. They are also effective in anxiety disorders and depression but they are not yet widely studied as the SSRIs [41].

One of the most commonly used antidepressants and SSRIs is Citalopram (CI) which is considered the more selective [59]. Although the drug is commercialized as a racemic mixture of (S) - (+) and (R) - (-), the SSRI activity resides essentially in the S- form being the R-enantiomer practically devoid of inhibitory potency of serotonin reuptake. This discovery led to the development of the drug as a single enantiomer, escitalopram (S-citalopram), which has a faster effect and effectiveness than the racemic mixture [60]. Citalopram is mostly used in the treatment of psychiatric disorders, but it is also suited to treat anxiety or eating disorders [57].

2. Technical Strategies

2.1. Metabolomic Studies

Metabolomics is the study of metabolism at the global level and its main goal is the identification and quantification of metabolites (dynamic set of all small molecules (<1500 Da)) present in organisms or biological samples. The concept of metabolomics covers the global analysis of all metabolites in a sample. In this project, pre-frontal cortex samples will be used, and the analysis will be focused on the metabolomic responses to HA, CL and CI drugs. These studies aim at helping in diagnose of diseases [61-64]. Currently, there are two complementary approaches for metabolomic studies: metabolomic profile and metabolomic fingerprinting [65]. The central objective of the first is the analysis of a group of metabolites that are related to a specific metabolomic pathway. The results of these approaches are quantitative and ideally independent of the technology used for data acquisition, and the data can be used to build or enlarge databases. The second approach does not intend to identify metabolites, but rather it compares patterns or 'fingerprints' of metabolites that change in response to disease [66]. Both strategies can be used in the search for new biomarkers as indicators of disease traits (or markers of risk), disease states, or illness progression [63, 66].

In SCZ, the identification of biomarkers is particularly important to improve diagnostics and therapy since there are no validated biomarkers [67]. Only in a few studies the metabolomics profile of SCZ was assessed. However, none of these succeed in validating a biomarker [61, 64, 68, 69].

Although metabolomics studies often rely on nuclear magnetic resonance (NMR) and mass spectrometry (MS) data, the latter is more often used, due to its higher sensitivity [70] and higher selectivity [62]. The coupling of MS with liquid chromatography (LC-MS) and in particular with high performance liquid chromatography has proven to be very successful in this area. HPLC separation is the most versatile method, because it allows the separation of compounds with a wide range of polarity [62].

Mass spectrometry uses versatile mass analyzers operating in tandem or hybrid configuration to more effectively perform the identification of metabolites by acquiring MS/MS spectra. The fragmentation of ions is made through collisional induced dissociation (CID) in a collision cell after precursor isolation in the quadrupole TOF (Q-TOF) instrument. The success of the separation also depends on the right choice of the HPLC columns. For semi-polar compounds the choice generally relies on the use of reversed phase C18

chromatography (RPC), while for polar compounds the hydrophilic interaction liquid chromatography (HILIC), that uses generally polar columns is the preferred one[62]. In what concerns sample preparation, solid phase extraction (SPE) has become one of the most important techniques in order to clean the samples [66].

2.2. Liquid Chromatography – Mass Spectrometry

Liquid chromatography coupled to mass spectrometry is a powerful analytical technique that gives qualitative and quantitative information about the samples to be analyzed. It is widely use in various fields including proteomics and metabolomics. This technique combines the physical separation capabilities of liquid chromatography (or HPLC) with the mass analysis capabilities of MS. The coupling with MS provides a better solution for the analysis of complex mixtures allowing, for example, the separation of isomers [71, 72].

2.2.1. High performance liquid chromatography

Liquid chromatography is an analytical technique that allows the separation of the components of a mixture based on their relative affinities with the stationary and mobile phases. In a simplified form, this technique involves the introduction of the sample into the column in which the compounds are distributed according to their affinity with the stationary phase [73]. The chromatographic separation is based on the different interactions occurring between the components of the mixture and the stationary and mobile phases. HPLC operates in two main modes, the normal phase chromatography (NPC) and reverse phase chromatography (RPC). In NPC, the least polar compounds elute first and more polar compounds elute last. Reversed-phase chromatography employs a polar mobile phase, and as a result the less polar molecules in the mobile phase tend to adsorb to the stationary phase, while polar molecules in the mobile phase will pass through the column and are eluted first. The mobile phase can also be of intermediate polarity by using a mixture of organic solvent with an extra level of water (e.g. ACN). A common example the stationary phase is RMe_2SiCl, where R is an alkyl chain such as $C_{18}H_{37}$ or C_8H_{17}. In this case the retention time is longer for non-polar molecules, while the polar ones are eluted more quickly [73].

Another method, less known is the hydrophilic interaction liquid chromatography (HILIC). It allows an efficient separation of smaller polar compounds in polar stationary phases. Several studies show that the HILIC mode is an "evolution" of NPC, however their mechanism of separation is more complicated. HILIC uses polar stationary phases with reversed-phase type eluents.

It is commonly believed that in HILIC, the mobile phase forms a water-rich layer on the surface of the polar stationary phase vs. the water-deficient mobile phase, creating a liquid/liquid extraction system. The analyte is distributed between these two layers. The more polar compounds will have a stronger interaction with the stationary aqueous layer than the less polar compounds. Thus, a separation based on compounds polarity and degree of solvation takes place. This mode shows some differences from the RPC, like the elution order of compounds that acts in the opposite direction, that is, from the less polar to more polar [74].

Mass Spectrometry

In mass spectrometry the sample components are ionized resulting in charged molecules or molecule fragments and their mass-to-charge (m/z) ratios are measured. In LC-MS, the MS works as a high sensitive and sophisticated detector [75]. In Figure 1 the main components of a mass spectrometer are schematically represented. These include sample introduction device, ionization source where the formation of gaseous ions occurs, mass analyzer that separates the ions formed according to their m/z, detection and quantification of ions and data acquisition [76].

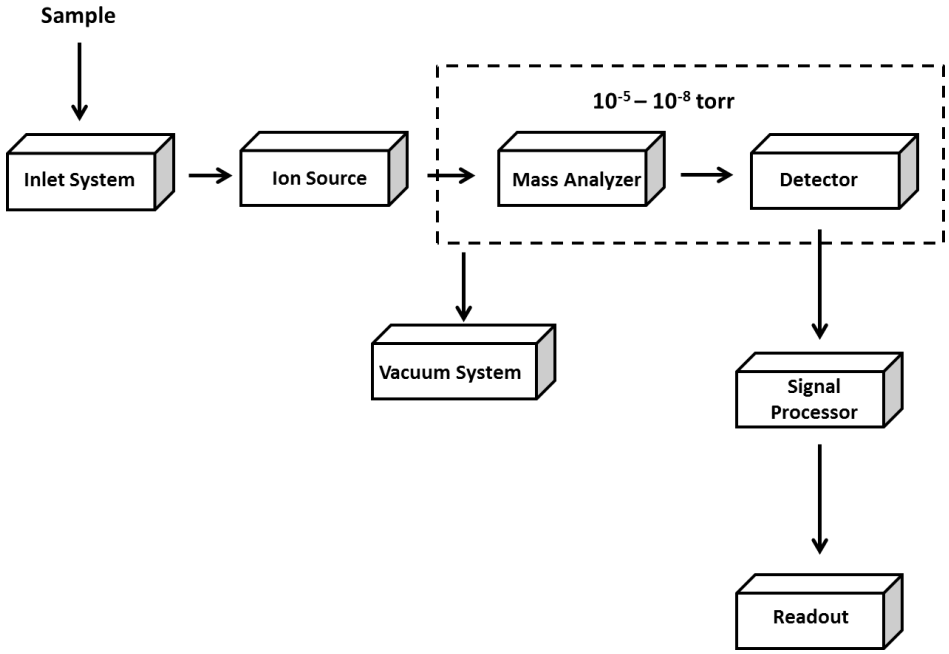

Figure 1 – Schematic representation of the components of a Mass Spectrometer. Adapted from [76]

Ion Source

The ion source is responsible for the ionization of the analyte into molecular ions. This step can occur by two different modes: (i) the hard mode, using ionization energy sources resulting in a high degree of fragmentation; (ii) the soft mode in which a low degree of fragmentation is obtained [75].

The most common LC-MS ionization sources are electrospray ionization (ESI), atmospheric-pressure chemical ionization (APCI) and atmospheric pressure photoionization (APPI). Over the last decade, electrospray ionization has become a very powerful technique capable of analyzing both small and large molecules of various polarities in complex biological samples [77]. It operates on the soft mode at atmospheric pressure and at room temperature, either in positive and negative ionization modes. Initially, the sample is nebulized by a N_2 flow and subjected to a 3-5 kV voltage to create an electrically charged spray of droplets. The subsequent evaporation of the solvent from the charged droplet makes it unstable upon reaching its Rayleigh limit. At that point, the droplet deforms as the electrostatic repulsion overcomes the surface tension holding the droplet together. The droplet undergoes Coulomb fission, and gives rise to many smaller, more stable droplets. The new droplets undergo desolvation and subsequently further Coulomb fissions (Figure 2) [78].

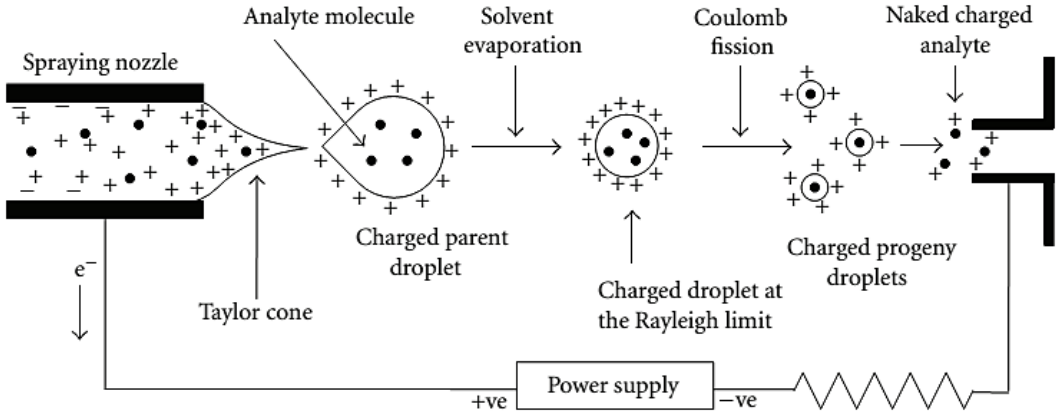

Figure 2 – Schematic representation of the electrospray ionization process. Adapted from [79].

Electrospray is considered a soft ionization technique, since fragmentation is low. This can be an advantage since the molecular ion is always observed, however little structural information can be gained from the simple mass spectrum obtained. This disadvantage can be overcome by coupling ESI with tandem MS (ESI-MS/MS), in which the fragmented ions after passing through the collision cell (CID) may undergo to analysis via a second analyser, the tandem mass spectrometer (MS/MS or MS^2) [77].

Mass Analyzer

All mass spectrometers combine ion formation, mass analysis and ion detection. Mass analyzers are used to separate ions according to their m/z ratio. Each mass analyzer has its own special characteristic and applications and its own benefits and limitations. Some mass analyzers are quadrupole (Q), quadrupole ion trap (QIT), linear ion trap (LIT), orbitrap, time-of-flight (TOF), ion cyclotron resonance (ICR) and magnetic/sector (less used). The performance of each analyzer takes into account aspects such as the speed of analysis, transmission, accuracy and mass resolution [80].

The quadrupoles are the most used mass analyzers [81]. They consist of four cylindrical electrodes (rods), parallel to each other (see Figure 3). In a quadrupole mass spectrometer, the quadrupole is the component of the instrument responsible for filtering sample ions based on their m/z ratio. Ions are separated in a quadrupole based on the stability of their trajectories in the oscillating electric fields that are applied to the electrodes.

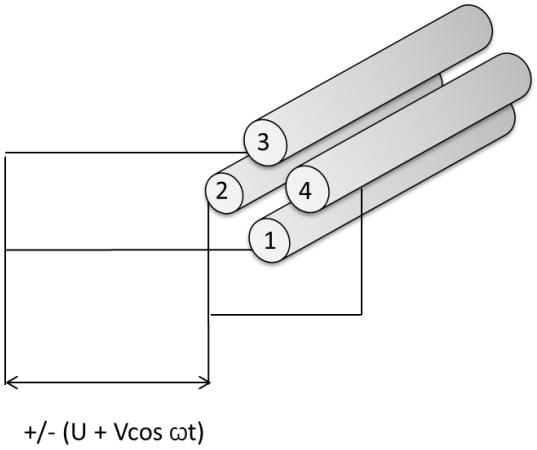

Figure 3 - Schematic representation of a quadrupole mass analyser. Adapted from [82]

Each opposing electrode pair is connected electrically and a radiofrequency voltage is applied between each pair of electrodes. A direct current voltage is then superimposed on the radiofrequency voltage. The ions travel down the quadrupole between the electrodes and only ions of a certain m/z ratio will reach the detector for a given ratio of voltages. Others, have unstable trajectories and will collide with the electrodes, being unable to reach the detector since the amplitude of oscillations became infinitive (see Figure 4). This allows scanning for a range of m/z values by continuously varying the applied voltage [75, 82].

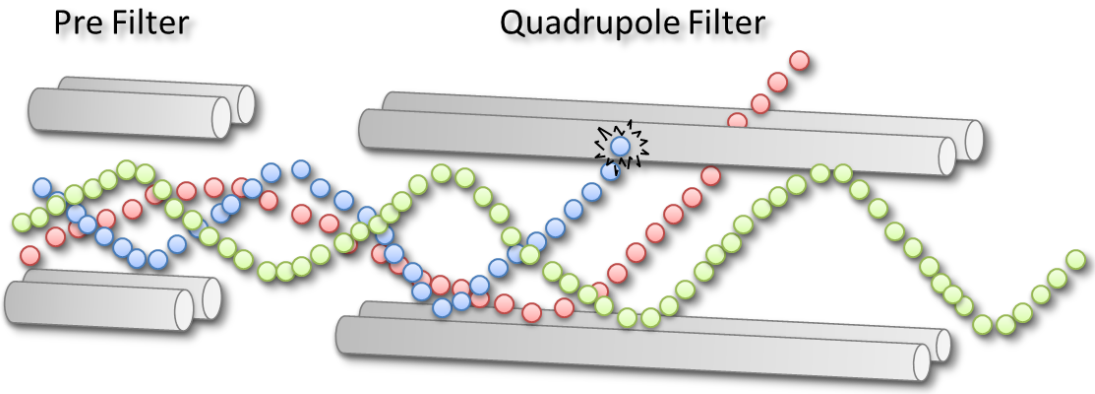

Figure 4 – Scheme of paths capable of acquiring from the filter action of the quadrupole. The blue circles represents m/z values with unstable trajectories that collide with the electrodes, being unable to reach the detector; the other ions represent stable trajectories thus being able to reach the detector. Adapted from [83]

To solve some specific problems, the MS/MS is a good solution and the most used in LC-MS, especially in biological compounds. This tool offers a faster screening of the sample, and offers a much higher sensibility [84]. In a MS/MS instrument, the system can be constituted by

three quadrupoles. The first and the third act as mass filters and the middle quadrupole (q2) is employed as a collision cell (Figure 5) [78].

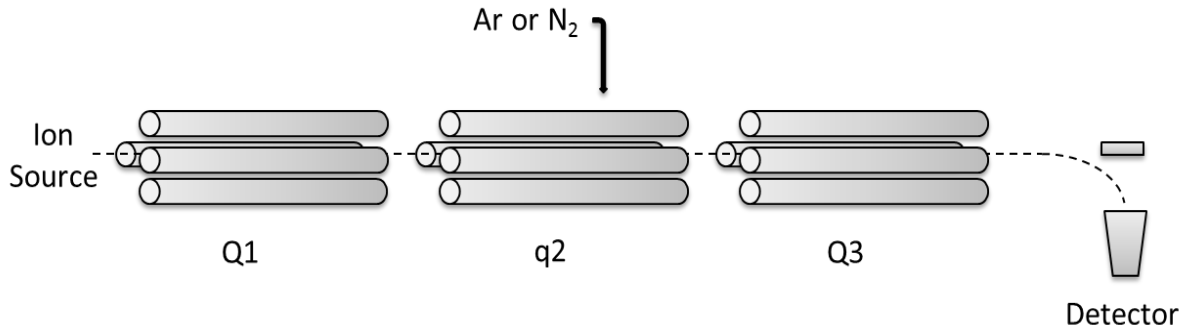

Figure 5 – A triple quadrupole mass spectrometer scheme. Q1 and Q3 act as mass filters and can be independently fixed, scanned or stepped; q2 is a collision cell that contains a low pressure inert gas. Adapted from [85].

The evolution of MS/MS leads to spectrometers in which the final quadrupole is replaced by a time-of-flight device as the second mass selection stage originating a hybrid instrument - Quadrupole time of flight (Q-TOF). This has the advantage of having a higher resolution. In TOF, the ions are accelerated by and electric field to the same kinetic energy and the velocity of the ion depends on their m/z ratio. The basic function of TOF is to measure the velocity, from which the m/z ratio can be determined. Therefore, TOF accelerate the ions according to their kinetic energy allowing the smaller ions to reach the detector faster [75, 82].

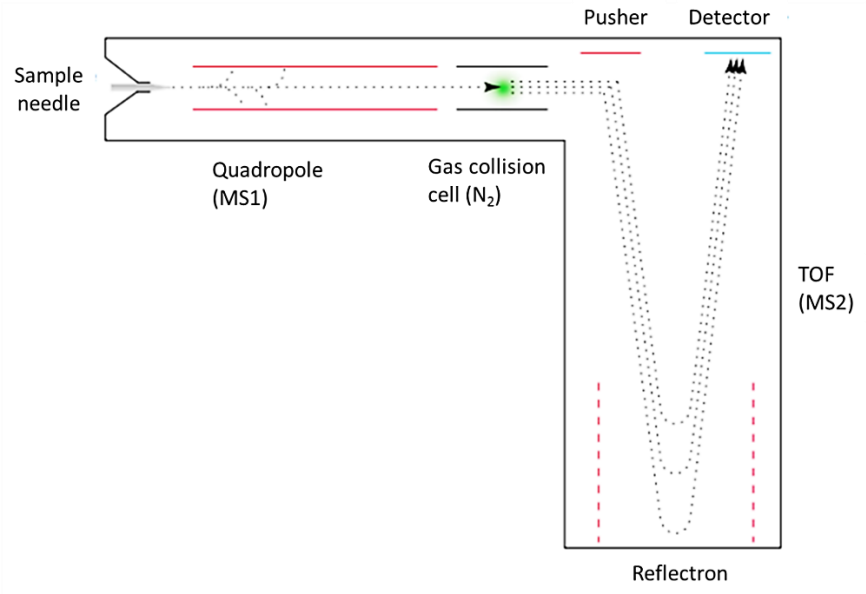

Figure 6 – Electrospray tandem mass spectrometer with TOF. Adapted from [86]

2.2.2. HPLC-MS data processing

In recent years, metabolomics has grown as an important tool for the analysis of biological systems and, particularly, in the diagnosis of diseases such as SCZ. The work done in this field produces a huge amount of data and therefore the identification and quantification of the metabolites requires a very careful and thorough manipulation of the dataset collected. The HPLC-MS data processing is an area still in development [87]. Data treatment can be divided in two steps: data processing and data analysis. The first consists essentially in transforming raw data into an accessible format while the latter consist in the interpretation of the information through multivariate analysis [88, 89].

In this project the main steps of data processing were peak detection (finds the peaks); alignment (data processing step specific to profiling experiments) and finally normalization (to reduce systematic errors by adjusting the intensities within each sample run) [87].

2.3. Data Analysis

The type of study performed in this project produce a huge amount of data being often difficult to choose the most relevant information. Multivariate data analysis and chemometrics methods have provided powerful tools for metabolomics data processing [90]. Chemometrics was introduced in 1972 by Svante Wold and Bruce R. Kowalski and includes a variety of mathematical and statistical methods to design and select procedures or experiments to simplify and characterize the system under study. The evaluation and interpretation of the relevant information aims at identifying the underlying patterns, and expressing the data in such a way as to highlight their similarities and differences. To this end different methods can be applied, such as cluster analysis and principal component analysis, among many others.

2.3.1. Principal Components Analysis

Principal Component Analysis (PCA) is an unsupervised technique that aims to reduce the dimensions of the dataset with minimal loss of information [91, 92]. PCA computes a compact and optimal description of the dataset, providing a roadmap to lay out a complex

dataset to a lower dimension and reveal the hidden, simplified structure that often is underlying.

PCA allows essentially to (i) extract the key information acquired from the data set, (ii) reduce the data dimensionality, keeping the most relevant information, (iii) establish the relationship between observations and variables, (iv) visualize and interpret the target system. In simple terms, the PCA procedure involves an orthogonal linear transformation that transforms the data into a new coordinate system in which the greatest variance lies on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on (see Figure 8).

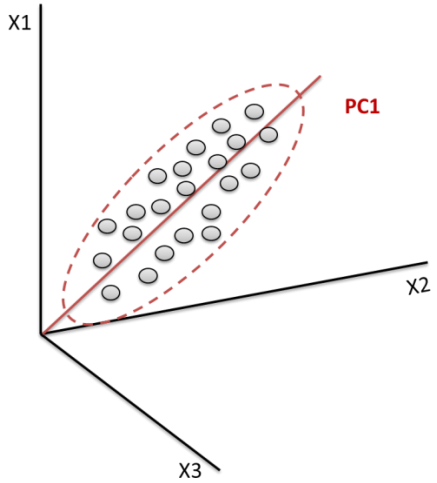

Figure 7 – Schematic representation of the PCA procedure. The greatest variance of the data lies on the first principal component.

These new variables called *Principal Components* (PC's) are obtained by linear combinations of the original variables, and these should gather most of the variability of the original dataset. The coordinates of the data in the new reference system are called “scores”, while the coefficient of the linear combination that describes each principal component (PC), or the weights of the original variables in each PC are called “loadings”. The principal components can be ordered by descending order of their variance. The PC´s with the higher variance are the ones that describe more properly the system. The most influential variables in the system are highlighted, and the most relevant factors may be identified. PCA is based on the assumption that most of the information about the structure of the data is contained in the directions along which the variations are the largest [91, 92]. Specifically, the original multi-dimensional space, defined by the parameters characterizing each sample, is contracted

into a few descriptive dimensions, which represent the main variation in the data. Each principal component can be displayed graphically and analyzed separately, and its meaning may often be interpreted according to simple descriptors. Essentially, the procedure is carried out by a linear transformation of the *m* analytical parameters $\mathbf{x_i}$ into a new set, the principal components u_i

$$u_i = w_{i1}x_1 + w_{i2}x_2 + \cdots + w_{im}x_m \quad (1)$$

where $w_{i1}. . .w_{im}$ are the loadings, i.e. the weights of the observed parameters in the linear combination. In PCA, the original matrix X ($n \times m$), in which *n* rows correspond to *n* samples and *m* columns correspond to the number of variables, is decomposed as a product of two matrices,

$$X = SW^T \quad (2)$$

in which S ($n \times m$) is the matrix of the scores, i.e. the coordinates of the samples in the principal components, and W ($m \times m$) is the loadings matrix. Since the first principal components retain most of the variance, several variables can be summarized by a few components and a graphical representation of the first two or three PCs enables the visualization of most relevant information contained in the data [93].

The principal components are not correlated with each other and altogether explain the total variance of the data. The transformation matrix W whose elements are the loadings w_{ij} and the vector λ, whose components correspond to the recovered variance λ_i in each i^{th} principal component, can be obtained via a singular value decomposition

$$C_xW = \lambda W \quad (3)$$

where C_x corresponds to the variance/covariance matrix of the original data. Also, $\sum_i^m \lambda_i$ gives the total variance of the data. Frequently, C_x is replaced by the correlation matrix, in a normalized approach. In this case, $\sum_i^m \lambda_i = m$.

A fundamental step to determine the number of significant PCs is the extraction of eigenvalues and eigenvectors.

The selection of the most relevant first *p* principal components can be done using different criteria [93]. The most common one is the Pearson criterion, which can be used in

conjunction with both the variance/covariance matrix and the correlation matrix. The value *p* is selected as the minimum integer that warrants

$$\sum_{i=1}^{p} \lambda_i \Big/ \sum_{i=1}^{m} \lambda_i \geq 0.8 \quad (4)$$

If the correlation matrix is used, the most common criteria correspond to retain the *p* components for which $\lambda \geq 1$. The $\lambda \geq 1$ rule, also known as Kaiser criterion, is selected when the correlation matrix is used. This rule takes into account all components with eigenvalues greater than one.

A more robust criteria is the scree plot. This representation displays the eigenvalues as a function of the corresponding principal components, as shown in Figure 8. According to this criteria the relevant principal components are those that markedly standout from the remaining in terms of variance.

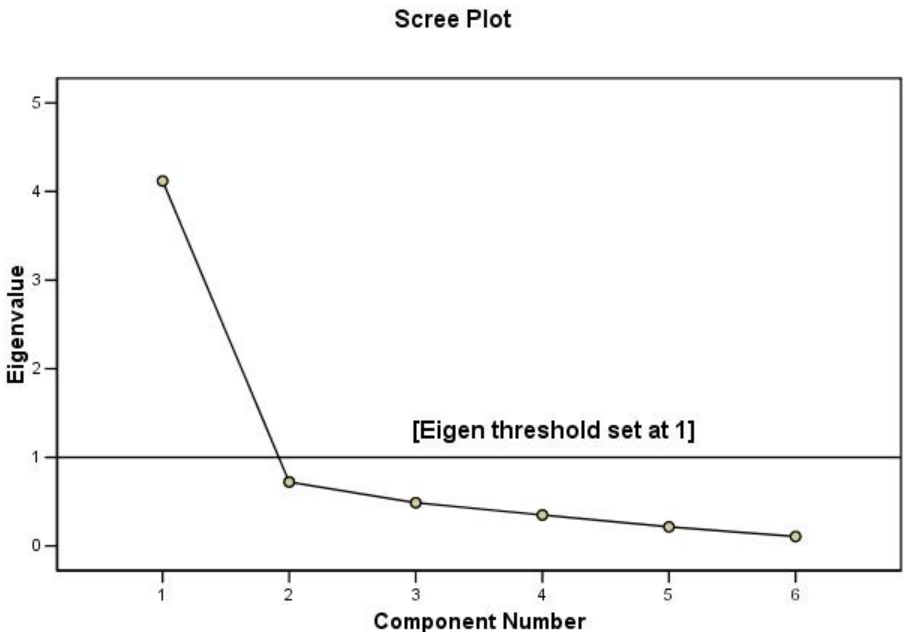

Figure 8 - Scree Plot example suggesting 1 PC solution.

In the present work, PCA is used to rationalize the information related to the identification of SCZ biomarkers. Specifically, the main focus is the study and identification of metabolites that may be altered by the effect of three different drugs, haloperidol, citalopram and clozapine. The data were collected from pre-frontal cortex samples of mice, resorting to

HPLC-MS/MS analysis. Efforts were made to establish correspondence between the most relevant metabolites and molecules available databases.

3. Materials and Methods

3.1. Drug administration and extraction of pre-frontal cortex in mice

Young black male C57BL/6J mice were purchased from Charles River, Laboratories International, Inc. (Spain) and kindly prepared in Dra. Graça Baltazar's lab with the help of Sandra Rocha (University of Beira Interior, Covilhã). Mice were divided into four groups, of 6 animals each. Each animal weighed around 20-25 g with access to food and water *ad libitum*. Animals were chronically treated for four weeks. The animals were injected, via intraperitoneal with clozapine, citalopram or haloperidol at a dose of 20 mg/Kg, 10 mg/Kg and 1 mg/Kg, respectively. All drugs were dissolved by diluting on a stock solution of 0.13% HCl 5M with 0.9% NaCl. An additional group was treated only with vehicle (saline solution) being considered the control group. The age of animals at the beginning of the habituation phase was 10 weeks and mice were injected daily for 30 days. After drug administration, the animals were weighed and anesthetized with a mixture of ketamine (100 mg/Kg) and xylazine (10 mg/Kg) 24 hours after the final injection. Finally, they were sacrificed and their brains removed and dissected bilaterally and the prefrontal cortex was collected. TEAB 0.5 M (triethylammonium bicarbonate buffer) with phosphatases and proteases inhibitors (Roche) was added to each tube. All samples were stored at -80˚C until use. These samples were already used by Susana Costa Saraiva master thesis for performance an analysis of PFC proteome. The metabolites were stored at -80˚C until now for their use in this the project.

3.2. Internal Standard Tests

Internal standard (I.S.) is a necessary step for quantification in HPLC-MS since it allows to see if any alteration occurs during the execution of the procedure/sample analysis. A mixture of 56 compounds (available in the laboratory) was prepared in a solution of 2% acetonitrile (ACN) and 1% formic acid (FA) to a final concentration of 10 µM. The mixture was then subject to C18 tips process (OMIX TIP C18 100 µl, Varian), which includes: 1) conditioning with 50% ACN solution; 2) equilibrate with 2% ACN + 1% FA solution; 3) sample; 4) rinse with 2% ACN + 1% FA solution; 5) elution with 70% ACN + 0.1% FA solution; 6) 100% ACN + 0.1% FA solution;

The two first steps were performed to prepare the tip for the sample. After that, the sample was added to C18 tip passed 5x, the next step, the rinse was performed by passing 1x (100 µl) 2% ACN + 1% FA solution and the elution was made by passing 4x (100 µl) of 70% ACN + 0.1% FA solution. This elution step is the most important one because it will show which compounds have affinity with C18 and were not wasted in the first two steps. This requirement is essential for being considered as internal standard. An extra step was performed by passing 4x (100 µl) 100% ACN + 0.1% FA solution for evaluation of the efficiency of SPE. The first two fractions were discarded, and the collected fractions were denoted as F1 (step 3), F2 (Step 4), F3 (step 5) and F4 (step 6). Samples were then evaporated and ressuspended in a 2% ACN + 0.1% FA solution, followed by sonication at 20% of amplitude for 2 minutes (1 second on 1 second off cycle) and transferred to vials for LC-MS analysis.

3.3. Metabolite Extraction

Metabolite extraction was performed by the membrane protein enrichment protocol. The tissue was removed from the storage at -80˚C and thawed at room temperature. Samples were transferred to centrifuge tubes and 1 mL of Tris 0.05 M (tris(hydroxymethyl)aminomethane) pH=7,4 with phosphatases and proteases inhibitors added. Subsequently, tissues were homogenized with ultrasonication (Vibra Cell 130 watts, Sonics) with the 2 mm probe for 30 seconds at 40% amplitude with 1 second cycles and for 15 seconds at 50% with 1 second cycles and centrifuged (Centrifuge 5417R, Eppendorf) at 5,000×g for 5 minutes at 4˚C. Supernatants were stored in a new microcentrifuge tube and 500 µL of Tris 0.05 M with phosphatases and proteases inhibitors were added to the pellets. Supernatants were transferred to ultracentrifuge tubes and the pellets were kept at -80˚C. In the next step, an ultracentrifugation (OptimalTM L-100 XP, Beckman Coulter) at 144,000xg for 1h at 4˚C was performed and the supernatant (soluble fraction, SF) was taken to a microcentrifuge tube. For the pellet (membrane fraction, MF) 500 µL of TEAB 0.5 M plus 500 µL of water was added (because TEAB has 1M of concentration) and the pellet was dissolved using the sonicator with the 2 mm probe for 15 seconds at 40%, once, after this the pellet was unstacked from the ultracentrifuge tube and homogenized again, 30 seconds at 40% amplitude with 1 second cycles and 30 seconds at 50% amplitude with 1 second cycles, this step was repeated until total dissolution. The MF and SF were subjected to protein

precipitation with MeOH (Methanol) and 1:4 (sample: methanol) was added in each. Fractions were vortexed (IKA Vortex 4basic) between 10 to 15s and stored for 20min at -80˚C. To obtain proteins and metabolites, it was performed a new centrifugation (Centrifuge 5810 R, Eppendorf) at 3.220×g at 4˚C for 20min. Supernatants (metabolites of MF and SF) were removed and transferred to microcentrifuge tubes, and the pellets (proteins) were stored at -80˚C. Supernatants were evaporated in the speedvac (Eppendorf, Concentrate plus) during 3h and 25min at 60˚C to perform a successful evaporation of all metabolites. When this was complete, the samples were resuspended and sonicated (Bioblock Scientific vibracell 75041) during 2min at 20% with 1 second cycles and subjected to vortex and centrifuge (Minispin plus, Eppendorf) with 14.100xg for 5 min. The pellets were stored and the supernatants were subjected to a solid-phase extraction protocol using C18 tips (OMIX TIP C18 of 100 µl from Varian) according to the process described above. Samples were evaporated during 1h and ressuspended with 30 µl of 2% ACN + 0.1% FA solution in each tube. After vortex, sonication and spin all the samples were transferred to vials for LC-ESI-QTOF-MS (Sciex) analysis. To optimize the protocol it was performed a set of different options with the internal standards mixtures. In Table 1 are summarized the four options that were considered for the addition of the internal standard.

Table 1 – Options considered to perform the I.S. spike in membrane protein enrichment protocol.

	1st	2st	3st	4st
Matrix	✔	✔	✔	✔
I.S	●			
Ultracentrifugation	✔	✔	✔	✔
I.S.		●		
Protein Precipitation	✔	✔	✔	✔
I.S.			●	
C18 tips	✔	✔	✔	✔
I.S.				●
LC-MS analysis	✔	✔	✔	✔

3.4. Preparation of pre-frontal cortex samples

The main focus of this project was to perform a metabolomic analysis with PFC samples. As previously mentioned, the metabolites of PFC obtained by Membrane protein enrichment protocol was already set, by Susana Saraiva masters' thesis. The metabolites PFC samples are divided in two categories: membrane fraction (MF) – fraction that was analyzed in the present

project - and soluble fraction (SF). Each category has 4 groups of samples, Control (CT), Haloperidol (HA), Citalopram (CI) and Clozapine (CL) and each group had 4 replicates.

Table 2 – Description of the SF and MF groups.

Soluble Fraction				**Membrane fraction**			
CT1	HA1	CI1	CL1	**CT1**	**HA1**	**CI1**	**CL1**
CT2	HA2	CI2	CL2	**CT2**	**HA2**	**CI2**	**CL2**
CT3	HA3	CI3	CL3	**CT3**	**HA3**	**CI3**	**CL3**
CT4	HA4	CI4	CL4	**CT4**	**HA4**	**CI4**	**CL4**

Samples were sonicated with 2min at 20% with 1 second cycles and ultra-sonicated (VWR Ultrasonic cleaner, USC-THP) with 3 min at 7 powers at 25˚C. The next step was to ressuspended the samples in 100 µl of 2% ACN + 1% FA in each. After that, they were divided and 50 µl were transferred to a new microcentrifuge tube and stored at -80˚C. Samples were identified according to the labels in table 2.

The samples were subject to vortex, sonication and centrifugation and then subject to C18 tips. The C18 protocol was performed as described above and for that three solutions were necessary: a) 50% ACN, b) 2% ACN + 1% FA and c) 70% ACN + 0.1% FA and make the five first steps that were required in this process. Subsequently, the samples were evaporated during 1h and ressuspended with 30 µl of 2% ACN + 0.1% FA solution and I.S. solution (Sulfamethazine-D4). Samples were then subject to vortex, sonicator and spin, and finally transferred to vials for HPLC-MS/MS analysis.

3.5. HPLC-MS/MS analysis

Samples were analyzed on a NanoLC Ultra 2D separation system (Eksigent) coupled to an electrospray ionization source (DuoSpray™ Source, Sciex) operated in positive mode, and a Triple TOF™ 5600 System mass spectrometer (Sciex). Metabolites were separated into a Halo C18 column (0.3 x 150 mm, 2.7 µm, 90 Å, Eksigent) at 5 µL/min, with an acetonitrile gradient (2% to 61% ACN, for 42 minutes) in 0.1% FA. Using the same chromatographic conditions, mass spectrometer was programmed for two different forms of data acquisition: information dependent acquisition (IDA) and information independent acquisition SWATH analysis. For IDA, a full mass spectra (30-1250 m/z) was acquired, followed by 11 MS/MS of ions with +1 to +4 charges and one MS/MS was performed before adding those ions to the exclusion list for

15 seconds. For SWATH experiments, the mass spectrometer was operated in a looped product ion mode. The instrument was specifically tuned to allow a quadrupole resolution of 50 m/z mass selection. Using an isolation width of 51 m/z (containing 1 m/z for the window overlap), a set of 17 overlapping windows was constructed covering the precursor mass range of 50–900 m/z. A 250 milliseconds survey scan (50-2000 m/z) was acquired at the beginning of each cycle and SWATH MS/MS spectra were collected from 50–1300 m/z for 120 milliseconds resulting in a cycle time of 2.34 seconds from the precursors ranging from 50 to 900 m/z. The collision energy for each window was determined according to the calculation for a charge 1+ ion centered upon the window with a collision energy spread of 15. The mass spectrometer was operated by Analyst® TF 1.6, Sciex.

3.6. HPLC-MS data processing

The characterization of biological samples generates a large amount of complex data which is difficult to describe and rationalize. Chemometrics provides useful tools to assist in the characterization of these systems. In this project different softwares were used: i) MarkerView™ software (version 1.2.1.1, Sciex) for peak detection, peak alignment and PCA analysis; ii) MultiQuant™ 2.1.1 software (version 2.1.1742.0, Sciex, 2012) for peak integration; iii) Predictive analytics software and solutions (SPSS) (version 18, PASW Statistic 18, Release 18.0.0) for Mann-Whitney significance test. An outlook of the data processing procedure is shown in Figure 9.

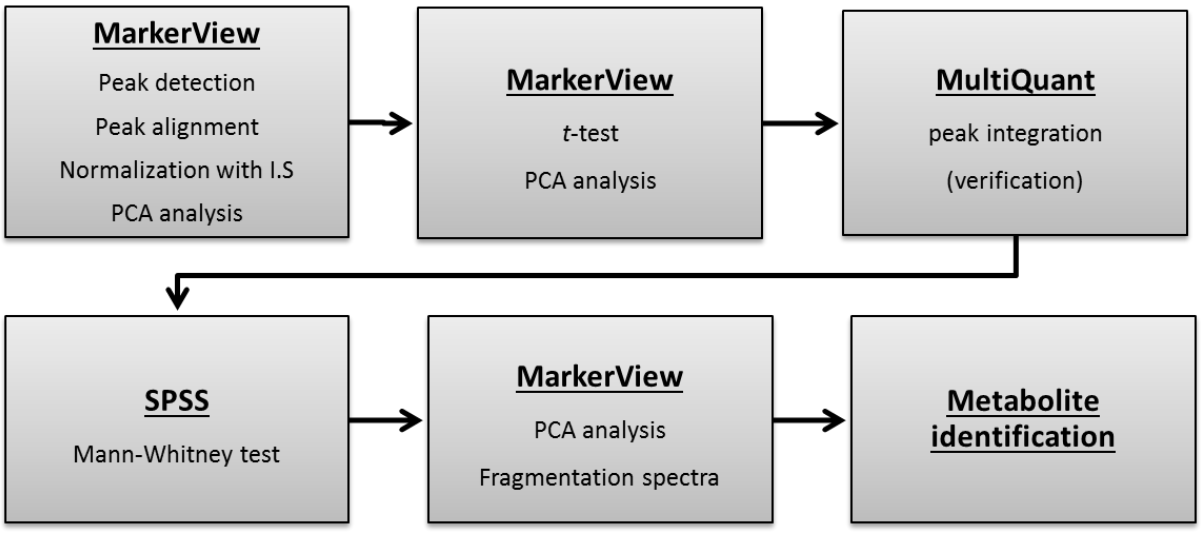

Figure 9 – Data processing workflow with the identification of the software used in each step.

In simple terms, the data files acquired by MS were directly imported to MarkerView™. For peak detection the following criteria were applied: a) retention times above 5 minutes; b) minimum spectral peak width of 0.02 Da (peaks narrower than this value are presumed to be noise); c) minimum retention time peak width of 4 scans; d) maximum retention time peak width of 10 scans (peaks wider than this value are assumed to be background ions). Data alignment was applied to compensate smaller variations in mass and retention times, to ensure that similar compounds in different samples are accurately compared. Additionally, retention time tolerance was set to 0.5 minutes, and mass tolerance set to 30.0 ppm. According to these criteria two or more peaks having the same m/z value were considered the same if the retention time did not exceeded 0.5 minutes. The same occurs when two or more peaks differ at the most in 30.0 ppm. Peak detection, alignment and normalization were carried out using Pareto[2] scaling approach in MarkerView. Preliminary PCA´s were performed directly on the global set of normalized integrated peak areas. Subsequent PCA´s were performed on filtered data, obtained through significance tests including t-test (performed in an "noise reduction approach") and Mann-Whitney (using SPSS

[2] According to MarkerView software, in Pareto each peak is subtracted by the average and divided by the *square root* of the standard deviation. This is a good initial choice for MS data since it prevents intense peaks from completely dominating the PCA analysis, but also allows peaks with good signal/noise to have more importance.

software). Peak areas were calculated by MultiQuant, using normalized data with internal standard[3].

[3] Internal standard used was Sulfamethazine-D4, a regular compound utilized in the laboratory.

4. Results and Discussion

4.1. Internal standards: preliminary tests

Internal standard tests were performed testing a set of 56 different compounds (see Table 7 in supplementary material), which should satisfy the criteria of being similar to the analyte, have a similar retention time, be stable and not interfere with the sample matrix.

The prepared samples were subjected to the same procedure of that of the pre-frontal cortex ones. Thus, C18 tips were used in order to remove interferences before the HPLC-MS/MS analysis. Three different fractions were obtained through this process, the loading sample, the washed and the eluted fractions, being the latter the most interesting one.

From the 56 compounds, only 29 revealed a high intensity peak in the elution step, thus being suitable for being considered as I.S, see Figure 10. In the left side of the plot are the compounds that were lost in loading sample fraction, while in the right side, starting in compound #56, are those that were retained in the loading sample and that were detected in the eluted fraction. These fulfill the profile for being selected as I.S.

Another requirement that the I.S. must fulfill is the adequate retention time, and since this parameter varied significantly along the 29 compounds selected, the subsequent step (Metabolite extraction) was conducted using mixtures of all of them.

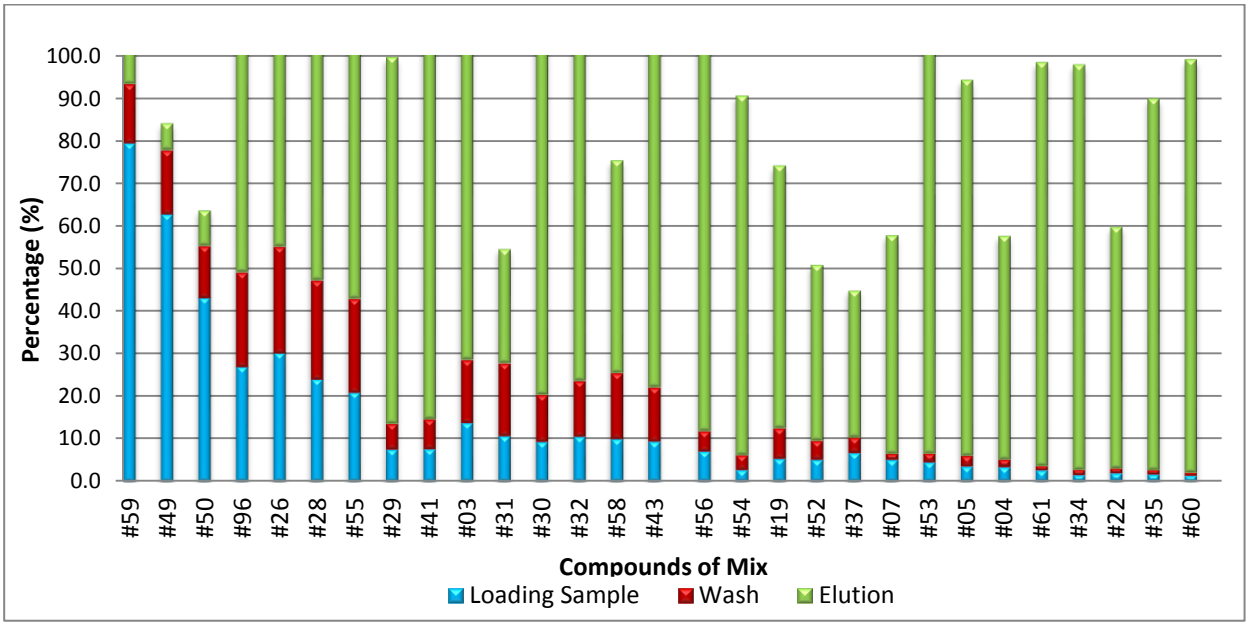

Figure 10 – Fractions obtained in I.S. tests resorting to C18 tips.

4.1.2. Metabolite Extraction

Membrane enrichment protocol allows separating proteins from metabolites, since the soluble part of the sample is separated from the membrane one, which is the focus of this project. During this protocol spikes of I.S (using the previously selected 29 compounds) in the matrix were made. Four different steps were considered to evaluate the recovery of each compound, before ultracentrifugation, before protein precipitation, before C18 tips and before HPLC-MS/MS analysis.

Figure 11 shows the results obtained in the different steps of this protocol. It can be seen that compounds #04, #19, #26, #35, #50, #52 and #59 presenting the higher intensities in all steps of the protocol are the most interesting ones to be used as I.S. These compounds represent also those with the smaller losses between the steps. It should be noted that additional tests would be necessary to complete this preliminary search for the ideal I.S. However, in order to comply with the main objective of the project on time, the work had to proceed with an I.S., Sulfamethazine-D4, commonly used in the laboratory for this type of application.

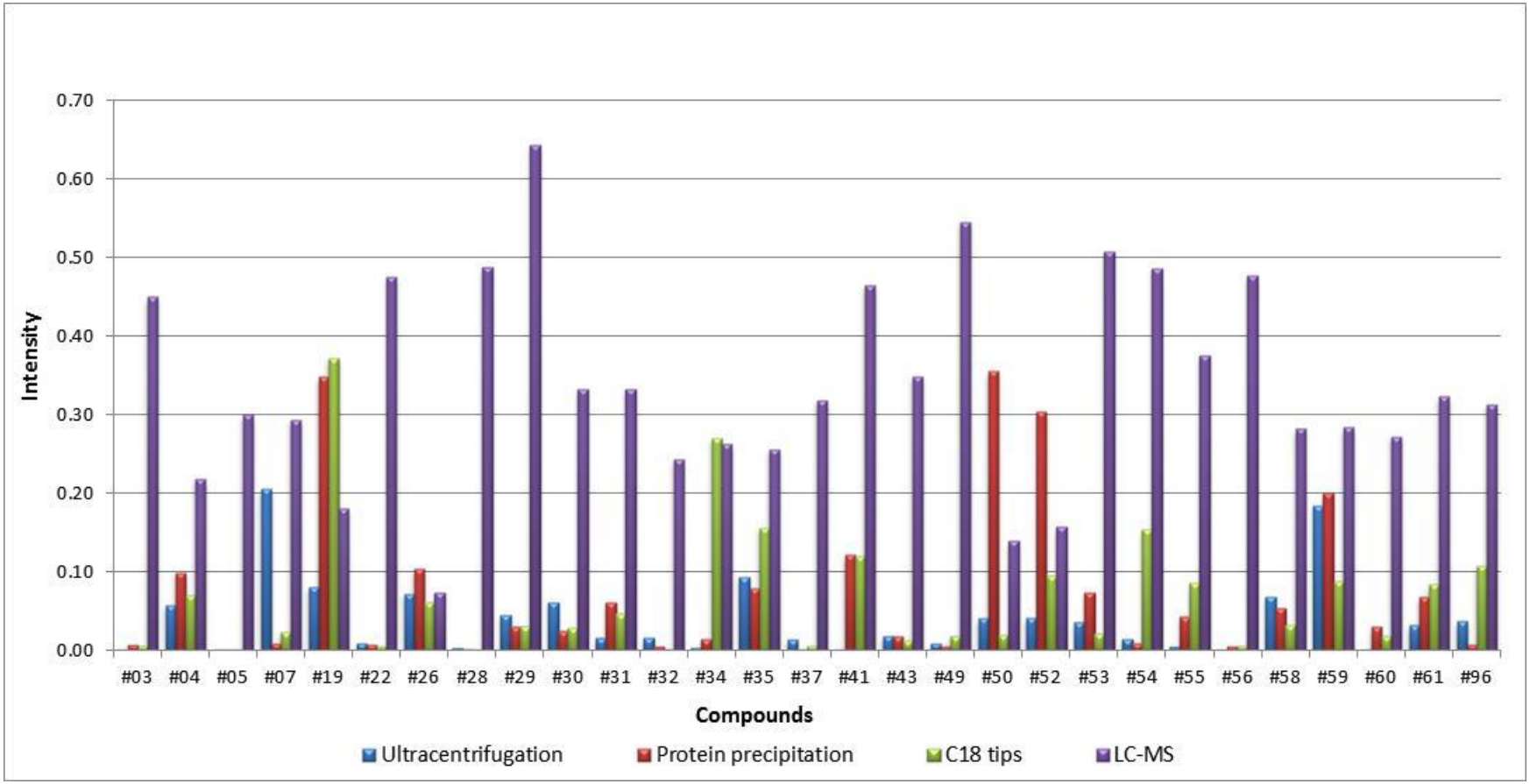

Figure 11 – Intensity detected by HPLC-MS, in each step of the membrane enrichment protocol. The blue bar comprises the step before ultracentrifugation; the red bar comprises the step before protein precipitation; the green bar, the step before C18 tips and the purple bar represents the step before the HPLC-MS analysis.

4.2. Metabolites in pre-frontal cortex samples

In order to investigate possible alterations in metabolites promoted by the effect of the three different drugs studied, each sample was compared with control samples obtained from mice free of any drug.

Control vs Citalopram

As referred before, after the normalization of the peak areas with that of the I.S, Sulfamethazine-D4 (283.11 m/z and 17.44 retention time) a PCA was performed. Figure 12 displays a composed view of the samples in the new orthonormal principal component system. This representation, in two dimensions, shows a significant overlap between citalopram(CI) and control (CT) samples. It is seen that the first two principal components are able to recover ca. 44%. and 21%, respectively, of the data variability.

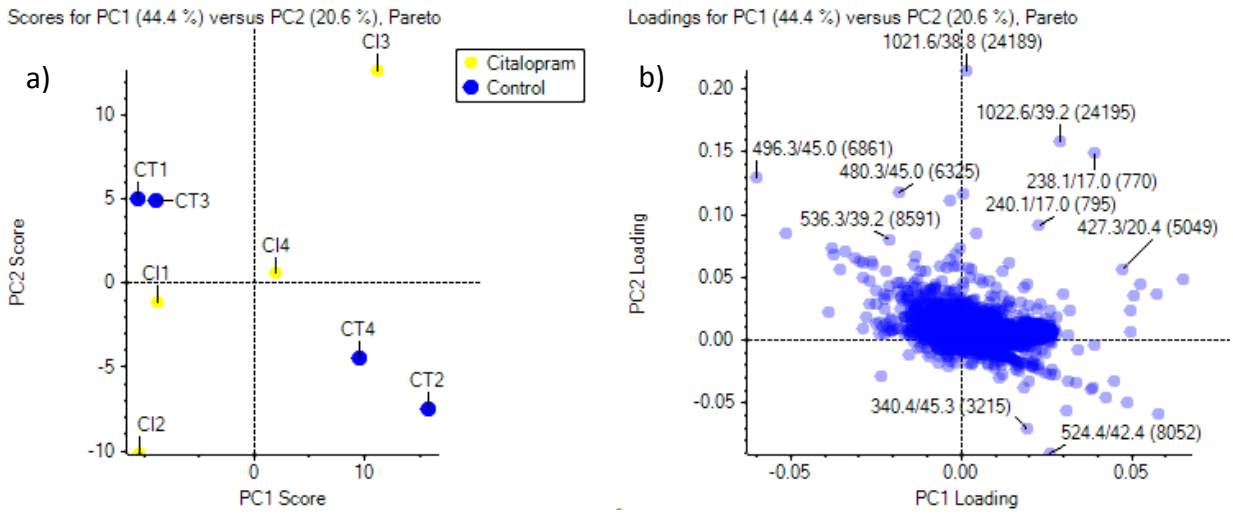

Figure 12 - Representation of samples on the main two principal components. Panel (a) corresponds to the scatter plot of correlation scores with ca. 65% of information recovery. The global set contains 4 samples for each type, control (blue) and citalopram (yellow). Panel (b) depicts the respective loadings.

In order to improve the distinction between both groups a new PCA was performed on 217 peaks selected with a t-test (p-value < 0.05). The results presented in Figure 13 show that the first component contains, in fact, the most relevant information for discrimination. The two first principal components are able to recover ca. 77% of the data variability. The inspection of panel (a) clearly shows that a graphical representation based on these two

components is clearly meaningful and the discrimination between citalopram samples and control ones lies essentially on PC1 which retain ca. 66% of the information. The relevance of each peak, in the first two components (PC1 and PC2) is presented in panel (b) of Figure 13. The criterion for selecting a significant loading is based on the comparison to the average value in each component, i.e., the loading is simply considered significant if above the average value defined by $1/\sqrt{m}$, and not significant otherwise, where m stands for the number of peaks. A full description of this approach can be found in reference [93].

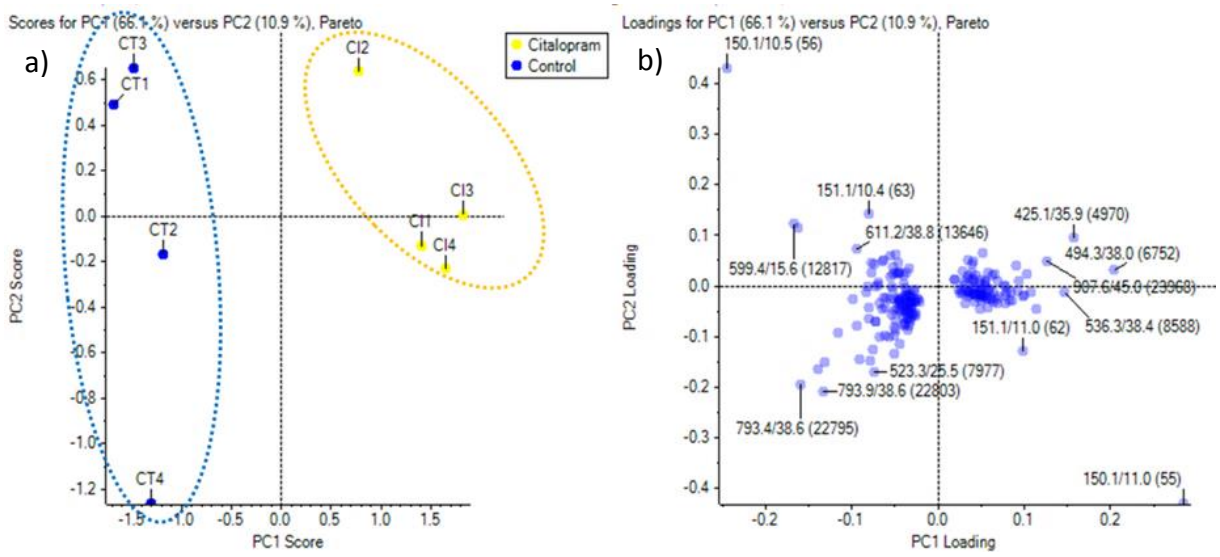

Figure 13- PCA results obtained for Control and Citalopram groups. Panel (a) shows the relative positioning of each sample in PC1 vs. PC2 plane corresponding to ca. 77% of information recovery. Panel (b) displays the respective loadings. Complete list of significant loadings in Table S5 in supplementary material.

Inspecting the scores and the loading values of each peak it is concluded that the first component retains mainly information over peaks (55), (56), (6752), (12817), (2865), (22795), (4970), (8588), (24237), (22803), (24235) and (23968) (see the complete list of significant loadings in Table S5 in supplementary material). It is seen that samples spread essentially along PC1, but PC2 also contributes for the intra-group discrimination. Most of the samples are placed in the PC1 vs PC2 plane in such a manner that the closest samples (for example Cl1, Cl3 and Cl4) are those with similar peak profile. To further improve the system resolution a Mann-Whitney test was performed and 75 peaks with p-value <0.05 were obtained (see Table S2 in supplementary material, with the ratio values). Figure 14 shows that using a significantly lower number of variables the discrimination between drug sample and control ones is preserved. It is possible to identify the peaks that characterize each group. In this case

discrimination is based mainly on peaks (22795), (22803), (18535), (24237), (4970), (24235), (8714), (2865), (3512) and (4686).

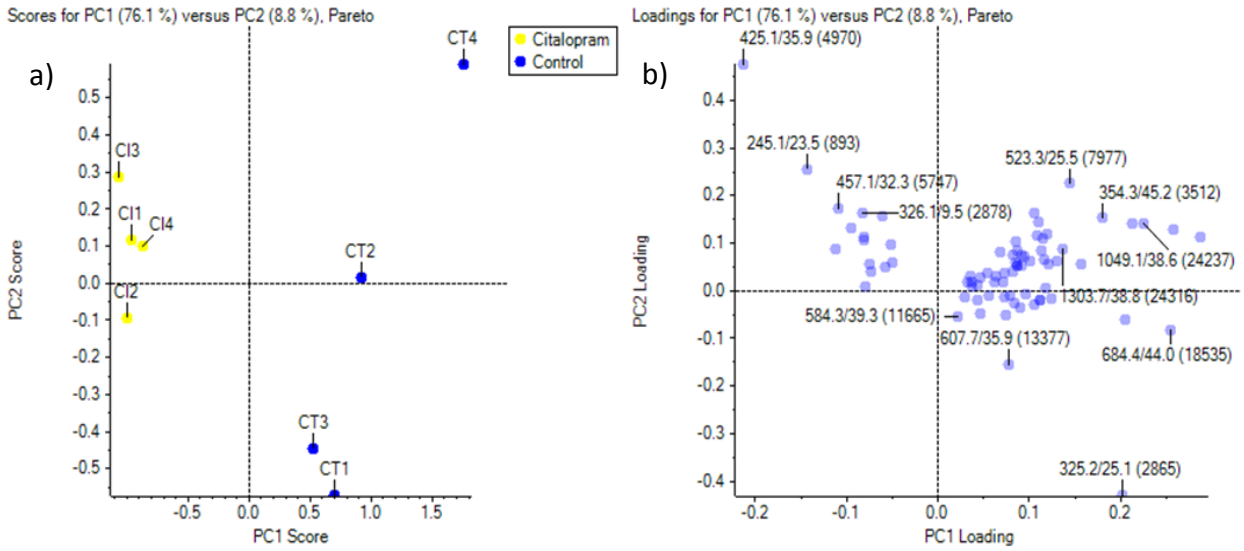

Figure 14 - Plot of the two principal components in a PCA between Control and Citalopram groups. The plot it was made by all interesting ions found after peak integration and a Mann Whitney test, without CV calculation. Complete list of loadings in Table S6 in supplementary material.

At a later stage from the 75 peaks previously selected, only those with a coefficient of variation lower than 30% and simultaneously present in the control and drug samples were considered (see Table 3). The former criterion intends to keep in each group the peaks that preserve the group homogeneity while the latter intends to remove potential errors arisen from the alignment procedures. Table 3 contains the values of the mean and standard deviation for each group together with the respective coefficient of variation. In the last column, and only to assist data interpretation, is qualitatively indicated which group has the greater variability, green arrow when citalopram samples shows higher variability and red when otherwise.

Table 3 – Values of mean, standard deviation and CV for citalopram and control samples. Last column represents for each peak which the group that presents the higher variability. Green arrow when citalopram samples show higher variability and red when otherwise.

Peak name $_{((m/z)/RT)}$	μ (CT)	σ (CT)	μ (CI)	σ (CI)	CV (CT) $_{(\%)}$	CV (CI) $_{(\%)}$	
321.1/17.4	0.024	0.006	0.043	0.012	25.9	28.1	↑
1304.2/38.9	0.058	0.017	0.023	0.001	28.7	4.2	↓
538.8/39.0	0.161	0.026	0.063	0.006	15.9	9.7	↓
800.8/35.8	0.040	0.006	0.019	0.002	14.9	11.5	↓
793.4/38.6	0.288	0.068	0.113	0.017	23.6	15.1	↓
1558.9/38.8	0.032	0.009	0.014	0.002	27.8	12.1	↓
384.0/18.9	0.053	0.014	0.025	0.005	26.6	18.7	↓
792.9/38.8	0.060	0.010	0.031	0.005	16.2	15.7	↓
252.1/18.7	0.023	0.005	0.012	0.002	22.7	17.7	↓
1891.1/38.8	0.009	0.002	0.004	0.001	22.8	17.1	↓
508.0/38.0	0.053	0.015	0.021	0.003	28.9	12.1	↓
288.3/35.4	0.056	0.017	0.027	0.008	29.8	29.3	↓
1048.1/38.9	0.034	0.009	0.019	0.003	25.1	13.6	↓
212.2/27.5	0.060	0.004	0.046	0.011	6.0	22.8	↑
195.1/29.5	0.012	0.001	0.009	0.002	9.4	21.9	↑
622.7/18.9	0.015	0.002	0.007	0.002	12.6	29.7	↑
527.2/34.5	0.011	0.003	0.007	0.001	23.3	14.5	↓
415.1/35.9	0.084	0.009	0.055	0.007	10.4	12.9	↑
792.4/38.9	0.027	0.003	0.010	0.001	11.5	13.6	↑

A new PCA was performed using the new set containing the 19 selected peaks. The results are displayed in Figure 15. It can be observed that a similar pattern is obtained in this simplified system. The first principal component is responsible for the separation between citalopram and control samples. Inspecting the scores and the loading values of each peak it can be seen that peaks (22795), (8714), (24317) and (7262) are those that allow to preserve the discrimination between control and citalopram samples. It should also be noted that the administration of citalopram seems to reduce some of the internal variability which is reflected in the projection of citalopram samples on the first principal component.

In order to support the previous results a parallel analysis was performed using the fragmentation mass spectra of each of the ions with significant loadings for PC1 (see Table S7 in supplementary material). From these, and due to software requirements, it was only possible to obtain the fragmentation mass spectra for those with m/z <1000 (m/z = 538.8, 793.4, 508.0, see Figures S1 to S3 in supplementary material). Finally, the precursors and their fragments were compared with the data available in MyCompoundID and Metlin metabolites databases. No matching could be found for these metabolites.

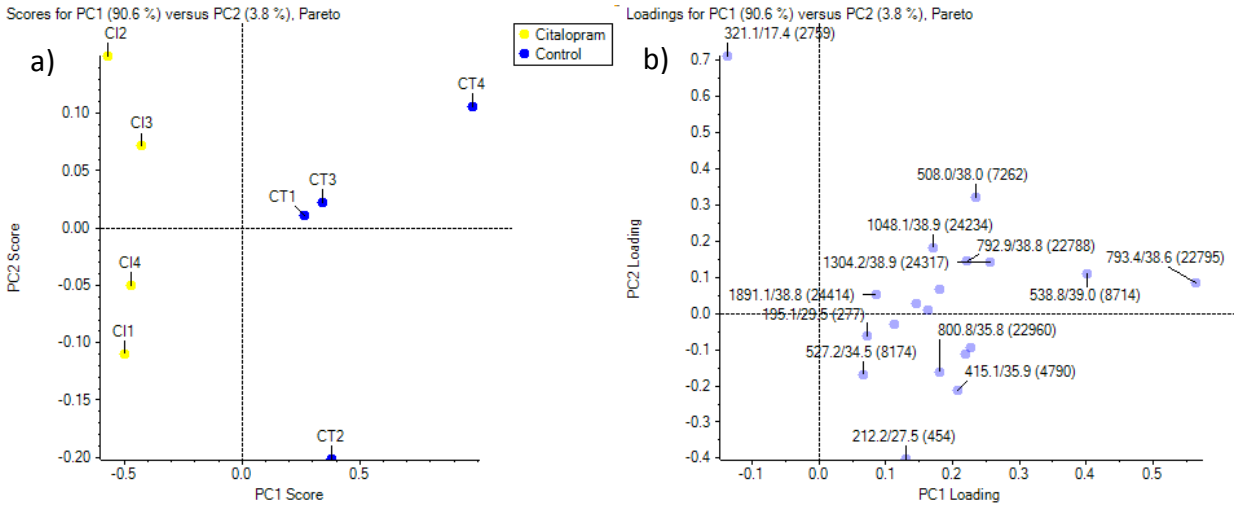

Figure 15- Representation of citalopram and control samples in the new coordinate system (with ca. 95% of information recovery) considering the 19 points with a CV<30%. Complete list of loadings in Table 13 in supplementary material.

The same type of approach has been conducted for haloperidol and clozapine. The same analysis steps previously described were applied in order to progressively reduce the dataset. Briefly, these include the following steps (i) PCA on the global set of normalized peak areas; (ii) PCA on peaks selected by a t-test (*p*-value < 0.05); (iii) Mann-Whitney test followed by PCA of the variables with *p*-value < 0.05; (iv) PCA on the variables with a coefficient of variation <30% and simultaneously present in the control and drug samples.

Finally, the fragmentation mass spectra of the selected ions was studied and the selected precursors and their fragments were compared with the data available in metabolites databases. Table 4 summarizes the size of the dataset studied by PCA in each step for haloperidol and clozapine.

Table 4 - Overview of dataset size in each of the four main steps for which PCA was performed.

	Size of the dataset in each subsequent analyses step			
	Global set (m/z)	t-test	Mann-Whitney test	CV< 30%
Citalopram	21 510	217	75	19
Haloperidol	30 803	147	17	4
Clozapine	20 481	279	57	19

Due to the similarity of the data analysis, and in order to avoid an exhaustive and repetitive description only a brief summary of the main results obtained for haloperidol and clozapine will be made hereafter.

In the case of haloperidol, and as observed for citalopram, no distinction could be found between drug and control samples in the first PCA (see Figure 16). PCA over the set of 147 selected peaks (step ii) revealed a clear separation of the two groups of samples along the first principal component provided mainly by peaks (938), (12197), (483), (944), (3843), (11780), (32401), (5077), (3593) and (7845), (see Figure 17). This distinction along PC1 remains as the dataset size decreases, see Figure 18 (and Table S3 in supplementary material). It is interesting to note that the information retained in the remaining 4 peaks (Table 5 and Figure 19) is sufficient to discriminate haloperidol and control samples. This discrimination is provided by peaks (11154) and (33604). In this case, and as occurred for citalopram, no correspondence could be found between the precursors selected or the corresponding fragments (Figure S4 in supplementary material) and the data available in the surveyed databases. It should be noted that the reduced size sets must be regarded with some caution, since they may be partially produced by accidental discriminative variables.

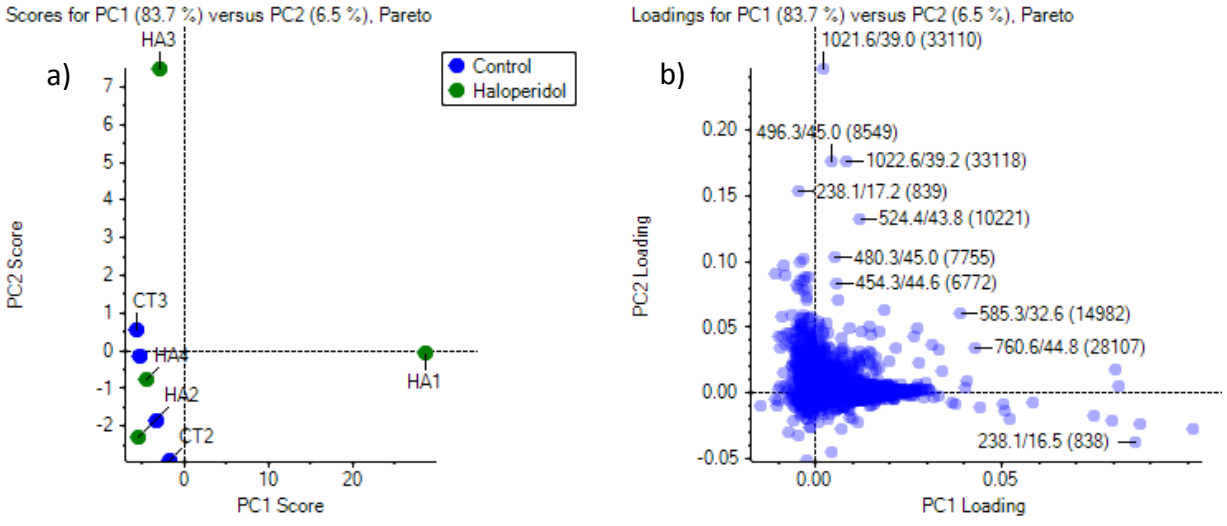

Figure 16 - Representation of Haloperidol and control samples in the new coordinate system. Blue refers to control samples and green to haloperidol.

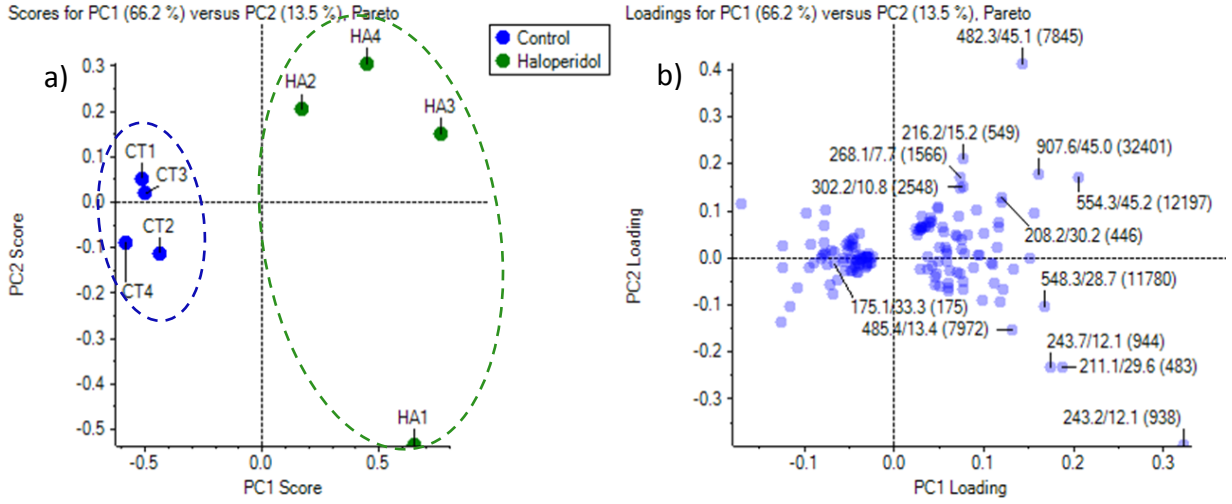

Figure 17 - Representation of haloperidol and control samples in the new coordinate system (with ca. 80% of information recovery). Blue refers to control samples and green to haloperidol ones. Complete list of significant loadings in Table S8 of supplementary material.

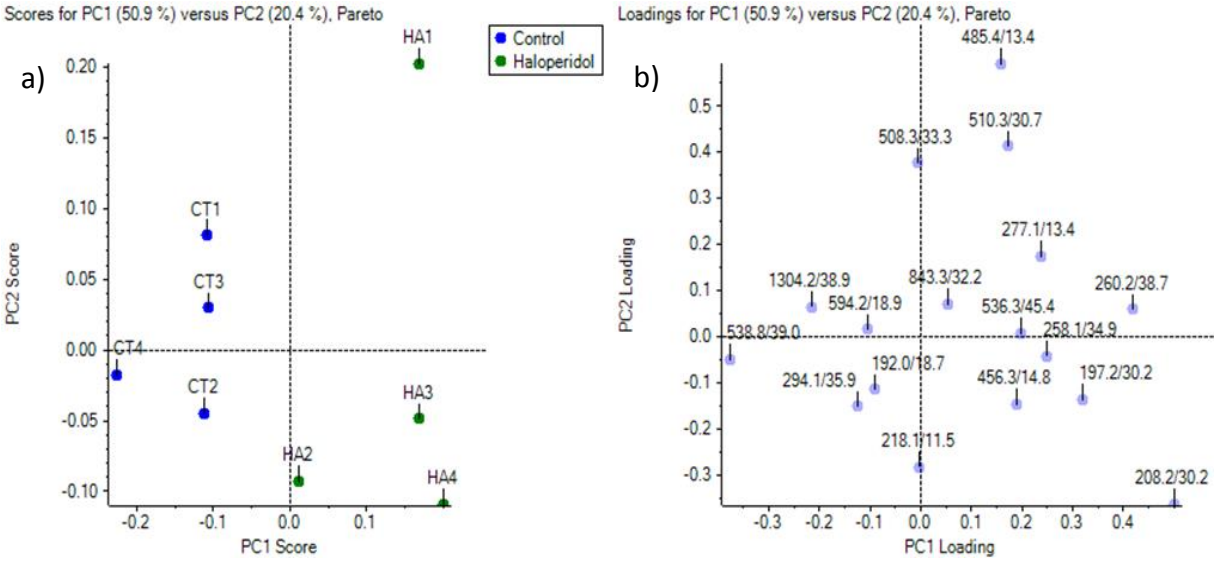

Figure 18 - Representation of haloperidol and control samples in the new coordinate system (with ca. 71% of information recovery). Blue refers to control samples and green to haloperidol ones. PCA was performed on ions selected after peak area integration and Mann Whitney test. Complete list of loadings in Table S9 of supplementary material.

Table 5 - Values of mean, standard deviation and CV for haloperidol and control samples. Last column represents for each peak which is the group that presents the higher variability. Green arrow when haloperidol samples shows higher variability and red when otherwise.

Peak name $_{((m/z)/RT)}$	**μ (CT)**	**σ (CT)**	**μ (HA)**	**σ (HA)**	**CV (CT)** $_{(\%)}$	**CV (HA)** $_{(\%)}$	
538.8/39.0	0.031	0.004	0.022	0.003	11.2	13.0	↑
1304.2/38.9	0.012	0.002	0.009	0.002	19.7	17.9	↓
843.3/32.2	0.002	0.000	0.002	0.001	14.4	27.0	↑
192.0/18.7	0.006	0.001	0.005	0.001	22.7	26.7	↑

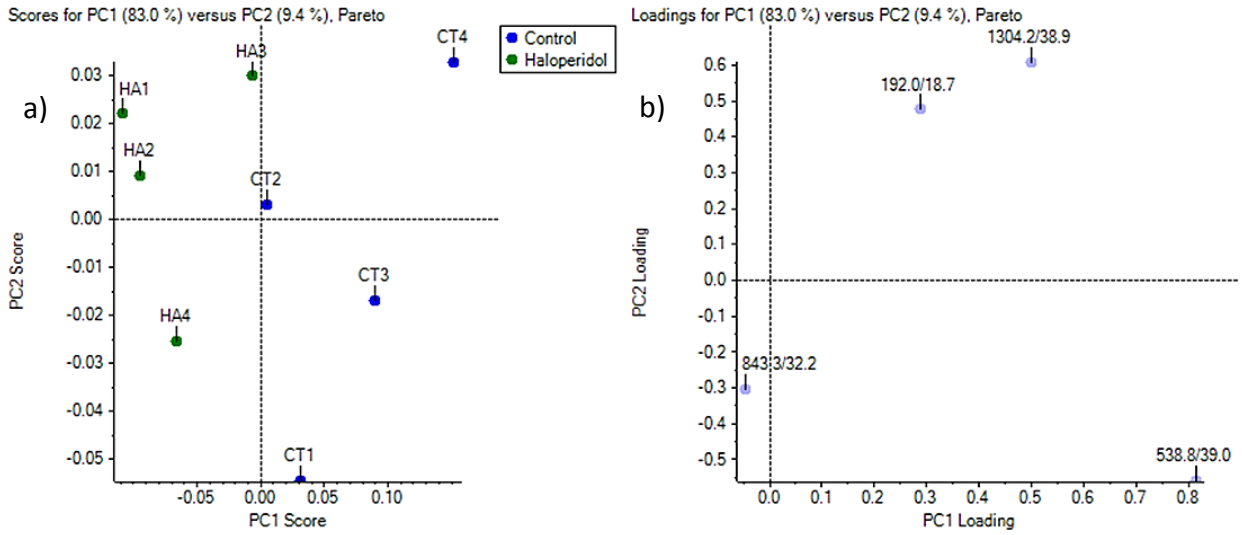

Figure 19- Representation of haloperidol and control samples in the PC1 vs. PC2 plane (panel a) and the contribution of each peak of Table 5 for the new coordinate system is represented in panel b. Complete list of loadings in Table S10 of supplementary material.

For Clozapine, the PCA on the global dataset shows that the discrimination between clozapine and control samples is made essentially along the second component (see Figure 20). The latter contains ca. 17% of the data variability. For example, peaks (7718), (713), (7768), (6592) and (23523) are the ones that contributes more for the discrimination between groups (along PC2). Note that the separation between samples of the same group occurs along PC1 (with 45.3% of information recovery), being more notorious in the control group.

The further simplification of the dataset (see Figures 21 to 23 and Table 6) clearly improves the discrimination between clozapine and control samples. The peaks responsible for the discrimination are (22143), (22151), (23574), (23572), (8385) and (9160). As previously described the study of the fragmentation of the precursors was carried out (Figures S5 to S7 in supplementary material) and the matching between one of the precursors and the data available in the databases was achieved. The molecule with m/z 554.3, matches with an entry of the MyCompoundID database (see the corresponding snapshot in Figure 24). According to this database, this molecule is a lysophospholipid (LPL), it is an endogenous molecule with biofunction in cell signaling, membrane integrity/stability, among others. The information on the database also indicates that this molecule is present in all tissues in the extracellular environment or in the membrane.

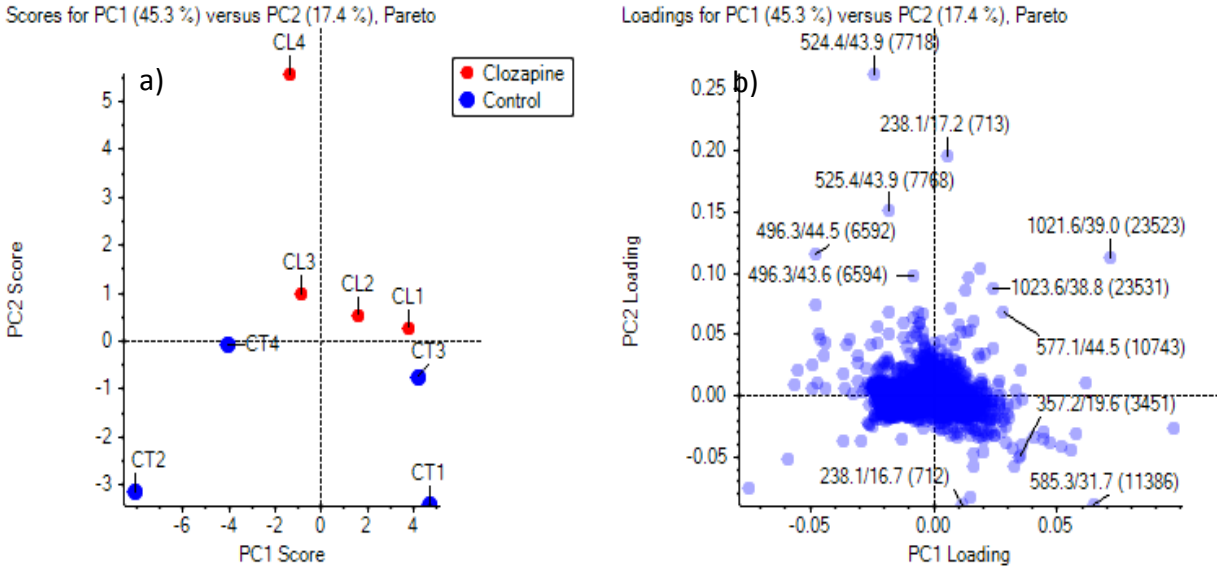

Figure 20 - Representation of Clozapine and control samples in the new coordinate system (with ca. 63% of information recovery). Blue refers to control samples and red to clozapine ones. PCA was performed on the global m/z dataset.

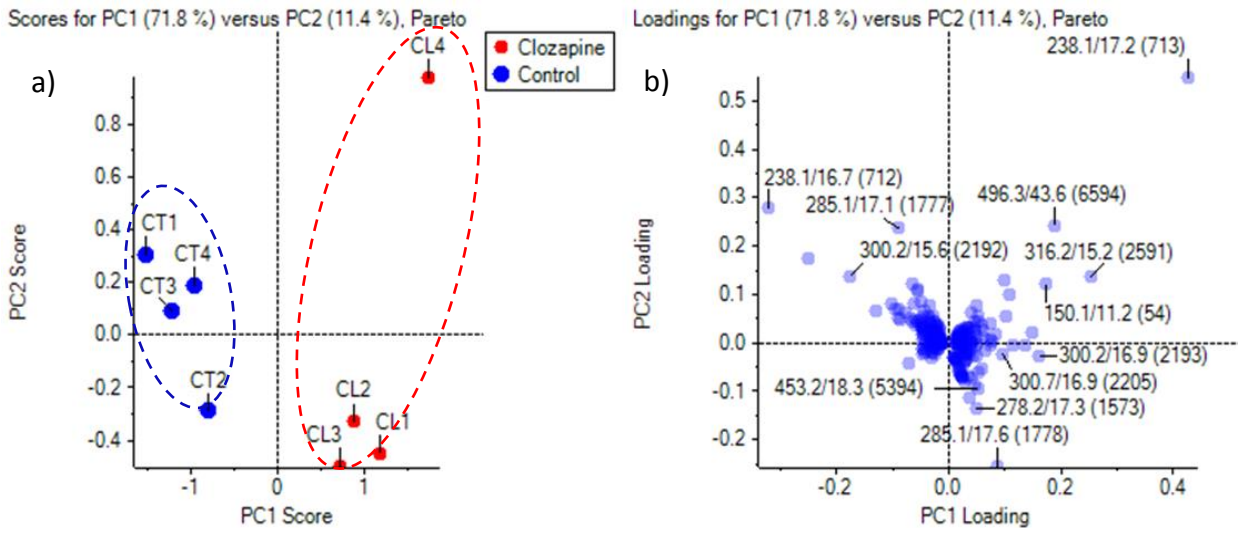

Figure 21 – Representation of Clozapine and control samples in the new coordinate system (with ca. 83% of information recovery). Blue refers to control samples and red to clozapine ones. PCA was performed on the selected 279 peaks. Complete list of significant loadings in Table S11 of supplementary material.

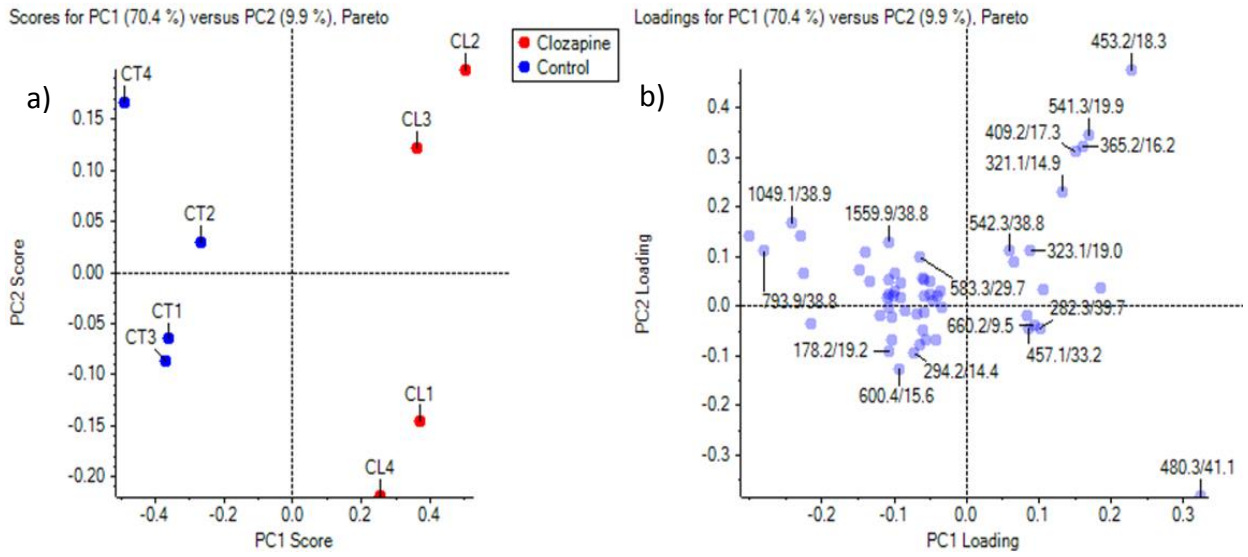

Figure 22 - Representation of Clozapine and control samples in the PC1 vs PC2 plane (with ca. 80% of information recovery). Blue refers to control samples and red to clozapine ones. PCA was performed on the ions selected after peak integration and Mann Whitney test. Complete list of loadings in Table S12 of supplementary material.

Table 6 - Results obtained by calculating the mean and standard deviation of ratios for each sample and after that calculate the CV.

Peak name $_{((m/z)/RT)}$	μ (CT)	μ (CL)	σ (CT)	σ (CL)	CV % (CT)	CV % (CL)	
513.3/34.5	0.012	0.007	0.002	0.001	17.5	13.7	↓
487.1/36.1	0.008	0.004	0.002	0.001	24.7	16.4	↓
1303.7/38.8	0.012	0.005	0.003	0.001	23.2	14.8	↓
1814.0/39.0	0.002	0.001	0.000	0.000	12.3	16.1	↑
1304.2/38.9	0.012	0.005	0.002	0.001	19.7	13.8	↓
538.8/39.0	0.032	0.016	0.004	0.002	11.3	13.5	↑
1049.1/38.9	0.034	0.015	0.007	0.002	20.8	13.4	↓
1558.9/38.7	0.005	0.003	0.001	0.000	20.8	9.5	↓
554.3/45.2	0.024	0.039	0.005	0.009	22.6	22.4	↓
793.4/38.8	0.057	0.029	0.009	0.004	15.5	14.7	↓
792.4/38.9	0.005	0.002	0.001	0.001	11.0	22.5	↑
473.1/33.3	0.004	0.002	0.001	0.001	22.5	29.3	↑
793.9/38.8	0.047	0.023	0.008	0.003	17.3	14.5	↓
1559.9/38.8	0.012	0.008	0.003	0.001	24.3	8.1	↓
1048.6/38.9	0.032	0.015	0.007	0.003	20.9	17.1	↓
1559.4/39.0	0.007	0.004	0.002	0.000	25.4	11.3	↓
627.3/22.9	0.004	0.003	0.000	0.001	8.9	21.1	↑
493.2/39.1	0.004	0.003	0.001	0.000	30.0	14.5	↓
486.1/36.1	0.005	0.004	0.000	0.000	8.4	12.2	↑

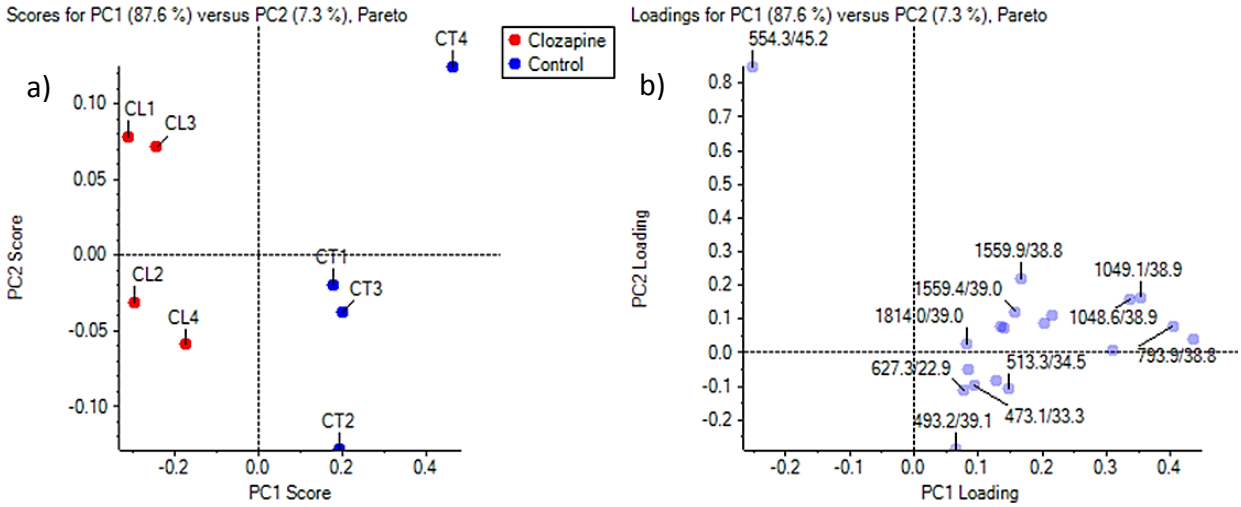

Figure 23- Representation of Clozapine and control samples in the PC1 vs PC2 plane (with ca. 95% of information recovery). Blue refers to control samples and red to clozapine ones. PCA was performed on the ions selected by CV<30%. Complete list of loadings in Table S13 of supplementary material.

HMDB ID	Common Name	Mass (Da)	Formula
HMDB11499	LysoPE(0:0/24:6(6Z,9Z,12Z,15Z,18Z,21Z))	553.316842	$C_{29}H_{48}NO_7P$

Figure 24 – Correspondence between the selected ion and the identification in the MyCompoundID database.

5. Conclusions and future perspectives

This project was intended to identify molecules that might be potential biomarkers of schizophrenia, in samples of prefrontal cortex of mice. Specifically, the main focus was the study and identification of metabolites that might be altered by the effect of three different drugs, haloperidol, citalopram and clozapine. The data were collected resorting to HPLC-MS analysis.

Relatively simple and well known chemometrics techniques provided the tools for an in depth scrutiny of the distribution of m/z peaks. PCA allowed the direct visualization of data structure and also provided the relative positioning of the samples. The most discriminating peaks were identified. A fundamental question concerns the normalization of the data, which in this case was achieved by the use an appropriate internal standard. Another important issue is the reduced size datasets, which must be analyzed and interpreted carefully mostly because they may contain artificially discriminative variables.

In summary, the use of standard multivariate analysis techniques facilitated interpretation, allowed graphical visualization of the potential biomarkers profile and can be used in an almost automated sequence in this kind of studies.

It must however be stressed that, although it was possible to identify a potential biomarker, research is underway to identify further molecules. Future work will also include the application of other techniques such as hierarchical cluster analysis and linear discriminant analysis and a detailed inspection of the behavior of the possible biomarkers, including a check of the respective variation, for example the effect of dosage on biomarker response. Finally, it will be also interesting to include in the analysis a positive control.

6. References

1. Bob, P. and G.A. Mashour, *Schizophrenia, dissociation, and consciousness.* Conscious Cogn, 2011. **20**(4): p. 1042-9.
2. Tandon, R., H.A. Nasrallah, and M.S. Keshavan, *Schizophrenia, "just the facts" 4. Clinical features and conceptualization.* Schizophr Res, 2009. **110**(1-3): p. 1-23.
3. Faludi, G., P. Dome, and J. Lazary, *Origins and perspectives of schizophrenia research.* Neuropsychopharmacol Hung, 2011. **13**(4): p. 185-92.
4. Tomasik, J., et al., *Neuroimmune biomarkers in schizophrenia.* Schizophr Res, 2014.
5. Kapur, J.v.O.a.S., *Schizophrenia*, in *The Lancet* 2009. p. 635-645.
6. Mikkelsen, J.D., et al., *Use of biomarkers in the discovery of novel anti-schizophrenia drugs.* Drug Discov Today, 2010. **15**(3-4): p. 137-41.
7. Kuzman, M.R., et al., *Genome-wide expression analysis of peripheral blood identifies candidate biomarkers for schizophrenia.* J Psychiatr Res, 2009. **43**(13): p. 1073-7.
8. Frangou, S. and J. Kington, *Schizophrenia.* Medicine, 2004. **32**(7): p. 21-25.
9. Abbott, C. and S. Keith, *Antipsychotic treatment and adherence in schizophrenia*. 2011, Oxford University Press, Oxford.
10. Jablensky, A.a.S., N. and Ernberg,G. and Anker,M. and Korten,A. and Cooper,J. E. and Day,R. and Bertelsen,A, *Schizophrenia - manifestations, incidence and course in different cultures A World Health Organization Ten-Country Study.* Psychological Medicine. Monograph Supplement, 1992. **Supplement**: p. 1--97.
11. Saha, S., D. Chant, and J. McGrath, *A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time?* Archives of general psychiatry, 2007. **64**(10): p. 1123-1131.
12. Courtenay M. Harding, J.Z., and John S. Strauss, *Chronicity in Schizophrenia: Fact, Partial Fact, or Artifact?* Psychiatric Services, 1987. **38**(5): p. 477-486.
13. Tandon, R., M.S. Keshavan, and H.A. Nasrallah, *Schizophrenia, "just the facts" what we know in 2008. 2. Epidemiology and etiology.* Schizophr Res, 2008. **102**(1-3): p. 1-18.
14. Emsley, R., et al., *The nature of relapse in schizophrenia.* BMC Psychiatry, 2013. **13**: p. 50.
15. Association, A.P., *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition: DSM-IV-TR®*. 2000: American Psychiatric Association.
16. Organization, W.H., *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. 1993: World Health Organization.
17. Sadock, B.J. and V.A. Sadock, *Kaplan and Sadock's Synopsis of Psychiatry: Behavioral Sciences/Clinical Psychiatry*. 2011: Wolters Kluwer Health.

18. Gladkevich, A., H.F. Kauffman, and J. Korf, *Lymphocytes as a neural probe: potential for studying psychiatric disorders.* Prog Neuropsychopharmacol Biol Psychiatry, 2004. **28**(3): p. 559-76.
19. Fenton, W.S., J. Hibbeln, and M. Knable, *Essential fatty acids, lipid membrane abnormalities, and the diagnosis and treatment of schizophrenia.* Biol Psychiatry, 2000. **47**(1): p. 8-21.
20. Tregellas, J.R., *Neuroimaging biomarkers for early drug development in schizophrenia.* Biol Psychiatry, 2014. **76**(2): p. 111-9.
21. Vita, A., et al., *Brain morphology in first-episode schizophrenia: a meta-analysis of quantitative magnetic resonance imaging studies.* Schizophr Res, 2006. **82**(1): p. 75-88.
22. Ellison-Wright, I. and E. Bullmore, *Meta-analysis of diffusion tensor imaging studies in schizophrenia.* Schizophr Res, 2009. **108**(1-3): p. 3-10.
23. Glahn, D.C., et al., *Meta-analysis of gray matter anomalies in schizophrenia: application of anatomic likelihood estimation and network analysis.* Biol Psychiatry, 2008. **64**(9): p. 774-81.
24. Focking, M., et al., *Common proteomic changes in the hippocampus in schizophrenia and bipolar disorder and particular evidence for involvement of cornu ammonis regions 2 and 3.* Arch Gen Psychiatry, 2011. **68**(5): p. 477-88.
25. Williamson, P.C. and J.M. Allman, *A framework for interpreting functional networks in schizophrenia.* Front Hum Neurosci, 2012. **6**: p. 184.
26. Cadet, J.L., et al., *Dopamine D1 receptors, regulation of gene expression in the brain, and neurodegeneration.* CNS Neurol Disord Drug Targets, 2010. **9**(5): p. 526-38.
27. Harley, C.W., *Norepinephrine and dopamine as learning signals.* Neural plasticity, 2004. **11**(3-4): p. 191--204.
28. Guidotti, A., et al., *GABAergic dysfunction in schizophrenia: new treatment strategies on the horizon.* Psychopharmacology (Berl), 2005. **180**(2): p. 191-205.
29. Lewis, D.A. and B. Moghaddam, *Cognitive dysfunction in schizophrenia: convergence of gamma-aminobutyric acid and glutamate alterations.* Arch Neurol, 2006. **63**(10): p. 1372-6.
30. Gray, J.A. and B.L. Roth, *The pipeline and future of drug development in schizophrenia.* Mol Psychiatry, 2007. **12**(10): p. 904-22.
31. Nakai, T., et al., *Alterations of GABAergic and dopaminergic systems in mutant mice with disruption of exons 2 and 3 of the Disc1 gene.* Neurochem Int, 2014. **74**: p. 74-83.
32. Chavez-Noriega, L.E., H. Schaffhauser, and U.C. Campbell, *Metabotropic glutamate receptors: potential drug targets for the treatment of schizophrenia.* Curr Drug Targets CNS Neurol Disord, 2002. **1**(3): p. 261-81.

33. Kantrowitz, J.T. and D.C. Javitt, *N-methyl-d-aspartate (NMDA) receptor dysfunction or dysregulation: the final common pathway on the road to schizophrenia?* Brain Res Bull, 2010. **83**(3-4): p. 108-21.
34. Weinberger, D.R., *Implications of normal brain development for the pathogenesis of schizophrenia.* Arch Gen Psychiatry, 1987. **44**(7): p. 660-9.
35. Davis, K.L., et al., *Dopamine in schizophrenia: a review and reconceptualization.* Am J Psychiatry, 1991. **148**(11): p. 1474-86.
36. Allman, J., *Evolving Brains*. 2000: Henry Holt and Company.
37. Graeff, F.G., et al., *Role of 5-HT in stress, anxiety, and depression.* Pharmacol Biochem Behav, 1996. **54**(1): p. 129-41.
38. Tandon, R., et al., *The cholinergic system in schizophrenia reconsidered: anticholinergic modulation of sleep and symptom profiles.* Neuropsychopharmacology, 1999. **21**: p. S189-S202.
39. Frangou, S., *Schizophrenia.* Medicine, 2008. **36**(8): p. 405-409.
40. Kapur, S., et al., *How antipsychotics work-from receptors to reality.* NeuroRx, 2006. **3**(1): p. 10-21.
41. Huffman, J.C. and J.E. Alpert, *An approach to the psychopharmacologic care of patients: antidepressants, antipsychotics, anxiolytics, mood stabilizers, and natural remedies.* Med Clin North Am, 2010. **94**(6): p. 1141-60, x.
42. Park, J., et al., *Haloperidol and clozapine block formation of autophagolysosomes in rat primary neurons.* Neuroscience, 2012. **209**: p. 64-73.
43. Gerace, E., et al., *Evidence of Haldol (haloperidol) long-term intoxication.* Forensic Sci Int, 2012. **215**(1-3): p. 121-3.
44. Ebrahimzadeh, H., et al., *Determination of haloperidol in biological samples using molecular imprinted polymer nanoparticles followed by HPLC-DAD detection.* Int J Pharm, 2013. **453**(2): p. 601-9.
45. Settle, E.C., Jr. and F.J. Ayd, Jr., *Haloperidol: a quarter century of experience.* J Clin Psychiatry, 1983. **44**(12): p. 440-8.
46. Adams, C.E., et al., *Haloperidol versus placebo for schizophrenia.* Cochrane Database Syst Rev, 2013. **11**: p. CD003082.
47. Newcomer, J.W., *Second-generation (atypical) antipsychotics and metabolic effects: a comprehensive literature review.* CNS Drugs, 2005. **19 Suppl 1**: p. 1-93.
48. Freudenreich, O. and D.C. Goff, *Antipsychotic combination therapy in schizophrenia. A review of efficacy and risks of current combinations.* Acta Psychiatr Scand, 2002. **106**(5): p. 323-30.

49. Popovic, I., et al., *First generation antipsychotics switch with Risperidone in the treatment of chronic schizophrenic patients.* Psychiatr Danub, 2011. **23**(4): p. 384-8.
50. Kucharska-Pietura, K. and A. Mortimer, *Can antipsychotics improve social cognition in patients with schizophrenia?* CNS Drugs, 2013. **27**(5): p. 335-43.
51. Kendall, T., *The rise and fall of the atypical antipsychotics.* Br J Psychiatry, 2011. **199**(4): p. 266-8.
52. Kane, J.M., *A user' guide to clozapine.* Acta Psychiatr Scand, 2011. **123**(6): p. 407-8.
53. Zhang, J.P. and A.K. Malhotra, *Pharmacogenetics of antipsychotics: recent progress and methodological issues.* Expert Opin Drug Metab Toxicol, 2013. **9**(2): p. 183-91.
54. Lett, T.A., et al., *Treating working memory deficits in schizophrenia: a review of the neurobiology.* Biol Psychiatry, 2014. **75**(5): p. 361-70.
55. Abi-Dargham, A. and M. Laruelle, *Mechanisms of action of second generation antipsychotic drugs in schizophrenia: insights from brain imaging studies.* Eur Psychiatry, 2005. **20**(1): p. 15-27.
56. Duncan, G.E., et al., *Comparison of the effects of clozapine, risperidone, and olanzapine on ketamine-induced alterations in regional brain metabolism.* J Pharmacol Exp Ther, 2000. **293**(1): p. 8-14.
57. Attia, S.M. and S.A. Bakheet, *Citalopram at the recommended human doses after long-term treatment is genotoxic for male germ cell.* Food Chem Toxicol, 2013. **53**: p. 281-5.
58. Richelson, E., *Pharmacology of antidepressants--characteristics of the ideal drug.* Mayo Clin Proc, 1994. **69**(11): p. 1069-81.
59. Beaune, S., et al., *Mechanisms of high-dose citalopram-induced death in a rat model.* Toxicology, 2012. **302**(2-3): p. 248-54.
60. Karlsson, L., et al., *Altered brain concentrations of citalopram and escitalopram in P-glycoprotein deficient mice after acute and chronic treatment.* Eur Neuropsychopharmacol, 2013. **23**(11): p. 1636-44.
61. Kaddurah-Daouk, R., et al., *Metabolomic mapping of atypical antipsychotic effects in schizophrenia.* Mol Psychiatry, 2007. **12**(10): p. 934-45.
62. Xiao, J.F., B. Zhou, and H.W. Ressom, *Metabolite identification and quantitation in LC-MS/MS-based metabolomics.* Trends Analyt Chem, 2012. **32**: p. 1-14.
63. Weckwerth, W. and K. Morgenthal, *Metabolomics: from pattern recognition to biological interpretation.* Drug Discovery Today, 2005. **10**(22): p. 1551-1558.
64. Xuan, J., et al., *Metabolomic profiling to identify potential serum biomarkers for schizophrenia and risperidone action.* J Proteome Res, 2011. **10**(12): p. 5433-43.

65. Dettmer, K. and B.D. Hammock, *Metabolomics--a new exciting field within the "omics" sciences.* Environ Health Perspect, 2004. **112**(7): p. A396-7.
66. Dettmer, K., P.A. Aronov, and B.D. Hammock, *Mass spectrometry-based metabolomics.* Mass Spectrom Rev, 2007. **26**(1): p. 51-78.
67. Kaddurah-Daouk, R., *Metabolic profiling of patients with schizophrenia.* PLoS Med, 2006. **3**(8): p. e363.
68. He, Y., et al., *Schizophrenia shows a unique metabolomics signature in plasma.* Transl Psychiatry, 2012. **2**: p. e149.
69. Webhofer, C., et al., *Metabolite profiling of antidepressant drug action reveals novel drug targets beyond monoamine elevation.* Transl Psychiatry, 2011. **1**: p. e58.
70. Schnackenberg, L.K. and R.D. Beger, *Metabolomic biomarkers: their role in the critical path.* Drug Discov Today Technol, 2007. **4**(1): p. 13-6.
71. Lei, Z., D.V. Huhman, and L.W. Sumner, *Mass spectrometry strategies in metabolomics.* J Biol Chem, 2011. **286**(29): p. 25435-42.
72. Berg, M., et al., *LC-MS metabolomics from study design to data-analysis - using a versatile pathogen as a test case.* Comput Struct Biotechnol J, 2013. **4**: p. e201301002.
73. Snyder, L.R., J.J. Kirkland, and J.W. Dolan, *Introduction to Modern Liquid Chromatography*. 2011: Wiley.
74. Buszewski, B. and S. Noga, *Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique.* Anal Bioanal Chem, 2012. **402**(1): p. 231-47.
75. Niessen, W.M.A., *Liquid Chromatography-Mass Spectrometry, Third Edition*. 2006: Taylor & Francis.
76. Skoog, D.A., F.J. Holler, and S.R. Crouch, *Principles of Instrumental Analysis*. 2007: Thomson Brooks/Cole.
77. Ho, C.S., et al., *Electrospray ionisation mass spectrometry: principles and clinical applications.* Clin Biochem Rev, 2003. **24**(1): p. 3-12.
78. Ni, J. and J. Rowe, *Microdosing Assessment to Evaluate Pharmacokinetics and Drug Metabolism Using Liquid Chromatography-Tandem Mass Spectrometry Technology.* 2012.
79. Banerjee, S. and S. Mazumdar, *Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte.* International journal of analytical chemistry, 2012. **2012**.
80. de Hoffmann, E. and V. Stroobant, *Mass Spectrometry: Principles and Applications*. 2013: Wiley.
81. Dawson, P., *Quadrupole mass analyzers: performance, design and some recent applications.* Mass Spectrometry Reviews, 1986. **5**(1): p. 1-37.

82. Ardrey, R.E., *Liquid Chromatography - Mass Spectrometry: An Introduction*. 2003: Wiley.

83. http://www.waters.com/webassets/cms/category/media/other_images/ms_primer_p2_fig1.jpg.

84. Glish, G.L. and R.W. Vachet, *The basics of mass spectrometry in the twenty-first century.* Nat Rev Drug Discov, 2003. **2**(2): p. 140-50.

85. Pitt, J.J., *Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry.* Clin Biochem Rev, 2009. **30**(1): p. 19-34.

86. Ashton, P.D., R.S. Curwen, and R.A. Wilson, *Linking proteome and genome: how to identify parasite proteins.* Trends in parasitology, 2001. **17**(4): p. 198-202.

87. Katajamaa, M. and M. Oresic, *Processing methods for differential analysis of LC/MS profile data.* BMC Bioinformatics, 2005. **6**: p. 179.

88. Milne, S.B., et al., *Sum of the parts: mass spectrometry-based metabolomics.* Biochemistry, 2013. **52**(22): p. 3829-40.

89. Katajamaa, M. and M. Oresic, *Data processing for mass spectrometry-based metabolomics.* J Chromatogr A, 2007. **1158**(1-2): p. 318-28.

90. Nyamundanda, G., L. Brennan, and I.C. Gormley, *Probabilistic principal component analysis for metabolomic data.* BMC bioinformatics, 2010. **11**(1): p. 571.

91. Davies, T. and T. Fearn, *Back to basics: the principles of principal component analysis.* Spectroscopy Europe, 2004. **16**(6): p. 20.

92. Jolliffe, I.T., *Principal Component Analysis*. 2002: Springer.

93. Cova, T.F., J.L. Pereira, and A.A. Pais, *Is standard multivariate analysis sufficient in clinical and epidemiological studies?* Journal of biomedical informatics, 2013. **46**(1): p. 75-86.

7. Supplementary data

A) Internal standard tests

Table S 1 – Compounds names and respective concentration used.

	Name	Conc (mM)
#01	Amoxicillin trihydrate VETRANAL®. analytical standard	1.08
#03	Brombuterol hydrochloride VETRANAL®. analytical standard	1.08
#04	Cefquinome Sulfate	0.73
#05	Ceftiofur VETRANAL®. analytical standard	0.37
#06	Chloramphenicol ≥98% (TLC)	0.63
#07	Chlortetracycline hydrochloride VETRANAL®. analytical standard	0.84
#08	Cimaterol VETRANAL®. analytical standard	1.82
#09	cimbuterol	1.71
#10	clenbuterol hydrochloride	1.44
#11	clencyclohexerol	1.25
#12	Clenproperol VETRANAL®. analytical standard	1.52
#19	Doxycycline hyclate VETRANAL®. analytical standard	0.88
#21	florfenicol	1.14
#22	Gamithromycin (Zactran)	0.51
#26	Isoxsuprine hydrochloride analytical standard. for drug analysis	1.41
#27	lincomycin hydrochloride monohydrate	1.13
#28	Mabuterol hydrochloride VETRANAL®. analytical standard	1.29
#29	Mapenterol hydrochloride VETRANAL®. analytical standard	0.61
#30	Medroxyprogesterone 17-acetate VETRANAL®. analytical standard	1.04
#31	Megestrol acetate VETRANAL®. analytical standard	1.2
#32	Melengestrol acetate VETRANAL®. analytical standard	1.05
#34	Monensin sodium salt hydrate	0.66
#35	Narasin from Streptomyces auriofaciens	0.26
#36	Metaproterenol hemisulfate salt (orciprenalin)	1.83
#37	Oxytetracycline hydrochloride VETRANAL®. analytical standard	0.85
#38	Prednisolone 21-acetate VETRANAL®. analytical standard	0.98
#39	Ractopamine hydrochloride VETRANAL®. analytical standard	1.37
#40	Ritodrine hydrochloride	1.39
#41	Robenidine hydrochloride VETRANAL®. analytical standard	1.2
#42	Salbutamol	1.57
#43	Salmeterol xinafoate	0.94
#45	Sulfachloropyridazine VETRANAL®. analytical standard	1.42
#46	Sulfadiazine VETRANAL®. analytical standard (Fluka)	1.69
#47	Sulfadimethoxine VETRANAL®. analytical standard	1.31
#48	Sulfadoxin VETRANAL®. analytical standard	1.40
#49	Sulfamethazine VETRANAL®. analytical standard	1.39
#50	Sulfamethoxypyridazine VETRANAL®. analytical standard	1.44
#51	Terbutalin sulfate	1.73
#52	Tetracycline hydrochloride	0.96
#53	Tiamulin VETRANAL®. analytical standard	8.09
#54	Tilmicosin - VETRANAL®. analytical standard (Fluka)	0.46
#55	Trenbolone acetate	1.30
#56	Triamcinolone acetonide analytical standard	0.91
#57	Trimethoprim VETRANAL®. analytical standard	1.38
#58	Tulathromycin (Draxxin)	0.50
#59	Tulobuterol hydrochloride VETRANAL®. analytical standard	1.72
#60	Tylosin tartrate VETRANAL®. analytical standard	0.41
#61	Valnemulin VETRANAL®. analytical standard	0.69
#62	α-Zearalanol ~97% (HPLC)	0.12
#63	Zilpaterol HCl	0.76
#96	Clenpenterol hydrochloride VETRANAL®. analytical standard	1.37
#98	Penicilline V potassium salt	1.28
#99	Sulfamethazine-D4	0.14

B) Ratios and fragmentation mass spectra

i. Control vs Citalopram

Table S 2 - Ratios obtained for Control and Citalopram samples, considering the I.S and the reduced set of 75 variables.

Peak name $_{((m/z)/RT)}$	CT1	CT2	CT3	CT4	CI1	CI2	CI3	CI4
175.0/19.1	0.007	0.012	0.006	0.007	0.006	0.006	0.004	0.003
607.7/35.9	0.042	0.021	0.034	0.032	0.018	0.017	0.005	0.013
321.1/17.4	0.026	0.015	0.027	0.029	0.032	0.058	0.050	0.035
282.3/39.4	0.000	0.000	0.011	0.000	0.020	0.013	0.021	0.012
1303.7/38.8	0.050	0.054	0.048	0.091	0.020	0.023	0.023	0.025
323.1/19.0	0.009	0.020	0.010	0.000	0.035	0.033	0.056	0.040
1304.2/38.9	0.049	0.051	0.051	0.083	0.022	0.024	0.024	0.023
538.8/39.0	0.151	0.134	0.166	0.195	0.065	0.057	0.071	0.061
385.1/18.9	0.034	0.056	0.036	0.042	0.016	0.030	0.029	0.017
245.1/23.5	0.012	0.026	0.020	0.058	0.067	0.134	0.098	0.127
800.8/35.8	0.036	0.044	0.034	0.046	0.020	0.021	0.016	0.019
1049.1/38.6	0.143	0.154	0.129	0.246	0.057	0.059	0.069	0.069
793.4/38.6	0.250	0.269	0.245	0.389	0.102	0.098	0.136	0.118
1059.6/38.9	0.003	0.003	0.005	0.009	0.012	0.013	0.014	0.016
1558.9/38.8	0.029	0.027	0.026	0.045	0.014	0.012	0.016	0.016
1048.6/38.9	0.136	0.140	0.128	0.232	0.061	0.058	0.072	0.075
457.1/32.3	0.002	0.006	0.010	0.009	0.044	0.024	0.073	0.031
325.2/25.1	0.211	0.126	0.149	0.100	0.006	0.003	0.015	0.003
354.1/36.1	0.017	0.025	0.021	0.033	0.013	0.010	0.006	0.007
1559.4/39.0	0.032	0.028	0.032	0.055	0.013	0.015	0.017	0.016
793.9/38.6	0.193	0.212	0.201	0.324	0.083	0.081	0.107	0.100
414.2/37.6	0.002	0.002	0.013	0.002	0.018	0.028	0.022	0.014
277.1/30.8	0.003	0.005	0.004	0.012	0.014	0.018	0.037	0.017
487.1/35.8	0.053	0.040	0.035	0.069	0.012	0.027	0.007	0.012
384.0/18.9	0.041	0.062	0.042	0.069	0.020	0.024	0.031	0.027
792.9/38.8	0.055	0.050	0.063	0.072	0.025	0.028	0.035	0.035
219.0/18.9	0.021	0.037	0.017	0.043	0.015	0.016	0.004	0.004
492.3/35.3	0.044	0.066	0.025	0.065	0.016	0.019	0.010	0.032
537.8/38.6	0.026	0.024	0.036	0.041	0.010	0.018	0.003	0.016
457.1/33.2	0.007	0.004	0.005	0.004	0.028	0.009	0.041	0.022
297.1/17.5	0.030	0.045	0.029	0.049	0.012	0.018	0.026	0.027
252.1/18.7	0.017	0.023	0.023	0.030	0.009	0.012	0.013	0.014
624.2/18.9	0.029	0.047	0.052	0.081	0.025	0.023	0.030	0.024
1891.1/38.8	0.010	0.006	0.008	0.011	0.005	0.003	0.004	0.004
660.2/9.5	0.001	0.002	0.004	0.006	0.017	0.017	0.026	0.034
433.3/31.9	0.012	0.024	0.008	0.013	0.008	0.002	0.011	0.004
486.1/36.1	0.021	0.024	0.022	0.030	0.014	0.018	0.006	0.009
1559.9/39.0	0.050	0.050	0.047	0.089	0.025	0.027	0.032	0.034
425.1/35.9	0.004	0.093	0.011	0.101	0.303	0.181	0.210	0.178
355.0/33.2	0.031	0.029	0.024	0.028	0.018	0.008	0.015	0.011
508.0/38.0	0.048	0.037	0.053	0.073	0.018	0.024	0.022	0.020
315.0/23.9	0.005	0.004	0.004	0.008	0.001	0.002	0.002	0.004
747.5/39.1	0.000	0.000	0.001	0.000	0.013	0.003	0.008	0.007
667.3/29.2	0.007	0.018	0.002	0.011	0.001	0.000	0.002	0.003
288.3/35.4	0.042	0.055	0.047	0.079	0.039	0.023	0.027	0.022
328.1/9.7	0.000	0.002	0.013	0.009	0.035	0.015	0.038	0.055
1048.1/38.9	0.029	0.029	0.032	0.047	0.016	0.018	0.022	0.018
326.1/33.1	0.002	0.007	0.003	0.001	0.016	0.008	0.012	0.012

Table S 2 – (cont.)

466.9/37.7	0.021	0.021	0.026	0.043	0.010	0.015	0.017	0.010
629.2/18.7	0.015	0.021	0.017	0.025	0.011	0.010	0.007	0.001
212.2/27.5	0.054	0.062	0.060	0.062	0.053	0.031	0.052	0.049
241.0/13.9	0.011	0.023	0.017	0.034	0.010	0.000	0.010	0.001
378.2/18.3	0.013	0.020	0.003	0.014	0.003	0.001	0.004	0.002
523.3/25.5	0.077	0.106	0.059	0.138	0.045	0.018	0.076	0.050
268.2/21.0	0.014	0.024	0.018	0.017	0.012	0.004	0.013	0.015
713.4/32.6	0.003	0.005	0.002	0.005	0.001	0.001	0.002	0.002
195.1/29.5	0.012	0.012	0.011	0.013	0.008	0.006	0.010	0.010
622.7/18.9	0.012	0.015	0.016	0.016	0.007	0.006	0.010	0.005
268.1/26.1	0.025	0.032	0.022	0.065	0.010	0.011	0.016	0.016
598.3/26.7	0.014	0.032	0.018	0.031	0.008	0.006	0.007	0.004
312.1/18.7	0.035	0.059	0.035	0.065	0.018	0.014	0.002	0.031
326.1/9.5	0.000	0.002	0.000	0.009	0.022	0.013	0.035	0.042
514.3/25.5	0.030	0.060	0.016	0.042	0.010	0.007	0.024	0.012
584.3/39.3	0.005	0.004	0.005	0.003	0.001	0.000	0.001	0.004
519.3/32.5	0.015	0.026	0.010	0.025	0.009	0.001	0.014	0.006
527.2/34.5	0.008	0.014	0.009	0.012	0.008	0.008	0.006	0.008
415.1/35.9	0.076	0.091	0.078	0.092	0.057	0.062	0.045	0.056
295.2/39.0	0.052	0.108	0.027	0.046	0.019	0.015	0.015	0.027
263.0/13.9	0.018	0.021	0.019	0.033	0.006	0.004	0.010	0.006
1074.5/39.2	0.032	0.053	0.047	0.072	0.023	0.024	0.027	0.028
792.4/38.9	0.024	0.025	0.027	0.031	0.008	0.011	0.012	0.011
354.3/45.2	0.083	0.088	0.032	0.151	0.021	0.032	0.007	0.027
555.3/38.6	0.014	0.007	0.009	0.012	0.003	0.006	0.005	0.001
409.3/21.1	0.055	0.117	0.055	0.096	0.045	0.013	0.011	0.010
684.4/44.0	0.293	0.271	0.268	0.371	0.201	0.128	0.012	0.218

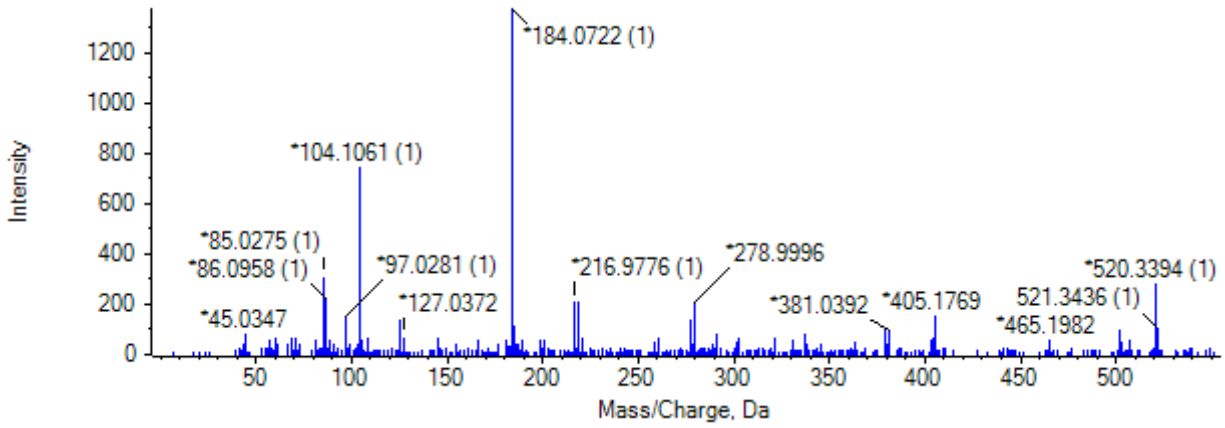

Figure S 1 – Fragmentation mass spectrum of the precursor with a m/z value of 538.8, in positive mode.

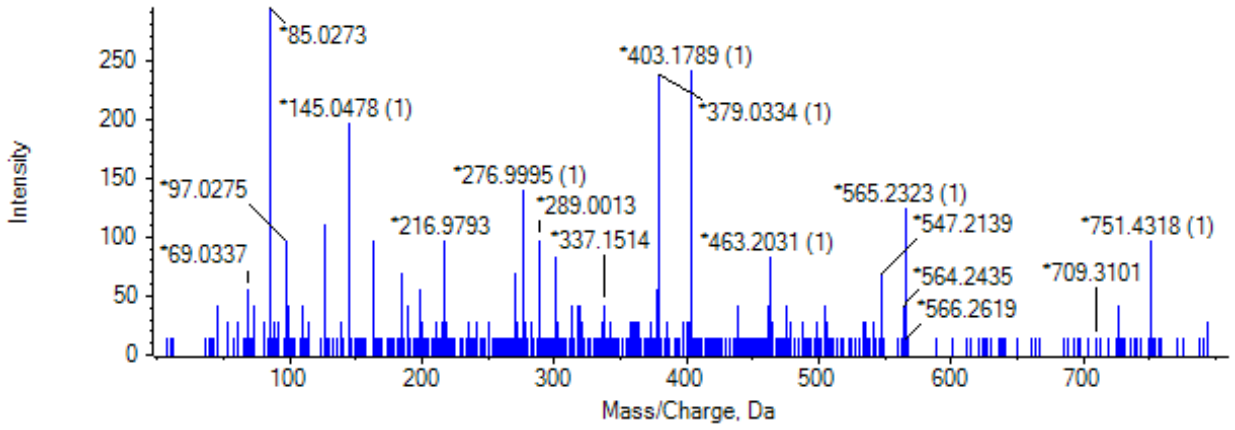

Figure S 2 -- Fragmentation mass spectrum of the precursor with a m/z value of 793.4, in positive mode.

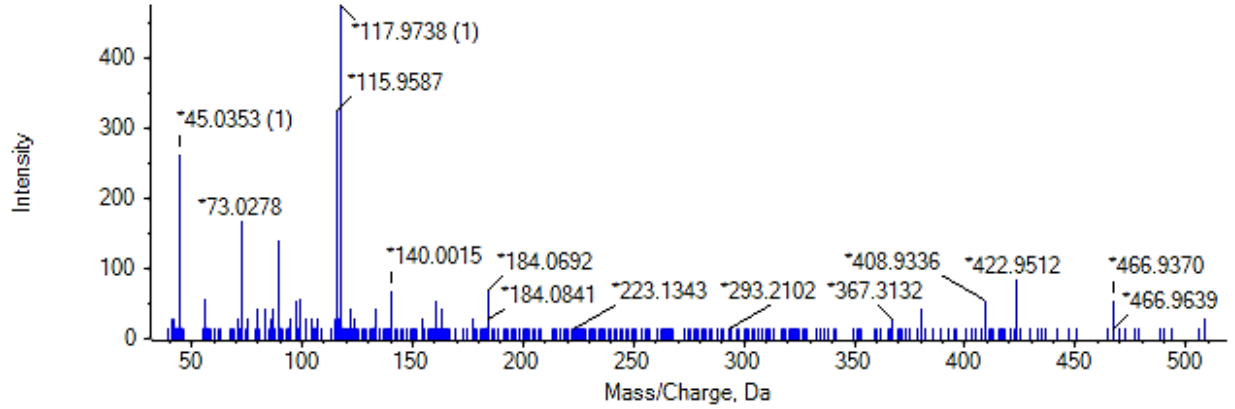

Figure S 3 -- Fragmentation mass spectrum of the precursor with a m/z value of 508.0, in positive mode.

ii. Control vs Haloperidol

Table S 3 - Ratios obtained for Control and Haloperidol samples, considering the I.S and the reduced set of 17 variables.

Peak name $_{((m/z)/RT)}$	CT1	CT2	CT3	CT4	HA1	HA2	HA3	HA4
538.8/39.0	0.031	0.027	0.033	0.035	0.019	0.021	0.025	0.024
294.1/35.9	0.002	0.004	0.002	0.003	0.001	0.004	0.002	0.001
197.2/30.2	0.005	0.005	0.005	0.004	0.008	0.006	0.012	0.012
536.3/45.4	0.004	0.001	0.002	0.001	0.004	0.004	0.004	0.005
1304.2/38.9	0.010	0.010	0.012	0.015	0.009	0.008	0.011	0.007
508.3/33.3	0.011	0.007	0.011	0.008	0.011	0.002	0.014	0.003
258.1/34.9	0.003	0.003	0.004	0.001	0.006	0.006	0.005	0.007
456.3/14.8	0.003	0.002	0.003	0.002	0.003	0.004	0.007	0.005
260.2/38.7	0.008	0.012	0.012	0.004	0.020	0.015	0.018	0.017
485.4/13.4	0.014	0.003	0.010	0.004	0.017	0.004	0.009	0.007
843.3/32.2	0.002	0.002	0.002	0.001	0.002	0.001	0.002	0.002
192.0/18.7	0.004	0.006	0.006	0.007	0.004	0.004	0.006	0.005
277.1/13.4	0.001	0.001	0.001	0.002	0.007	0.001	0.004	0.005
510.3/30.7	0.003	0.003	0.003	0.006	0.014	0.003	0.005	0.005
218.1/11.5	0.002	0.002	0.002	0.003	0.000	0.003	0.004	0.003
594.2/18.9	0.002	0.001	0.003	0.002	0.001	0.002	0.001	0.001
208.2/30.2	0.006	0.007	0.010	0.007	0.013	0.011	0.029	0.029

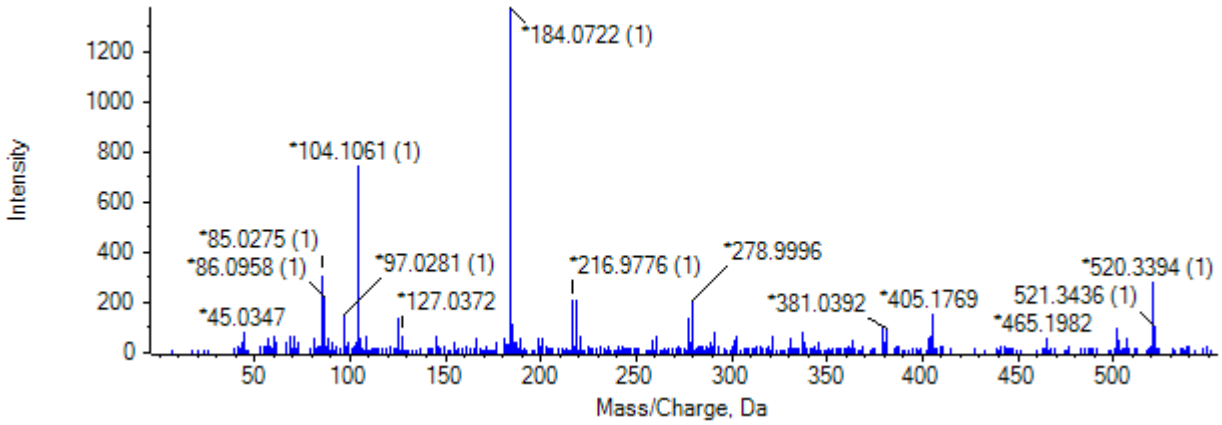

Figure S 4 - Fragmentation mass spectrum for the precursor with a m/z value of 538.8, in positive mode.

iii. Control vs Clozapine

Table S 4 - Ratios obtained for Control and Clozapine samples, considering the I.S and the reduced set of 57 variables.

Peak name $_{((m/z)/RT)}$	CT1	CT2	CT3	CT4	CL1	CL2	CL3	CL4
354.1/36.1	0.004	0.005	0.005	0.006	0.002	0.002	0.001	0.001
294.1/35.9	0.002	0.004	0.002	0.003	0.001	0.001	0.001	0.001
660.2/9.5	0.002	0.000	0.002	0.001	0.006	0.004	0.004	0.003
607.2/35.8	0.011	0.008	0.008	0.008	0.006	0.006	0.002	0.003
178.2/19.2	0.007	0.004	0.008	0.006	0.004	0.001	0.002	0.002
800.8/35.8	0.006	0.006	0.006	0.006	0.003	0.003	0.001	0.002
457.1/33.2	0.001	0.001	0.001	0.001	0.005	0.004	0.003	0.003
513.3/34.5	0.013	0.014	0.009	0.012	0.008	0.006	0.008	0.007
487.1/36.1	0.011	0.006	0.008	0.009	0.005	0.004	0.004	0.003
613.3/22.0	0.002	0.002	0.003	0.002	0.002	0.001	0.002	0.001
800.3/35.8	0.009	0.010	0.009	0.008	0.005	0.004	0.002	0.003
292.2/15.6	0.004	0.002	0.002	0.003	0.001	0.000	0.000	0.000
607.7/35.9	0.007	0.004	0.005	0.004	0.003	0.004	0.003	0.004
323.1/19.0	0.001	0.003	0.001	0.002	0.004	0.005	0.006	0.004
301.7/15.6	0.009	0.006	0.006	0.007	0.006	0.000	0.004	0.000
588.4/44.1	0.004	0.001	0.004	0.004	0.006	0.009	0.007	0.008
1303.7/38.8	0.010	0.011	0.011	0.016	0.005	0.004	0.005	0.006
1814.0/39.0	0.002	0.002	0.002	0.003	0.001	0.001	0.001	0.001
1304.2/38.9	0.010	0.010	0.012	0.015	0.005	0.005	0.006	0.006
538.8/39.0	0.031	0.027	0.033	0.035	0.014	0.014	0.018	0.017
790.4/23.4	0.002	0.005	0.002	0.003	0.002	0.001	0.001	0.002
1049.1/38.9	0.030	0.031	0.030	0.044	0.014	0.013	0.018	0.016
1558.9/38.7	0.005	0.005	0.005	0.007	0.003	0.002	0.003	0.003
554.3/45.2	0.026	0.017	0.022	0.029	0.048	0.034	0.045	0.030
224.1/16.3	0.036	0.042	0.037	0.036	0.009	0.007	0.034	0.030
793.4/38.8	0.052	0.054	0.051	0.070	0.025	0.026	0.030	0.034
600.4/15.6	0.012	0.000	0.008	0.000	0.000	0.000	0.000	0.000
792.4/38.9	0.005	0.005	0.004	0.006	0.002	0.002	0.002	0.003
473.1/33.3	0.004	0.005	0.003	0.004	0.003	0.002	0.002	0.002
793.9/38.8	0.040	0.042	0.046	0.058	0.021	0.019	0.023	0.027
1559.9/38.8	0.010	0.010	0.011	0.016	0.008	0.008	0.007	0.007
1891.1/38.9	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.001
321.1/14.9	0.002	0.006	0.002	0.005	0.007	0.013	0.014	0.007
464.8/26.7	0.003	0.006	0.001	0.002	0.001	0.001	0.002	0.001
294.2/14.4	0.005	0.002	0.004	0.002	0.001	0.001	0.001	0.001
1048.6/38.9	0.028	0.028	0.029	0.042	0.013	0.013	0.017	0.018
808.8/36.1	0.002	0.002	0.002	0.001	0.001	0.000	0.000	0.001
1559.4/39.0	0.007	0.006	0.007	0.010	0.003	0.003	0.003	0.004
627.3/22.9	0.005	0.004	0.004	0.004	0.003	0.002	0.003	0.003
542.3/38.8	0.001	0.001	0.001	0.002	0.003	0.003	0.004	0.001
453.2/18.3	0.006	0.014	0.003	0.011	0.019	0.046	0.032	0.011
493.2/39.1	0.004	0.006	0.004	0.003	0.002	0.003	0.003	0.003
541.3/19.9	0.005	0.010	0.005	0.010	0.013	0.025	0.025	0.010
583.3/29.7	0.004	0.007	0.002	0.004	0.001	0.002	0.003	0.001
409.2/17.3	0.004	0.008	0.001	0.006	0.009	0.021	0.014	0.006
713.4/32.6	0.001	0.001	0.000	0.001	0.000	0.000	0.000	0.000
282.3/39.7	0.007	0.003	0.004	0.006	0.013	0.010	0.008	0.008
506.5/30.8	0.002	0.005	0.002	0.003	0.001	0.001	0.002	0.001
588.3/31.8	0.007	0.007	0.004	0.006	0.001	0.002	0.003	0.003
365.2/16.2	0.004	0.007	0.003	0.006	0.009	0.023	0.017	0.007
508.0/38.0	0.010	0.007	0.012	0.013	0.004	0.004	0.006	0.006
480.3/41.1	0.000	0.023	0.000	0.000	0.047	0.036	0.037	0.070
486.1/36.1	0.004	0.005	0.005	0.005	0.004	0.004	0.003	0.004

Table S 4 – (cont.)

283.3/41.8	0.000	0.002	0.000	0.004	0.006	0.004	0.004	0.005
277.1/13.5	0.001	0.001	0.001	0.002	0.002	0.004	0.006	0.004
577.2/36.7	0.008	0.004	0.009	0.011	0.008	0.003	0.002	0.004
550.6/31.4	0.001	0.002	0.001	0.001	0.000	0.000	0.001	0.001

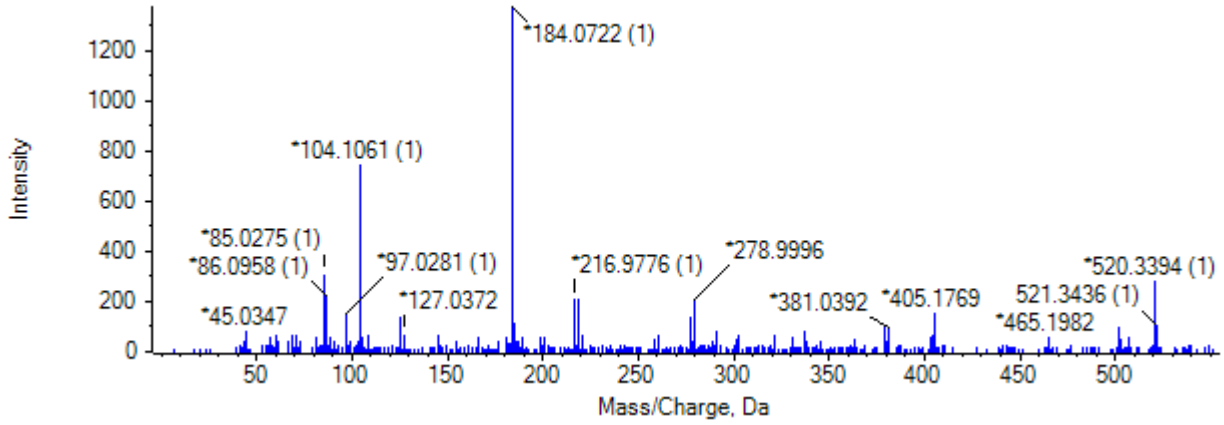

Figure S 5 – Fragmentation mass spectrum for the precursor with a m/z value of 538.8, in positive mode.

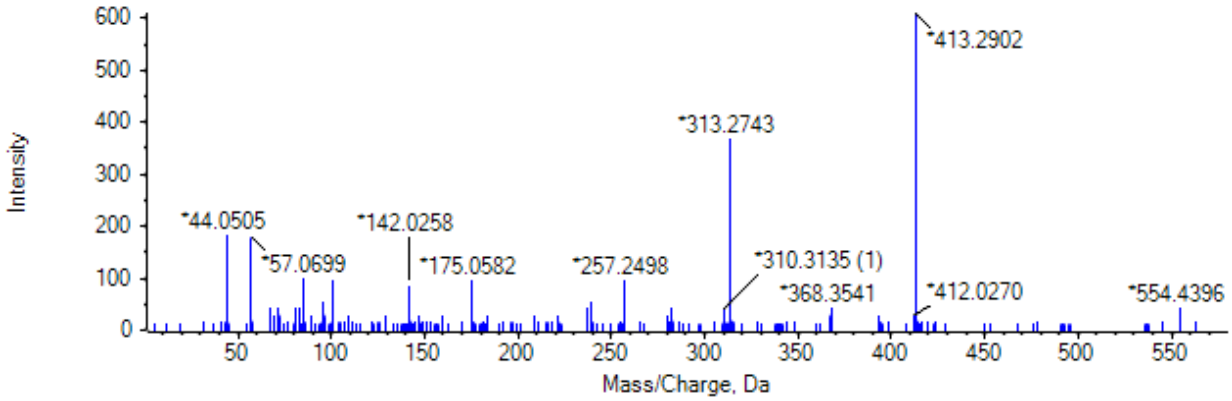

Figure S 6 – Fragmentation mass spectrum for the precursor with a m/z value of 554.3, in positive mode.

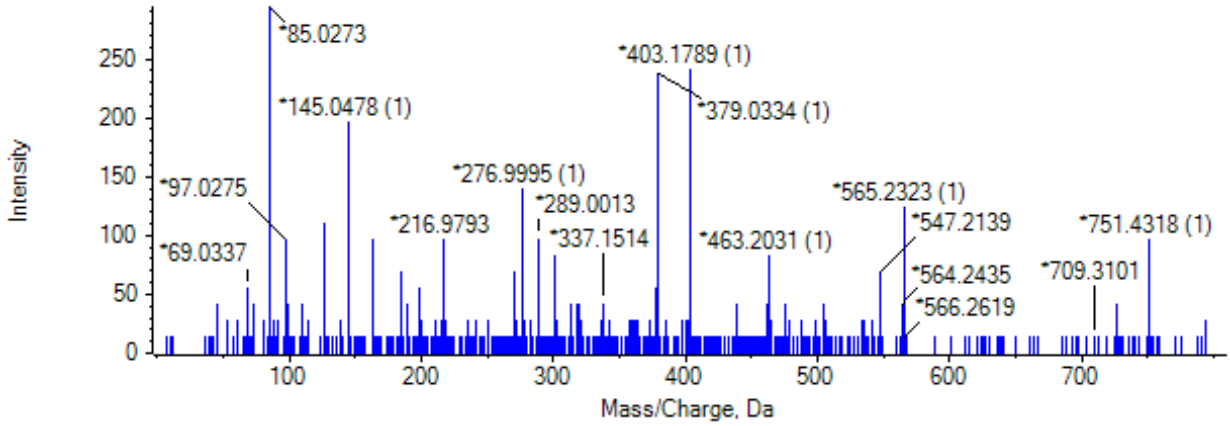

Figure S 7 - Fragmentation mass spectrum for the precursor with a m/z value of 793.4, in positive mode.

C) Loading Values

i. Control vs Citalopram

Table S 5 - Significant loading values (>0.068) of each variable on the first two principal components. The contributions are sorted in descending order along PC1.

.	PC1(66.1%)	PC2 (10.9 %)
150.1/11.0 (55)	0.285	-0.431
150.1/10.5 (56)	0.245	0.431
494.3/38.0 (6752)	0.205	0.031
599.4/15.6 (12817)	0.168	0.122
325.2/25.1 (2865)	0.162	0.114
793.4/38.6 (22795)	0.160	-0.194
425.1/35.9 (4970)	0.158	0.095
536.3/38.4 (8588)	0.147	-0.013
1049.1/38.6 (24237)	0.139	-0.164
793.9/38.6 (22803)	0.134	-0.209
1048.6/38.9 (24235)	0.132	-0.150
907.6/45.0 (23968)	0.126	0.049
538.8/39.0 (8714)	0.115	-0.093
485.4/13.1 (6474)	0.115	-0.046
245.1/23.5 (893)	0.107	-0.013
478.3/37.1 (6262)	0.102	0.025
360.3/41.9 (3662)	0.100	-0.019
1304.2/38.9 (24317)	0.099	-0.027
151.1/11.0 (62)	0.099	-0.130
457.1/32.3 (5747)	0.098	-0.021
1303.7/38.8 (24316)	0.095	-0.078
611.2/38.8 (13646)	0.094	0.073
548.3/34.0 (9168)	0.093	0.011
652.4/43.9 (16593)	0.091	-0.010
254.2/17.0 (1071)	0.091	-0.145
321.1/17.4 (2759)	0.091	0.041
302.2/10.8 (2301)	0.086	-0.029
457.1/33.2 (5746)	0.084	-0.012
323.1/19.0 (2798)	0.083	-0.030
607.2/35.7 (13345)	0.083	-0.012
302.2/10.3 (2304)	0.082	-0.003
151.1/10.4 (63)	0.081	0.142
318.2/21.8 (2696)	0.080	-0.066
537.3/38.4 (8637)	0.080	-0.011
1559.4/39.0 (24354)	0.080	-0.043
508.3/34.1 (7277)	0.079	-0.148
254.2/17.5 (1074)	0.078	-0.035
734.5/44.8 (21016)	0.078	0.000
542.3/19.9 (8890)	0.077	-0.018
567.3/28.5 (10362)	0.077	0.002
326.2/25.1 (2886)	0.077	0.046
800.8/35.8 (22960)	0.076	0.026
297.2/44.3 (2188)	0.076	-0.126
800.3/35.7 (22944)	0.074	0.043
523.3/25.5 (7977)	0.074	-0.169
355.0/32.4 (3520)	0.073	-0.072
1559.9/39.0 (24355)	0.073	-0.070
607.7/35.9 (13377)	0.072	-0.001
505.3/19.6 (7158)	0.070	-0.021

Table S 5 – (cont.)

508.0/38.0 (7262)	0.070	0.007
434.8/28.8 (5223)	0.070	-0.034
178.2/19.2 (175)	0.069	0.001
1558.9/38.8 (24353)	0.068	-0.035
462.3/45.1 (5862)	0.068	0.009
583.3/39.2 (11586)	0.068	-0.025

Table S 6 - Loading values of each variable on the first two principal components. The contributions are sorted in descending order along PC1(Significant loadings >0.115)

Peak Name	PC1 (76.1 %)	PC2 (8.8 %)
793.4/38.6 (22795)	0.287	0.111
793.9/38.6 (22803)	0.258	0.129
684.4/44.0 (18535)	0.255	-0.082
1049.1/38.6 (24237)	0.225	0.142
425.1/35.9 (4970)	0.213	0.476
1048.6/38.9 (24235)	0.212	0.141
538.8/39.0 (8714)	0.204	-0.061
325.2/25.1 (2865)	0.202	-0.430
354.3/45.2 (3512)	0.180	0.153
409.3/21.1 (4686)	0.156	0.054
523.3/25.5 (7977)	0.144	0.225
245.1/23.5 (893)	0.144	0.255
1303.7/38.8 (24316)	0.136	0.088
1304.2/38.9 (24317)	0.130	0.063
487.1/35.8 (6519)	0.124	-0.016
312.1/18.7 (2541)	0.120	0.056
1559.9/39.0 (24355)	0.120	0.119
508.0/38.0 (7262)	0.118	0.006
492.3/35.3 (6687)	0.116	0.066
288.3/35.4 (1949)	0.115	0.110
384.0/18.9 (4143)	0.113	0.086
323.1/19.0 (2798)	0.112	0.089
415.1/35.9 (4790)	0.112	-0.019
792.9/38.8 (22788)	0.111	-0.020
624.2/18.9 (14619)	0.110	0.145
1074.5/39.2 (24258)	0.109	0.116
457.1/32.3 (5747)	0.109	0.173
268.1/26.1 (1431)	0.106	0.163
295.2/39.0 (2113)	0.105	-0.029
1559.4/39.0 (24354)	0.101	0.061
800.8/35.8 (22960)	0.096	-0.008
219.0/18.9 (517)	0.095	0.071
328.1/9.7 (2922)	0.094	0.130
1558.9/38.8 (24353)	0.091	0.054
514.3/25.5 (7588)	0.091	0.076
537.8/38.6 (8652)	0.089	-0.035
297.1/17.5 (2166)	0.087	0.084
598.3/26.7 (12734)	0.087	0.053
263.0/13.9 (1283)	0.087	0.052
1048.1/38.9 (24234)	0.085	0.060
241.0/13.9 (809)	0.085	0.103
792.4/38.9 (22774)	0.083	-0.027
326.1/9.5 (2878)	0.083	0.164
354.1/36.1 (3504)	0.083	0.041

Table S 6 – (cont.)

466.9/37.7 (5948)	0.081	0.076
660.2/9.5 (17101)	0.081	0.113
457.1/33.2 (5746)	0.081	0.106
414.2/37.6 (4778)	0.079	0.009
385.1/18.9 (4181)	0.078	-0.010
607.7/35.9 (13377)	0.077	-0.155
321.1/17.4 (2759)	0.075	0.055
355.0/33.2 (3519)	0.075	-0.052
282.3/39.4 (1762)	0.074	0.039
252.1/18.7 (1022)	0.073	0.036
486.1/36.1 (6492)	0.073	-0.014
629.2/18.7 (14973)	0.071	0.019
519.3/32.5 (7790)	0.068	0.080
212.2/27.5 (454)	0.064	0.032
378.2/18.3 (4033)	0.062	0.020
277.1/30.8 (1625)	0.061	0.158
326.1/33.1 (2879)	0.057	0.051
622.7/18.9 (14499)	0.056	-0.012
667.3/29.2 (17555)	0.055	0.036
1059.6/38.9 (24249)	0.051	0.098
747.5/39.1 (21525)	0.049	0.060
433.3/31.9 (5205)	0.047	0.029
555.3/38.6 (9544)	**0.047**	**-0.050**
268.2/21.0 (1447)	0.044	0.013
1891.1/38.8 (24414)	0.044	-0.020
195.1/29.5 (277)	0.037	0.010
527.2/34.5 (8174)	0.036	0.019
315.0/23.9 (2624)	0.036	0.030
713.4/32.6 (20076)	0.032	0.018
175.0/19.1 (157)	0.030	-0.013
584.3/39.3 (11665)	0.022	-0.053

Table S 7 - Loading values of each variable on the first two principal components. The contributions are sorted in descending order along PC1(Significant loadings >0.229)

Peak Name	PC1 (90.6 %)	PC2 (3.8 %)
793.4/38.6 (22795)	0.564	0.085
538.8/39.0 (8714)	0.402	0.110
1304.2/38.9 (24317)	0.256	0.141
508.0/38.0 (7262)	0.235	0.322
288.3/35.4 (1949)	0.227	-0.096
792.9/38.8 (22788)	0.222	0.145
384.0/18.9 (4143)	0.220	-0.113
415.1/35.9 (4790)	0.208	-0.212
800.8/35.8 (22960)	**0.181**	**-0.161**
1558.9/38.8 (24353)	0.180	0.068
1048.1/38.9 (24234)	0.171	0.182
792.4/38.9 (22774)	0.162	0.009
252.1/18.7 (1022)	0.145	0.029
321.1/17.4 (2759)	0.137	0.712
212.2/27.5 (454)	0.131	-0.402
622.7/18.9 (14499)	0.112	-0.031
1891.1/38.8 (24414)	0.085	0.051
195.1/29.5 (277)	0.072	-0.063
527.2/34.5 (8174)	0.067	-0.169

ii. Control vs Haloperidol

Table S 8 - Significant loading values (> 0.082) of each variable on the first two principal components. The contributions are sorted in descending order along PC1.

Peak Name	PC1 (66.2 %)	PC2 (13.5 %)
243.2/12.1 (938)	0.322	-0.398
554.3/45.2 (12197)	0.206	0.173
211.1/29.6 (483)	0.187	-0.231
243.7/12.1 (944)	0.174	-0.233
350.2/17.8 (3843)	0.170	0.116
548.3/28.7 (11780)	0.167	-0.103
907.6/45.0 (32401)	0.161	0.179
399.1/14.6 (5077)	0.157	0.094
339.3/45.1 (3593)	0.150	-0.001
482.3/45.1 (7845)	0.143	0.413
220.1/28.8 (588)	0.133	-0.025
485.4/13.4 (7972)	0.131	-0.154
793.4/38.8 (29646)	0.126	-0.135
130.2/13.9 (17)	0.124	-0.020
1049.1/38.9 (33225)	0.123	0.027
251.2/38.0 (1087)	0.122	-0.068
208.2/30.2 (446)	0.120	0.128
175.1/33.5 (174)	0.119	0.118
548.3/34.0 (11773)	0.118	-0.095
514.3/33.1 (9571)	0.117	0.064
214.2/30.2 (527)	0.117	0.074
1048.6/38.9 (33223)	0.116	-0.103
549.3/28.7 (11856)	0.115	0.021
508.3/33.3 (9174)	0.112	-0.055
622.4/43.7 (18317)	0.110	-0.014
212.2/27.3 (506)	0.109	-0.011
538.8/39.0 (11154)	0.107	-0.062
151.1/11.0 (65)	0.099	0.052
550.3/30.5 (11929)	0.098	-0.091
607.2/36.1 (16949)	0.097	0.094
281.2/35.1 (1915)	0.096	0.006
652.4/44.4 (20875)	0.095	-0.022
343.2/15.7 (3685)	**0.094**	**0.030**
656.4/35.6 (21239)	0.092	0.005
569.4/45.1 (13536)	0.092	0.101
1303.7/38.8 (33603)	0.090	-0.011
260.2/38.7 (1315)	0.088	-0.021

Table S 9 - Loading values of each variable on the first two principal components. The contributions are sorted in descending order along PC1(Significant loadings >0.243)

Peak Name	PC1 (50.9 %)	PC2 (20.4 %)
208.2/30.2 (446)	0.500	-0.365
260.2/38.7 (1315)	0.419	0.059
538.8/39.0 (11154)	0.377	-0.049
197.2/30.2 (344)	0.319	-0.137
258.1/34.9 (1275)	0.247	-0.043
277.1/13.4 (1805)	0.237	0.173
1304.2/38.9 (33604)	0.214	0.064
536.3/45.4 (10966)	0.197	0.006
456.3/14.8 (6845)	0.189	-0.147
510.3/30.7 (9296)	0.171	0.413
485.4/13.4 (7972)	0.157	0.592
294.1/35.9 (2305)	0.125	-0.152
594.2/18.9 (15766)	0.106	0.015
192.0/18.7 (265)	0.091	-0.115
843.3/32.2 (31373)	0.054	0.070
508.3/33.3 (9174)	0.007	0.376
218.1/11.5 (561)	0.002	-0.284

Table S 10 - Loading values of each variable on the first two principal components. The contributions are sorted in descending order along PC1 (Significant loadings >0.500).

Peak Name	PC1 (83.0 %)	PC2 (9.4 %)
538.8/39.0 (11154)	0.815	-0.559
1304.2/38.9 (33604)	0.500	0.608
192.0/18.7 (265)	0.289	0.477
843.3/32.2 (31373)	0.046	-0.302

iii. Control vs Clozapine

Table S 11 - Significant loading values (> 0.060) of each variable on the first two principal components. The contributions are sorted in descending order along PC1.

Peak Name	PC1 (71.8 %)	PC2 (11.4 %)
238.1/17.2 (713)	0.427	0.548
238.1/16.7 (712)	0.321	0.277
316.2/15.2 (2591)	0.253	0.135
316.2/14.7 (2596)	0.251	0.173
496.3/43.6 (6594)	0.190	0.242
300.2/15.6 (2192)	0.174	0.138
150.1/11.2 (54)	0.174	0.121
300.2/16.9 (2193)	0.162	-0.028
243.2/13.3 (809)	0.150	0.021
522.3/44.9 (7616)	0.137	-0.005
150.1/10.5 (53)	**0.128**	**0.066**
494.3/37.9 (6484)	0.115	-0.006
221.1/18.3 (517)	0.107	0.099
523.4/44.4 (7663)	0.103	0.055
300.7/15.6 (2204)	0.102	0.081
497.3/43.8 (6640)	0.100	0.129
300.7/16.9 (2205)	0.097	-0.023
350.2/17.8 (3294)	0.095	0.071
612.4/17.8 (13396)	0.090	0.050
285.1/17.1 (1777)	0.089	0.237
599.4/15.6 (12449)	0.088	0.071
285.1/17.6 (1778)	0.088	-0.255
306.7/17.8 (2336)	0.086	0.048
243.7/13.3 (818)	0.083	0.018
317.2/13.4 (2618)	0.078	0.063
178.2/17.7 (172)	0.077	0.002
536.3/38.3 (8265)	0.075	0.039
350.2/18.3 (3293)	0.075	0.007
317.2/13.9 (2617)	0.074	-0.007
224.1/16.3 (556)	0.070	-0.041
178.1/12.6 (164)	0.068	0.050
301.2/15.6 (2214)	0.065	0.045
599.4/16.9 (12451)	0.065	-0.017
793.4/38.8 (22143)	0.064	0.124
1049.1/38.9 (23574)	0.063	0.082
130.2/14.5 (13)	0.062	0.009
216.2/15.2 (460)	0.062	0.026
554.3/45.2 (9160)	0.060	-0.053

Table S 12 - Loading values of each variable on the first two principal components. The contributions are sorted in descending order along PC1 (Significant loadings >0.132).

Peak Name	PC1 (70.4 %)	PC2 (9.9 %)
480.3/41.1 (6063)	0.323	-0.382
793.4/38.8 (22143)	0.300	0.141
793.9/38.8 (22151)	0.279	0.112
1049.1/38.9 (23574)	0.241	0.169
1048.6/38.9 (23572)	0.228	0.141
453.2/18.3 (5394)	0.228	0.476
538.8/39.0 (8385)	0.225	0.068
224.1/16.3 (556)	0.216	-0.034
554.3/45.2 (9160)	0.185	0.038
541.3/19.9 (8503)	0.169	0.346
365.2/16.2 (3596)	0.160	0.322
409.2/17.3 (4472)	0.150	0.311
1303.7/38.8 (23650)	0.147	0.073
1304.2/38.9 (23651)	0.139	0.110
321.1/14.9 (2694)	0.133	0.231
508.0/38.0 (6993)	0.133	0.050
800.3/35.8 (22292)	0.119	-0.019
513.3/34.5 (7246)	0.110	0.016
1559.9/38.8 (23689)	0.108	0.130
1559.4/39.0 (23688)	0.108	0.052
487.1/36.1 (6256)	0.108	0.023
178.2/19.2 (170)	0.108	-0.091
607.2/35.8 (12991)	0.107	-0.001
588.4/44.1 (11594)	0.106	0.033
577.2/36.7 (10748)	0.103	-0.067
301.7/15.6 (2226)	0.102	-0.021
588.3/31.8 (11592)	0.102	0.022
282.3/39.7 (1680)	0.101	-0.044
800.8/35.8 (22309)	0.100	0.029
354.1/36.1 (3386)	0.099	0.067
600.4/15.6 (12510)	0.093	-0.126
660.2/9.5 (16696)	0.093	-0.040
1558.9/38.7 (23687)	0.092	0.046
792.4/38.9 (22123)	0.090	0.018
323.1/19.0 (2725)	0.089	0.111
457.1/33.2 (5503)	0.086	-0.044
292.2/15.6 (1966)	0.086	-0.010
283.3/41.8 (1723)	0.084	-0.019
294.2/14.4 (2016)	0.072	-0.095
473.1/33.3 (5865)	0.068	-0.014
277.1/13.5 (1550)	0.066	0.091
627.3/22.9 (14456)	0.064	-0.076
583.3/29.7 (11241)	0.064	0.098
294.1/35.9 (2008)	0.060	0.057
607.7/35.9 (13026)	0.060	-0.047
506.5/30.8 (6933)	0.060	0.052
486.1/36.1 (6230)	0.059	-0.013
542.3/38.8 (8560)	0.059	0.112
1814.0/39.0 (23720)	0.058	0.021
808.8/36.1 (22472)	0.056	-0.068
464.8/26.7 (5652)	0.050	0.052
493.2/39.1 (6445)	0.050	0.024

Table S 12 – (cont.)

790.4/23.4 (22079)	0.047	0.012
613.3/22.0 (13465)	0.042	-0.068
713.4/32.6 (19574)	0.040	0.021
550.6/31.4 (8956)	0.037	0.031
1891.1/38.9 (23747)	0.035	-0.004

Table S 13 - Loading values of each variable on the first two principal components. The contributions are sorted in descending order along PC1(Significant loadings >0.229).

Peak Name	PC1 (87.6 %)	PC2 (7.3 %)
793.4/38.8 (22143)	0.436	0.039
793.9/38.8 (22151)	0.403	0.079
1049.1/38.9 (23574)	0.354	0.164
1048.6/38.9 (23572)	0.336	0.157
538.8/39.0 (8385)	0.310	0.005
554.3/45.2 (9160)	0.252	0.849
1303.7/38.8 (23650)	0.216	0.109
1304.2/38.9 (23651)	**0.204**	**0.087**
1559.9/38.8 (23689)	0.167	0.219
1559.4/39.0 (23688)	0.158	0.122
513.3/34.5 (7246)	0.148	-0.109
487.1/36.1 (6256)	0.140	0.073
1558.9/38.7 (23687)	0.134	0.077
792.4/38.9 (22123)	0.127	-0.081
473.1/33.3 (5865)	0.093	-0.097
486.1/36.1 (6230)	0.085	-0.051
1814.0/39.0 (23720)	0.082	0.026
627.3/22.9 (14456)	0.078	-0.113
493.2/39.1 (6445)	0.065	-0.287