PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS E INGENIERÍA

"Chemical analysis of cuticular compounds in Alphitobius diaperinus via Chemical Ionization - Mass Spectrometry"

Tesis para obtener el título profesional de Licenciado en Química

AUTOR:

Aldo Gustavo Guimas Mazzetti

ASESOR:

Dra. Madina Mansurova

CO-ASSOR:

Dr. Carlos Gabriel Martel Gora

Lima, octubre, 2021

Al sensei Alejandro Fukusaki, por abrirme las puertas al mundo de la Química!

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Professor Madina Mansurova, for all the trust and confidence placed in me while carrying out this work, as well as the help and guidance provided.

The ICOBA Team, in particular Professor Alfredo Ibañez, Carlos Martel and Erika Calla, for their valuable help learning the techniques utilized in this study, as well as their unending patience.

My family, Giuliana, Gustavo, Pilar, Kantú, for their presence throughout the long process of not just this thesis but my career as a whole, their ever-present support and comprehension, their encouragement, their believe in me.

My friends throughout the career in chemistry, in particular Alvaro Castañeda and Alonso Donayre, for all the moments shared and the encouragement provided.

Ethel Ramos and Ricardo Bustamante, for all the help and clarity provided.

Finally, I want to express my gratitude to PROCIENCIA (formerly FONDECYT), for supporting the project “Identificación de semioquímicos y desarrollo de biotrampas contra escarabajos que son plagas agropecuarias, con énfasis en la avicultura” (proyecto 185-2020) granted to the ICOBA and the Pontificia Universidad Católica del Perú, under which my thesis project was funded.

ET LUX IN TENEBRIS LUCET
MCMXVII

SUMMARY

Alphitobius diaperinus represents a serious plague problem for the poultry industry. The possibility of utilizing biotraps for control of *A. diaperinus* is promising, but still merits more research, in particular with regards to the identification of compounds for their possible use in biotraps for *A. diaperinus*. Herein, data from a GC-MS experiment carried out on a cuticular extract of *A. diaperinus* is statistically analyzed, preparing the data for analysis and carrying out data filtration by a one-way ANOVA (p-value < 0.01) and finally data visualization by PCA. Three compounds of interest are successfully identified: 2-methyl-1,4-benqouinone, 2-ethyl-1,4-benzoquinone and 2-ethyl-1,4-hydroquinone.

INDEX

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

- ***A. diaperinus***: *Alphitobius diaperinus*
- **Alk:** Sample identifier, alkane mix.
- **ANOVA**: Analysis of variance.
- **BK**: Sample identifier, blank.
- **C16**: Hexadecane
- **CI**: Chemical Ionization
- **EBQ**: 2-ethyl-1,4-benzoquinone
- **EI**: Electron Ionization
- **F-exp:** Sample identifier, experienced females.
- **Fv:** Sample identifier, virgin females.
- **GC**: Gas Chromatography
- **Hex:** Sample identifier, hexane.
- **LC**: Liquid Chromatography
- **M-exp**: Sample identifier: experienced males,
- **MBQ**: 2-methyl-1,4-benzoquinone
- **MS**: Mass Spectrometry
- **Mv**: Sample identifier: virgin males.
- **NaOEt**: Sodium ethoxide.
- **PCA**: Principal Component Analysis
- **PC#**: Principal Component of number #.
- **RI**: Retention index.
- **TIC**: Total ion current.
- **TOF**: Time of Flight.

1. INTRODUCTION

A. diaperinus is a beetle and a plague for the poultry industry, which represents a serious problem. Severe economic losses have been attributed to this beetle, resulting from structural damage, as well as higher mortality rates of birds [1,2].

There is a demand for finding different control methods for *A. diaperinus,* since all the current working methods for controlling this pest present drawbacks. Chemical control methods can potentially result in environmental concerns. Physical control methods often result in stress for birds, leading to economic losses. And finally, the current alternative for biological control methods are so far not statistically reliable, presenting irreproducibility between replicate samples [2].

Therefore, a different method for controlling *A. diaperinus* is proposed, whereby biotraps loaded with pheromones are utilized in a push-pull system to remove the beetles from the poultry farm. The advantages of this method would result in a high degree of beetle capture, as well as little environmental impact.

Due to the potentially high cost of utilizing different mixtures of compounds as pheromones, it becomes essential to correctly identify which specific compounds can elicit a response (e.g., as an attraction or as an alarm pheromone) in *A. diaperinus*. Correct identification of compounds would allow to optimize this method, and reduce the cost.

This study aims to identify cuticular compounds of *A. diaperinus* which may potentially be attraction or alarm pheromones, to be used as active compounds in biotraps, as well as propose a possible potential synthesis for such compounds.

2. OBJECTIVES

General Objective:

Carry out data-mining of GC-MS (CI) data from cuticular extracts of *A. diaperinus*

Specific Objectives:

1. Analyze MS-CI data of cuticular extracts from *A. diaperinus*.
2. Identify the selected compounds of interest
3. Propose a chemical synthesis for target compounds.

3. THEORETICAL FRAMEWORK

3.1 *Alphitobius diaperinus:* A PROBLEMATIC SITUATION

Alphitobius diaperinus (*A. diaperinus*) is a beetle belonging to the Tenebrionidae family, commonly known as the "lesser mealworm beetle". Its taxonomic classification is shown in Table 1 [1]. This plague is of great interest to the poultry industry due to the severe economic damage it causes [2]. Apart from generating structural damage, *A. diaperinus* is also associated with the ability to transmit disease agents such as bacteria of the genera *Escherichia*, *Salmonella*, and *Campylobacter* [3–5], viruses such as Fowl Pox and Newcastle [6], fungi of the genera *Aspergillus*, *Penicillium* and *Candida* [7], and protozoans such as *Eimeria* (Coccidiosis) [8]. *A. diaperinus* also acts as a vector of cecal worms and avian tapeworms [9]. Therefore, controlling this pest in poultry farms is of extreme importance.

Table 1. Taxonomic classification of *Alphitobius diaperinus* (left) and individual (right). Table adapted from reference [1]. Image is taken from reference [10].

Category	**Classification**	**Individual**
Kingdom	Animalia	
Subkingdom	Bilateralia	
Infrakingdom	Protostomia	
Superphylum	Ecdysozoa	
Phylum	Arthropoda	
Subphylum	Hexapoda	
Class	Insecta	
Subclass	Pterygota	
Infraclass	Neoptera	
Superorder	Holometabola	
Order	Coleoptera	
Suborder	Polyphaga	
Infraorder	Cucujiformia	
Superfamily	Tenebrionoidea	
Family	Tenebrionidae	
Genus	*Alphitobius*	
Species	*Alphitobius diaperinus*	

There are several methods developed for managing and controlling this pest. These methods can be divided into chemical, physical, and biological, based on their operation mode. In terms of mortality alone, the most effective control methods are the chemical ones, based on the use of insecticides [2]. However, though chemical methods are efficient, there are problems associated with their use, especially their toxicity towards other species. Pesticides pose serious health and environmental hazard, having a severely damaging effect on ecosystems, and as such, alternatives to pesticides are continually being sought [11].

An alternative to the chemical control of this pest is the use of physical control methods. These are based on changing the humidity and temperature conditions of farms [2] and are highly efficient, achieving very high mortality rates [12]. The main issue associated with these methods is economic: the stress-induced in chickens during the manipulation of farm conditions diminishes their growth potential [13].

The third alternative for controlling this pest is with biological control methods. These methods are based on using biological species that are specifically pathogenic towards *A. diaperinus*. However, most methods show high irreproducibility or a delay in terms of when the insects' mortality rate becomes evident [2]. In summary, the advantages and disadvantages of each of these methods are shown in Table 2.

Table 2. Advantages and disadvantages of methods of control for *Alphitobius diaperinus*

Method of Control	**Advantages**	**Disadvantages**
Chemical	High mortality rates.	High impact on the environment due to the use of insecticides.
Physical	High mortality rate and low impact on the environment.	High cost associated with them.
Biological	Low impact on the environment.	Irreproducibility, some of them present low mortality rates.

It is clear, that chemical control methods are the most efficient. Furthermore, if the compounds' toxicity can be reduced (or eliminated), it would make them the method of choice for the poultry industry.

A possible alternative to traditional chemical pesticides is the use of natural compounds such as pheromones. Pheromones can be used in the poultry farms in devices called biotraps (Figure 1).

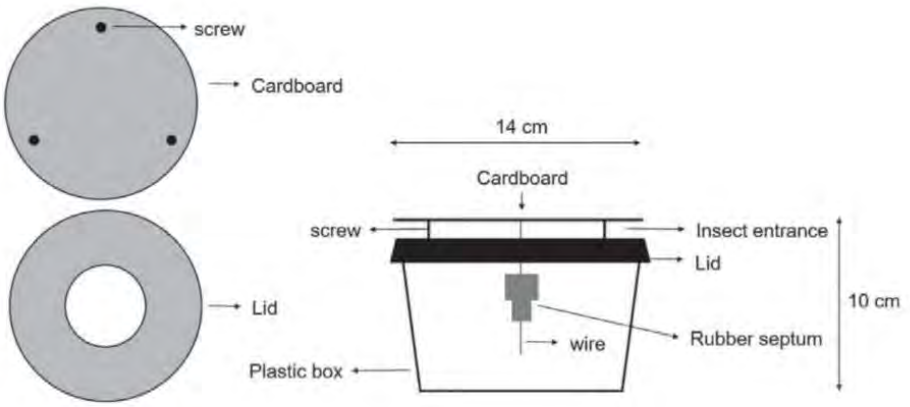
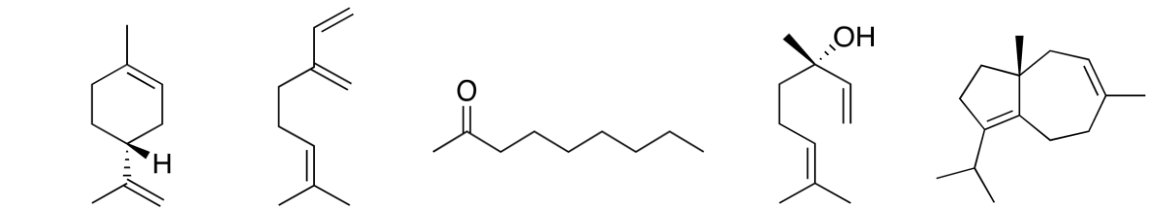


Figure 1. General Scheme for a Biotrap. Taken from reference [14].

Pheromones are semiochemicals. A semiochemical is an organic compound used by insects to convey specific chemical messages. So, pheromones are emitted by members of the species to modify the behavior or physiology of other individuals of the same species. Pheromones can be volatile and non-volatile organic compounds; selected examples of pheromones are shown in Figure 2. They can be classified by their function: (i) aggregation, (ii) alarm, (iii) aphrodisiac and anti-aphrodisiac (*i.e.*, sexual behavior), and (iv) social hierarchy.



(R)- (+)-limonene (R)-β-ocimene 2-nonanone (S)- (+)-linalool (R)- (+)-daucene

Figure 2. Aggregation pheromones for *Alphitobius diaperinus.* Taken from reference [15].

By using a combination of pheromones, *e.g.*, alarm (repelling effect) and aggregation (attraction effect), insects can be herded out from the poultry farms to specific locations (biotraps). Subsequently, the insects can be removed alongside the traps shown by Hassemer *et al*. (Figure 3) [14]. The use of aggregation pheromones alone utilizes the "pull-effect", or the attraction effect of the pheromones to lure insects into the traps, loaded with aggregation pheromones. On the other hand, aggregation pheromones can be combined with alarm pheromones ("push-pull" effect) to direct insects towards the traps. In this scenario, biotraps loaded with alarm pheromones will drive insects away from their location, directing them towards biotraps loaded with attraction pheromones, which will drive insects towards their location (hence, a "push-pull" effect).

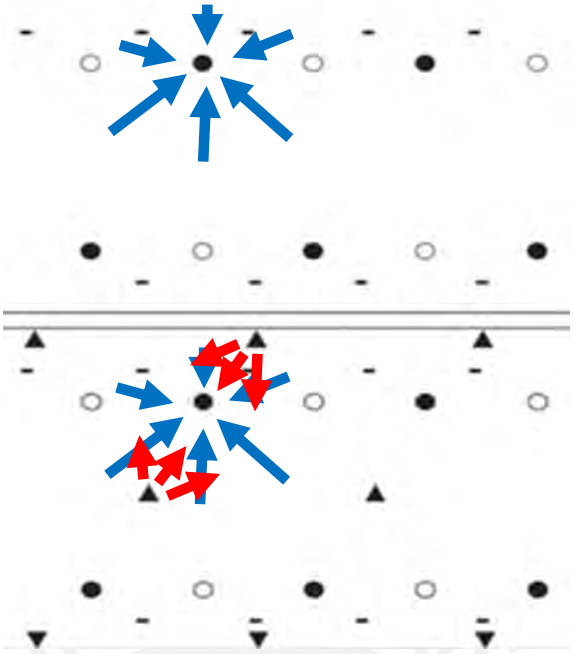


Figure 3. Comparison of layouts for deploying biotraps: Top image shows a push system, lower image shows a push-pull system. The legend is as follows: Black circles indicate a biotrap loaded with attraction pheromones, white circles indicate a control trap, triangles indicate biotraps loaded with alarm pheromones, lines indicate sampling sites. The arrows overlayed on the image indicate the movement of insects: Blue arrows indicate movement of insects due to an aggregation pheromone (*i.e.*, pull effect). Red arrows indicate the movement of insects due to an alarm pheromone (*i.e.*, push effect). Adapted from [14].

It is important to note that push-pull effects worked better than pull effect alone. The pull effect (aggregation pheromones) was described by Hassemer *et al*. [15] during a previous study on a population of *A. diaperinus* in Brazil. The following six aggregation pheromones were: (R)-limonene, (R)- β-ocimene, 2-nonanone, (S)-linalool, (R)-daucene (Figure 2). At the same time, the push effect (alarm pheromones) can vary depending on the location of the species. For example, Hassemer *et al.* [15] identified as alarm pheromones a mix of 1,4-benzoquinone, 2-methyl-1,4-benzoquinone, and 2-ethyl-1,4-benzoquinone. While, Baran *et al*. in a different study [16] also identified the following compounds as repellents for *A. diaperinus*: formic acid, acetic acid, propionic acid, and butyric acid.

In summary, to deploy successful biotraps, it is necessary to find the correct combination of attraction pheromones, as well as repellent compounds. Additionally, there is a demand for the analysis and understanding of *A. diaperinus* pheromones for improving the deployment of biotraps in poultry farms. Continuing previous work, this thesis is devoted to identifying further compounds relevant to the mating behavior in *A. diaperinus.* And having those identified compounds of interest, this thesis will attempt to propose synthetic methodologies for potentially biologically active compounds that are not commercially available.

3.2 MASS SPECTROMETRY AS AN ANALYTICAL TECHNIQUE

Mass spectrometry (MS) is an analytical technique which allows to obtain a spectrum of masses of molecules and fragments of such molecules present in a given sample. It is especially useful for metabolomic studies and identification of compounds within a large sample matrix [17].

The molecules are ionized via different methods, which will sometimes involve fragmentation, and the resulting ions are then sequentially analyzed, determining their charge to mass ratio.

MS is coupled to a separation technique, usually gas chromatography (GC) or liquid chromatography (LC). As such, analytes will enter the mass spectrometer at different points in time; whenever this happens all analytes will be ionized, sometimes fragment, and a mass spectrum will be generated, for each point in time. Thus, multiple mass spectrums will result from an analysis of a given sample via GC-MS or LC-MS.

A 3D spectrum will be obtained, where one can find the chromatogram and, for each time cut of the chromatogram, a corresponding mass spectrum [18].

3.3 SEPARATION TECHNIQUE: GC

Gas chromatography (GC) is a separation technique whereby a mixture of compounds is injected in gas phase into a chromatography column. A carrier, inert gas will then ensure the mixture of compounds to travel along the column. The column is coated with a stationary phase, which varies according to the type of column selected; there is a very thin film of this phase coating all sides of the column. As the analytes travel down the column, they will interact, at varying degrees, with this stationary phase, according to their polarity. Compounds which present a stronger affinity for the stationary phase will be further retained in the process of chromatographic separation, whereas compounds with a weaker affinity for the stationary phase will be less retained. This will lead to compounds separating, and thus eluting off the column at different points in time [19].

3.4 IONIZATION TECHNIQUES

3.4.1 ELECTRON IONIZATION

In an electron ionization technique, the ionization chamber contains a source of electrons, usually a heated metal filament, and an electron trap. A "flow of high energy electrons" (about 70 eV) [18] is thus established within the ionization chamber. When the sample is injected, analyte molecules will collide with this "electron stream" (Figure 4).

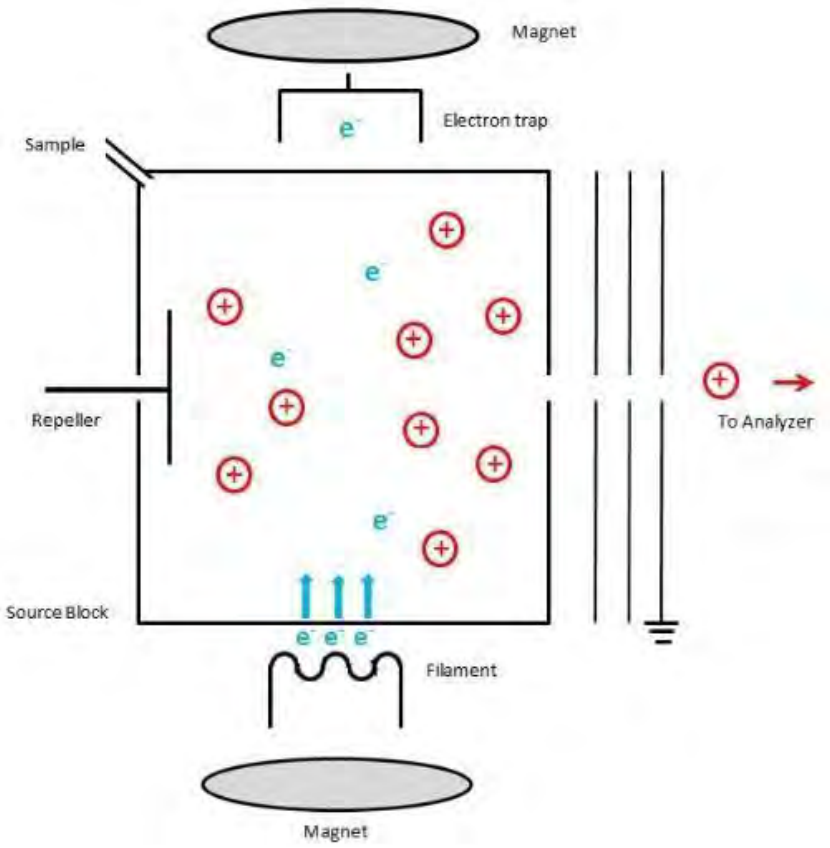


Figure 4. Electron ionization technique. Taken from reference [18]

The impact will result in a charged radical being formed, with the abstraction of one electron from the analyte. As such, the following equation describes this process:

$$\boldsymbol{Analyte} + \boldsymbol{e^-} \rightarrow \boldsymbol{Analyte^{+\cdot}} + \boldsymbol{2\,e^-}$$

In the situation described above, the ion formed is termed the "molecular ion", it corresponds to the original analyte molecule with a charge. Its mass, having only lost one

electron, will be virtually identical to the original analyte's mass. However, due to the high energy of electrons which will be impacting the injected analyte molecules, the ionization process can sometimes result in the breaking of a chemical bond, and the formation of two (or more) fragments. Two results can thus occur from ionization via EI: a non-dissociative result, producing the molecular ion or a dissociative result, producing two or more different fragments (Figure 5).

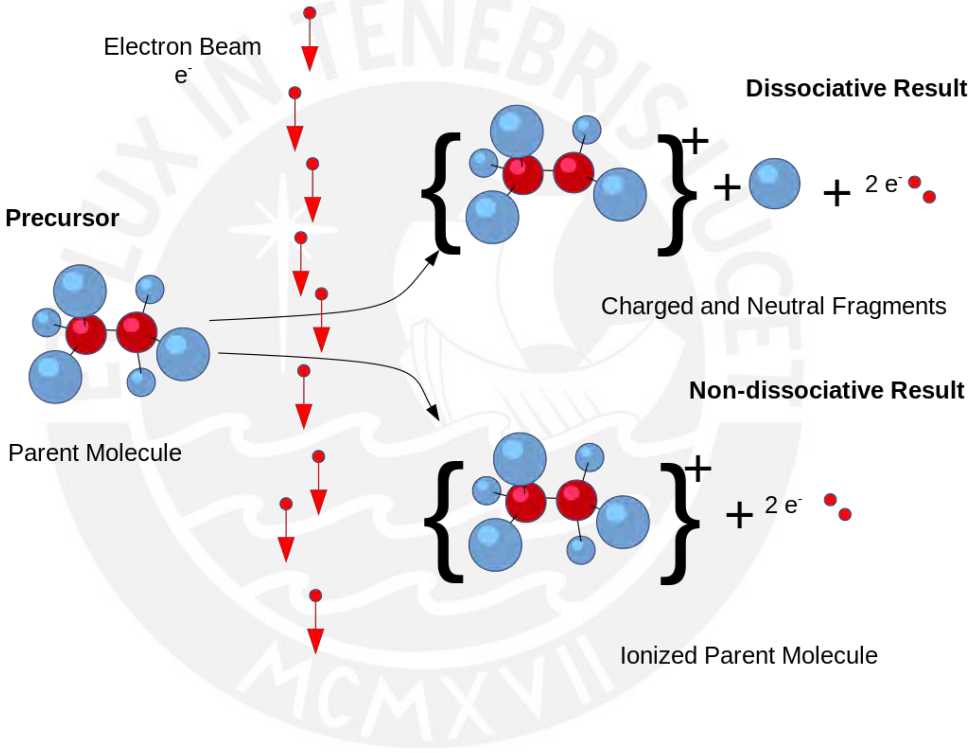


Figure 5. Fragmentation in electron ionization. Taken from reference [20]

Fragmentation in EI is complex. Once a radical is formed, several different rearrangements can occur resulting in different fragments being formed.

The combination of high energy electron-impacts and multiple possible rearrangement routes result in complex mass spectra originating from EI. In practice, multiple fragments are formed, to the point where sometimes the molecular ion will not be present at all within these fragments. The presence of fragmentation is what classifies EI as a "hard"

ionization technique, as opposed to “soft” ionization techniques, where fragmentation is reduced by comparison to EI [21].

3.4.2 CHEMICAL IONIZATION

In chemical ionization (CI) the method of electron ionization described beforehand is used to ionize molecules of a given reagent gas. This reagent gas can be ammonia, methane, ethane amongst others. The ionization of the reagent gas forms multiple ions from the gas molecules, these ions are termed "reagent ions". Upon injection of the sample to the ionization chamber, the reagent ions will collide with the analyte molecules, producing ionization [22]. It is important to note that this ionization is of lower energy than EI, as no direct collision with electrons is occurring with the analyte molecules. Instead, lower energy collisions between the reagent gas and the analyte results in a "soft" ionization of the analyte, with little to no fragmentation (Figure 6)

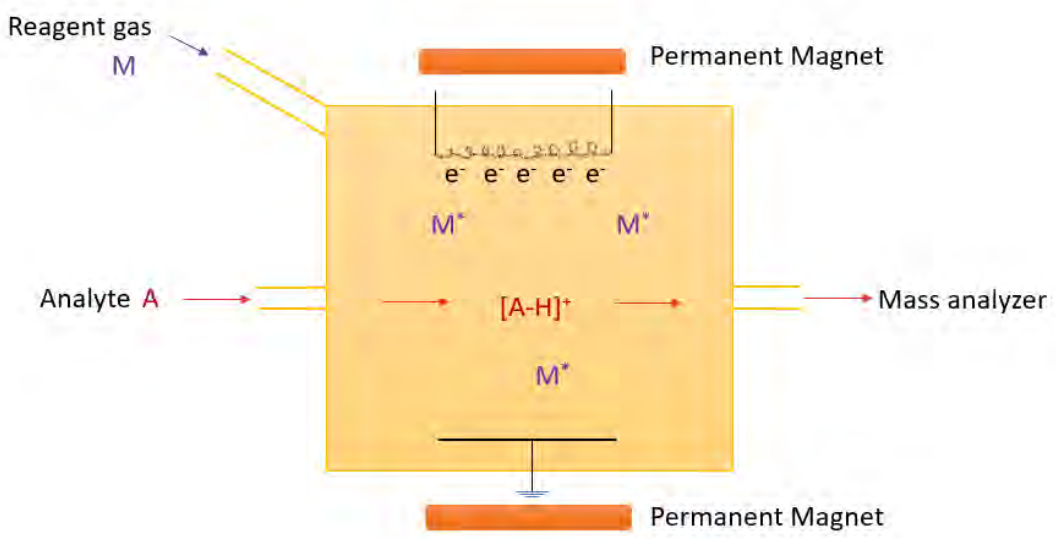


Figure 6. Chemical Ionization Technique. Taken from reference [23]

Adducts are formed, however, the most typical of which is $(\boldsymbol{Analyte} + \boldsymbol{H})^+$, the addition of a proton resulting from an ionization collision. This adduct is commonly referred to as $(\boldsymbol{M} + \boldsymbol{H})^+$. Multiple different adducts can be potentially formed between the reagent gas ions and the analyte molecule, these adducts will present a mass-to-charge ratio greater than that of the molecular ion.

Some degree of fragmentation, however, is also possible. This usually results when the formed adduct is unstable, and a mechanistic pathway is available to the lower energy of the product ions.

It is important to note, however, that for CI to be an efficient ionization technique, the analyte must have functional groups which will readily react with the ions produced from the reagent gas. Functional groups which can accept a proton are an excellent example of this. On the other hand, molecules such as alkanes lack readily ionizable functional groups, and will result in much less efficient ionization when CI is used. Given a particular sample consisting of a mixture of compounds, while EI will readily ionize nearly all of the compounds present, it is expected that CI might fail to ionize some of them. However, CI will usually provide more useful information with regards to the molecular ion (easily identifiable from the M+H fragment), making it still a valuable analytical technique [22].

3.5 ION DETECTION: TIME OF FLIGHT

The selected ion detection method was time-of-flight mass spectrometry (TOF-MS). In TOF-MS, a group of ions is accelerated through a flight tube, towards a detector. The ion's charge ensures that they follow a flight path that allows them to reach the detector. The mass is determined based on the time the ion spends "flying" through the flight tube (hence, time-of-flight). Larger ions will move slowly, and have a longer time-of-flight than smaller ions.

The main advantage of TOF-MS is that it can analyze ions of several different mass-to-charge ratios at the same time. This is ideal when a full-spectrum scan is needed, such as in our case where we wish to ensure detection of every compound present. This detection is made possible by having carried out previous analysis on samples of linear alkanes of increasing molecular weight. Simultaneous detection of numerous chemical species is of key importance in our work [24].

3.6 GAS CHROMATOGRAPHY – MASS SPECTROMETRY (GC-MS)

In GC-MS, gas chromatography is coupled to mass spectrometry. Molecules exiting the chromatograph will then be inserted into the mass spectrometer, allowing for separate analysis of each compound.

Spectra obtained from a MS experiment are in three dimensions, corresponding to the chromatogram on two axes and the mass spectra for each given time on two axes, with the intensity axis being shared. When multiple chromatography-MS experiments are run on samples which might contain the same analytes, it is important to first align the chromatograms. The chromatogram peak for one analyte for instance, might be further displaced in one chromatogram with comparison to the other. If this is the case, and the data is taken as such, this analyte will be interpreted as being two different analytes, at different retention times [25].

Programs such as MS-DIAL will align the chromatograms via different mathematical techniques to obtain aligned data, where chromatogram peaks deemed to belong to the same analyte due to their similarity in shapes and retention times are merged together. The same happens for similar mass fragments. This is a necessary step for analysis of multiple mass spectra; however it must be taken into account that this alignment has its limitations, and different entered parameters will, sometimes, lead to some loss in information from the mass spectra [26].

Once the mass spectra have been aligned, different statistical techniques can be applied to filter out the compounds which are deemed to be unimportant for the experiment in question. These compounds will include those which are present in insignificant amounts, those which are more closely correlated with the blank samples, and those which do not strongly correlate with the desired samples for analysis. The different statistical techniques will be expanded upon in a subsequent chapter.

3.7 PRINCIPAL COMPONENT ANALYSIS AND P-ANOVA FILTERING

Given a particular data set, one can generate a series of points in multi-dimensional space. Each axis represents a variable, the distance along the axis at which a point is found represents the value for such a variable, each point represents data. In our case, the data set consists of multiple intensities of different mass fragments.

In Principal Component Analysis (PCA), a linear fit is carried out onto this series of points. The data is then projected onto this linear fit, generating a score value for each data set, corresponding to the value of the projection of the data point onto the linear fit. The linear fit is called the First Principal Component or PC1, each data point has a score on PC1. The unit-vector corresponding to the PC1 line is termed the “Eigenvector”. The “Eigenvalue” is the sum of squares of projections of data onto PC1.

Subsequently, a new linear fit is carried out on the data, which must be orthogonal to PC1 in the nth-dimensional space. This second linear fit is called the Second Principal Component or PC2, and once again, each data point has a score value for PC2. Multiple linear fits are carried out subsequently, giving a final number of Principal Components equivalent to n-1, where n is the dimension of the nth-dimensional space [27].

Each Principal Component describes variation in the data to a certain extent. A PCA scores plot would be a plot of the PCA scores, the scores each data point has for each principal component. If only the first two PCA scores are graphed, this can show variation in the data to a certain extent. The extent to which it represents variation in the data will be determined by a Scree Plot, which shows what percentage of the variation of the data each Principal Component accounts for.

It has been mentioned that Principal Components account for the variation within a data

set, they account for the varied influences or loadings the original variables have on the variation of the data set. Such influences or loadings can be graphed: Each Principal Component will have a corresponding Eigenvector. The Eigenvector will be a linear combination of all the variables present in the dataset, with different coefficients or loadings for each variable. As such, each Principal Component will have corresponding loadings for each variable present. Graphing out these loadings will produce a PCA loadings plot [27].

Before PCA is carried out, it is important to ensure that the data is normalized. This is done by taking the logarithm of all the values for the mass intensities. The result will be a normalized data set.

Finally, once the data has been normalized, the p-ANOVA of the data is taken. Anova represents an "Analysis of Variance". This will indicate which variables have a significant variance so as to be included in the study. A value of p-ANOVA greater than 0.05 is chosen, based on previous work [28].

4. EXPERIMENTAL PART

4.1. MATERIALS AND EQUIPMENT

Analytical Instrument:

- Agilent 7250 GC/Q-TOF at University of California Davis (CA, USA)

Computer specifications:

- **Processor:** Intel® Core™ i7-4790 CPU @ 3.60 GHz (8 CPUs) ~ 3.6 GHz
- **RAM:** 32,768 GB.

Materials used for extraction procedure:

- 4mL, 2 mL vials.
- 100 µL, 500 µL syringes.
- 9x9 vial storage box.
- ParaFilm ™.
- Plastic Petri dishes, 6 cm diameter.

Equipment used for extraction procedure:

- Growth chamber (Memmert HPP750, Memmert GmbH, Schwabach, Germany)
- Ultrafreezer (-20 ℃)

Reagents used for extraction procedure:

- Hexane (HPLC grade, purchased from Sigma - Aldrich).
- Nitrogen gas.

4.2. METHODOLOGY

A flow diagram for the followed methodology is shown in Figure 7.

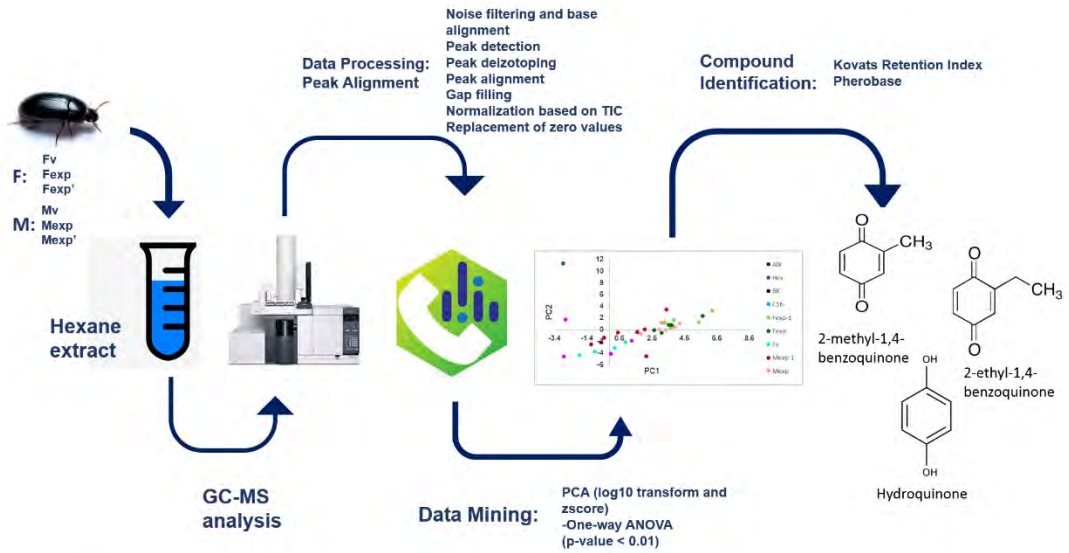


Figure 7. Methodology flow diagram

4.1.1 INSECTS

Work with insects (both insect rearing and sample extraction) was done in our group (ICOBA – PUCP) by Erika Calla Quispe.

In brief: Adult *A. diaperinus* beetles were obtained from a farm in Lima – Perú, they were reared at ICOBA – PUCP in glass boxes, with food and water provided. Individuals were isolated in glass boxes by sexual conditions (virgins vs. sexually experienced) and gender (male vs. female). Glass boxes were kept in an environmentally controlled climate chamber (Memmert HPP750, Memmert GmbH, Schwabach, Germany) at a constant temperature of 30 °C and humidity of 50 %, with a photoperiod of 12:12 h (light : dark). Adults being 21-d-old or older were used to guarantee sexual maturity.

4.1.2 CUTICULAR EXTRACTION AND CHEMICAL ANALYSIS

For the extraction process, 12 individuals belonging to each group (experienced males, experienced females, virgin males, virgin females, Table 3) were killed by freezing. Extraction was then carried out by immersion in 2 mL hexane (HPLC grade, purchased from Sigma - Aldrich). 1 µL of 1000 ppm hexadecane in hexane (C16, MS grade, purchased from Sigma-Aldrich) was added as an internal standard to all samples. Then, the cuticular extracts were concentrated to 120 µL using a gentle stream of nitrogen. Samples were stored at -20 ºC for subsequent chemical analysis.

The identities for each sample are shown in Table 3. These are the abbreviated names used in the present study.

Table 3. Identifiers for each sample utilized in the present study

Sample	Identifier
Alkane Mix	Alk
Hexane	Hex
Blank	BK (Hexane)
Hexadecane standard	C16
Experienced females, group 1	F-exp1
Experienced females, group 2	F-exp
Virgin females	Fv
Experienced males, group 1	M-exp1
Experienced males, group 2	M-exp
Virgin males	Mv

Cuticular extracts were analyzed by gas chromatography coupled to quadrupole time-of-flight mass spectrometry Agilent 7250 GC/Q-TOF at University of California Davis (CA, USA) with chemical ionization. A DB-5 column was used for gas chromatography. 1 µL aliquots of cuticular extract were injected at 50 ºC for 2 minutes, the temperature was increased at a rate of 10 ºC / minute until reaching 250 ºC and it was held there for 20 minutes. Then, splitless-mode injections were made at an injector temperature of 250 ºC,

and detector temperature of 270 ºC. Helium was used as the carrier gas, at a flowrate of 1.5 mL / min.

4.1.3 DATA PROCESSING

Processing of data was carried out using the MS-DIAL software (ver 4.48), which was utilized to carry out peak alignments and normalization of the GC-MS (CI) data. Normalization was based on total ion current (TIC).

The MS data was converted into .abf format, which is accepted by MS-DIAL. Once this was done, the data was uploaded into MS-DIAL v4.48, and chromatogram alignment was carried out, as well as sample normalization based on TIC (total ion chromatogram). This gave as a result an unfiltered table of multiple mass fragments, totaling over 400 features.

Filtering was now applied. For this step, the base ten logarithm of the intensities of each mass fragment was taken, following this Principal Component Analysis (PCA) was carried out with MatLab™ R2021. PCA was utilized to track data quality, reduce the data dimensionality, identify potential outliers in the dataset, as well as to identify sample clusters.

The ANOVA values were also taken. The chosen p-value for the filter was 0.01. Any mass fragments whose p-value was lower than 0.01 were kept, any mass fragments whose p-value was greater than 0.01 were rejected. This filtration allowed to reduce the least amount of compounds. Additionally, fragments which had a concentration in samples which was at least three times greater than the concentration of the same fragment in the blank sample and at least three times greater than the concentration of the same fragment in the standard sample were kept, all others were rejected. These two filtration criteria allowed for a significant reduction of the sample size (see section 4.3).

PCA score plots and loading plots were obtained, as well as a Scree plot (plot of

percentage of variance explained by each principal component) and a “volcano” plot (plot of p-ANOVA vs. PC1) to confirm the correct filtration of samples.

For compound identification, their Kovats retention index was calculated, following this a database search was carried out in the pheromone compound database PheroBase™. Compounds within the database with a +/- 20 match in RI and a +/- 0.05 match in m/z fragments were selected as positive potential identifications for our list of fragments. Following this, a comparison with compounds previously reported in the literature [32] allowed for correct identification of our mass fragments.

4.3. RESULTS

4.3.1 PROCESSING RAW DATA

Data extracted from the mass spectrometer was aligned in MS-DIAL. For this analysis, all zero values within the dataset were replaced by a fraction (one tenth) of the minimal peak value. The retention time tolerance for identification was set to 10 minutes, the accurate mass tolerance was set to 0.01 Da. The minimum peak height for peak detection was set to an amplitude of 10000.

4.3.2 NORMALIZATION AND Z-SCORE

The data obtained from mass spectrometry was aligned in MS-DIAL. The raw aligned data did not possess a normal distribution: as such, the logarithm of the raw aligned data is taken. The distribution of the data thus changes to a normal distribution, now suitable for carrying out PCA. Furthermore, the z-score for the data is also taken, to ensure further analysis is accurate.

4.3.3 DATA ANALYSIS: ONE-WAY ANOVA AND PCA

An initial PCA is carried out to better visualize the data. The scores plot of this initial PCA is shown in Figure 8.

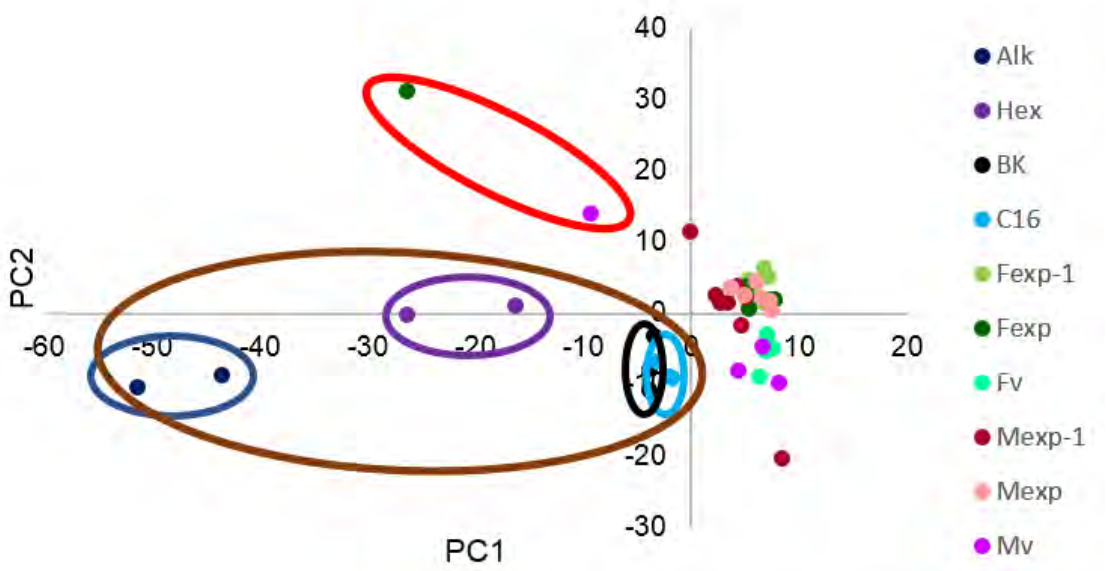


Figure 8. Initial PCA on dataset

The blue circle contains the alkane mix samples, the purple circle contains the hexane standards, the black circle contains the blanks, the cyan circle contains the C16 standards. Notice how each of these samples group together in the represented circles and also group together within the larger brown circle, away from most of the data to the right. This initial PCA has already succeeded, to a large extent, to represent an adequate visualization of data. However, two samples (shown within the red circle, one for 'Fexp' and one for 'Mv' do not group with the rest of the samples. Additionally, the loadings plot (Figure 9) shows how each sample is being affected by several different mass fragments, including potentially mass fragments which are not relevant to the study (compounds present in blank, etc.).

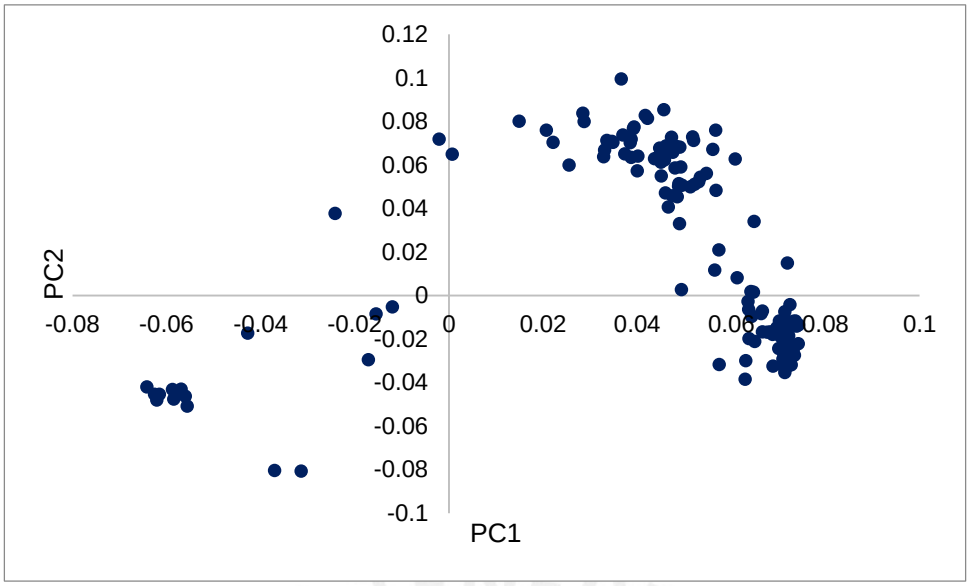


Figure 9. Loadings plot of first PCA

The Scree plot (Figure 10) corresponding to this PCA shows that two principal components alone only account for around 55% of the variation in the data, with the third and fourth principal components accounting for an additional 9.77% and 7.13% of variation respectively.

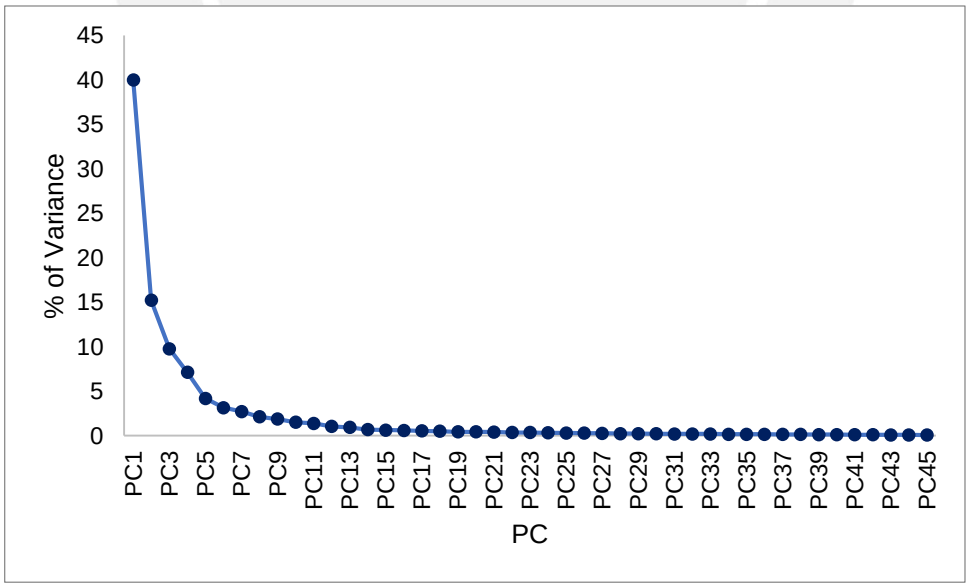


Figure 10. Scree plot for first PCA

As such, the data is filtered, by a $-\log_{10}$ p-ANOVA ≥ 5, shown in the volcano plot in Figure 11, and additionally by a conditional filter whereby a mass fragment passes the

filter if and only if its concentration is greater in one of the experimental samples than its concentration in the blank.

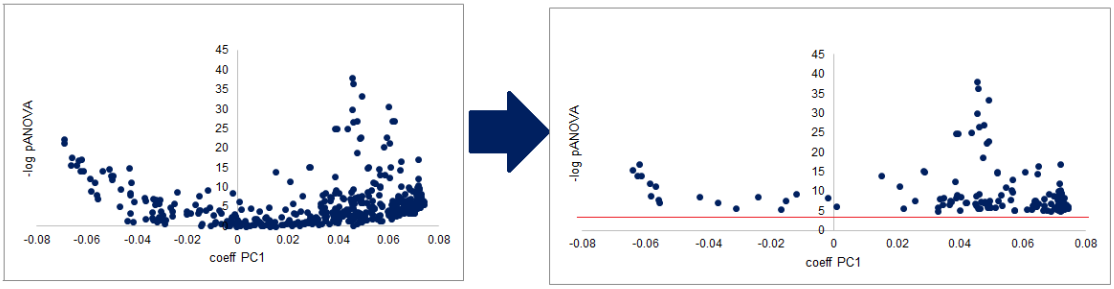


Figure 11. Filtration by ($-\log_{10}$ p-ANOVA $\geq$ **5**), notice how the number of data points is diminished. Cut-off line shown in red.

After filtration, a new PCA is carried out on the data. The new PCA scores plot is shown in Figure 12.

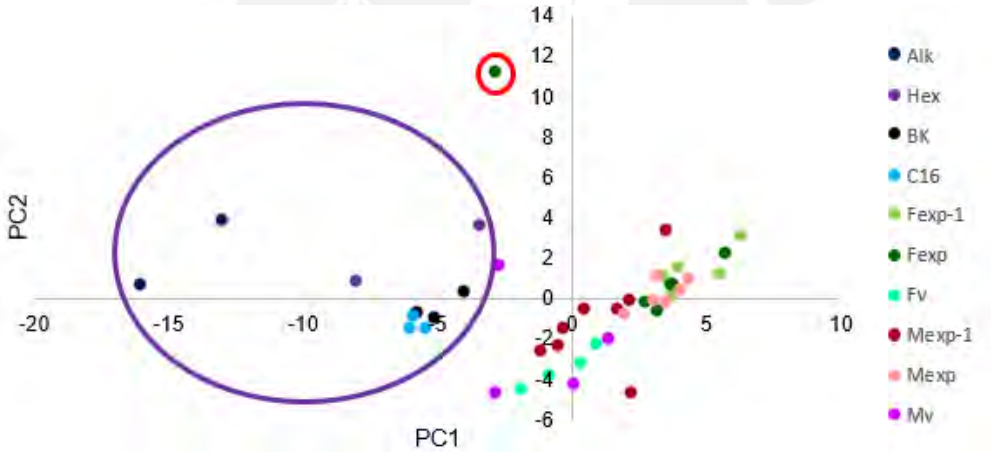


Figure 12. PCA scores plot for second PCA

Notice how all of the non-experimental samples cluster together in the purple circle. This time, only one of the experimental samples (within the red circle, one for Fexp) does not cluster with the rest of the experimental samples. The graph is zoomed in Figure 13 on the region of interest for the experimental samples.

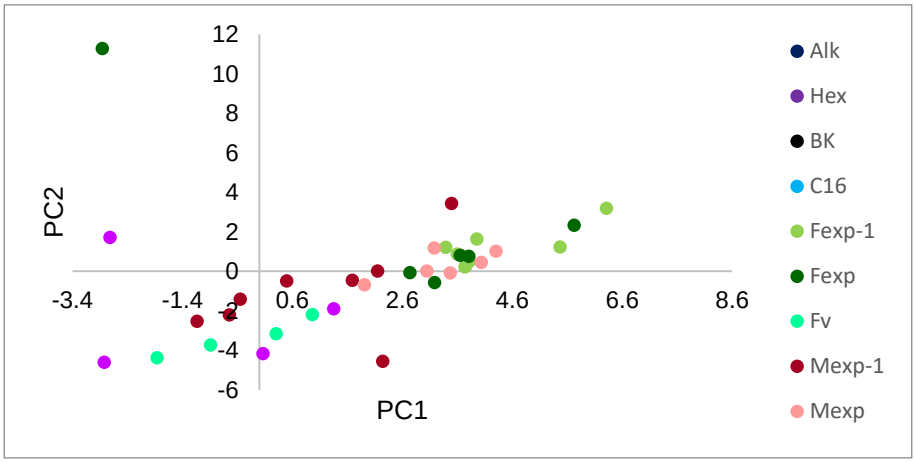


Figure 13. PCA scores plot for second PCA, region of interest is zoomed in

Additionally, the new loadings plot for the second PCA is shown in Figure 14.

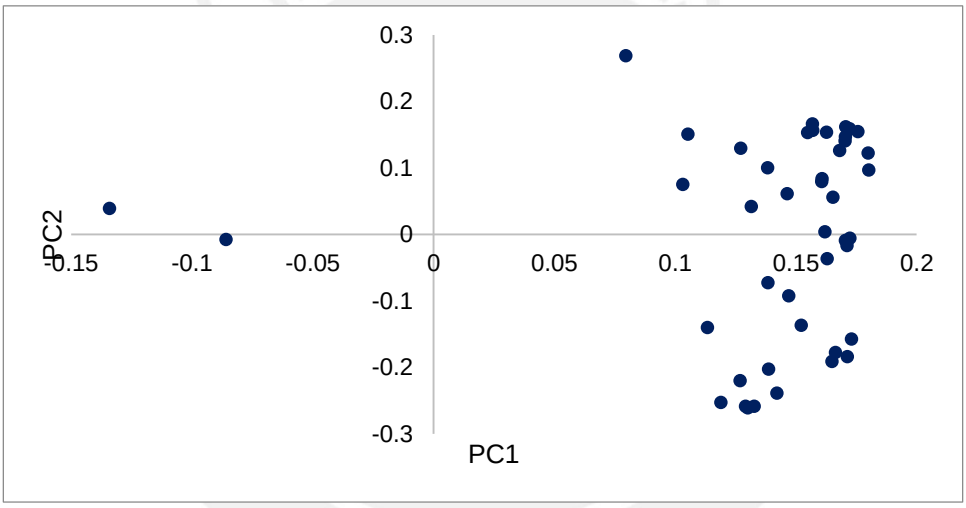


Figure 14. Loadings Plot for second PCA

Now, it can be observed that the significant compounds are all providing loadings in the region of the samples.

Finally, the new Scree plot for the second PCA is show in Figure 15.

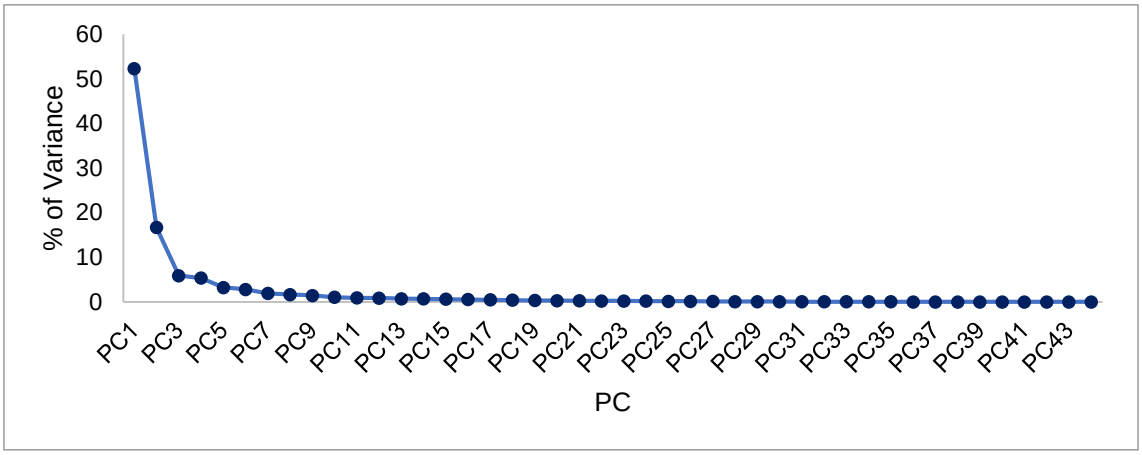


Figure 15. Scree plot for second PCA

This Scree plot now shows that the two first principal components account for over 65% of the variation in the data.

After filtration by the method described in section 4.1, the number of features obtained was reduced to 44, from over 400 initial features.

According to the statistical processing, the following compounds of interest were identified after filtration (Table 4).

Table 4. Identified compounds of interest

ID	Compound	Formula	Calculated RI	Theoretical Mass	Experimental Mass
ID_249	2-Methyl-1,4-benzoquinone (MBQ)	$C_7H_6O_2$	1017	122.03733	122.03632
ID_279	2-Ethyl-1,4-benzoquinone (EBQ)	$C_8H_8O_2$	1110	136.05188	136.05176
ID_282	2-Ethyl-1,4-hydroquinone	$C_8H_{10}O_2$	1430	138.06753	138.06735

5. DISCUSSION

5.1. RESULTS FROM MASS SPECTROMETRY

Figures 12 and 13 (Chapter 4) show the obtained PCA scores plot from our data filtering. Some tendencies can be clearly observed.

Firstly, after the PCA technique the filtered signals were associated to a non-trivial property. This can be seen by the successful separation between samples with biological information from the blank and control samples. Secondly, it can be noticed in Figure 15 that while there is not much of a clear distinction between male and female samples, there is some degree of separation between experienced individual's samples taken as a whole (both male and female) and virgin individual's samples taken as a whole (both male and female). Notice that the distinction is principally along the first principal component (PC1, X-axis). A small distinction along this axis carries a high degree of significance as seen in the Scree plot: PC1 accounts for 52% of the variation within the dataset.

These results suggest that cuticular compounds for beetles are more closely related to their reproductive status than to their actual sex: Within the separate experienced and virgin individual's clusters within the PCA, beetles of different sex are shown to cluster together without many degrees of difference between them. This may imply, for instance, that a given virgin beetle will produce pheromones (likely, aggregation pheromones) that are not closely related to sex. The poor effect that sex has on the chemical composition of the beetles cuticular compounds could potentially lead homosexual behavior within these species of beetle. For the case of sexually experienced beetles, they are likely to produce pheromones which repel other beetles, regardless of their sex – sexual status has a more significant effect on chemical composition than sex.

5.2. COMPOUND IDENTIFICATION

For each mass fragment, the Kovats retention index was calculated by utilizing the retention times of the alkane mix. Utilizing both the Kovats retention index and the exact mass of each ion, a search was carried out in the pheromone database Pherobase [29].

Compounds with a similar Kovats retention index (+/- 30) and a similar exact mass (+/- 0.05) were selected as possible candidates for a positive identification. Then, fragmentation of these compounds was proposed to account for mass fragments. The exact mass of each proposed fragment was calculated utilizing an online mass calculator [30] and matched to the fragments on the table.

Alongside this, previous literature on *A. diaperinus* compounds [15,16,31] was revised and compounds which matched the Kovats retention index and mass were positively identified.

The identified compounds have all been previously reported as pheromones (in the online database PheroBase [29]) for different species within the same family such as *Epanitus obscurus, Zadenos delandei* and *Zadenos multanti*, all within the Coleoptera family , which would suggest plausibility for them being *A. diaperinus* compounds. Additionally, several of them, benzoquinones, have been previously reported specifically for *A. diaperinus* [32]. According to the method utilized for compound identification, it is proposed that the selected compounds have been identified correctly.

Further analysis of the filtered compound table would be possible by utilizing several different databases and cross-checking between them to ensure a positive identification.

5.3 PROPOSED SYNTHESIS: 2-ETHYL-1,4-BENZOQUINONE

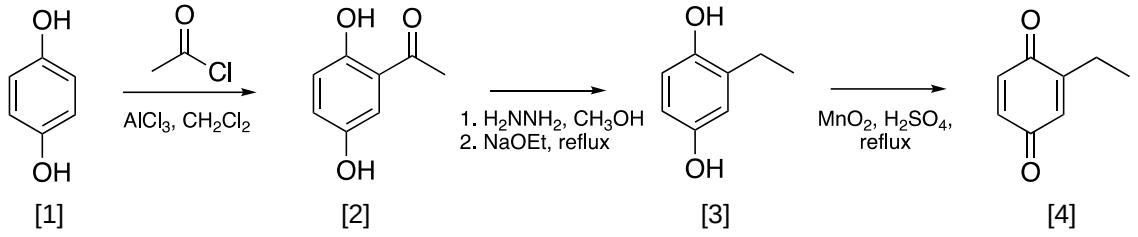


Figure 16. Proposed synthetic scheme for obtaining 2-ethyl-1,4-benzoquinone

A synthesis for one of the identified compounds, 2-ethyl-1,4-benzoquinone is herein proposed (based on [33-35]) (Figure 16), since this material is not widely available from commercial suppliers, and the additional difficulty of shipping chemicals to Peru. Synthesizing it in a simple fashion would be highly beneficial. Additionally, its synthesis would allow for further study into its possible applications for biotraps, given the results previously obtained which positively identify this compound as one of importance for *A. diaperinus*.

The proposed synthetic scheme is shown in Figure 16. With 1,4-hydroquinone as the starting material (1), a Friedel-Crafts acylation is proposed, whereby the acetyl group is introduced into the aromatic ring by treatment with acetyl chloride and anhydrous aluminum chloride using methylene chloride as the solvent. This initial reaction requires no reflux, instead cooling is applied during addition and then the reaction is kept at room temperature for its progression. Workup includes extraction and distillation to recover the acetylated product [33].

Then, the acetylated product (2) is reduced via a Wolff-Kishner reduction. This reduction will not affect the hydroxyl group, due to the reaction mechanism requiring formation of an imine at the carbonyl position for reduction to proceed. The acetylated product obtained from the previous reaction is treated with hydrazine hydrate, allowing the imine product to form under reflux conditions, utilizing methanol as the solvent. The resulting

imine is treated with sodium ethoxide (formed via reaction of metallic sodium with ethanol) under reflux. Workup includes extraction, yielding an ethyl-substituted aromatic ring [34].

Finally, (3) is oxidized via an adaption of a previously reported synthetic route [35] for oxidation of hydroquinone to 1,4-benzoquinone. Herein, 2-ethyl-1,4-hydroquinone is treated with manganese dioxide in acidic medium yielding the desired 2-ethyl-1,4-benzoquinone. Despite manganese dioxide not being the best oxidation agent available, the added stabilization by delocalized electrons found on the product would signify a much higher yield when utilizing this fairly simple oxidant. The reaction is carried out under distillation conditions, where the desired quinone (4) is the first to distill over.

The synthesis was not possible to perform due to restrictions on laboratory work because of the global COVID-19 pandemic at the time of developing this thesis work.

6. CONCLUSION

In conclusion, the data received from GC-MS (CI) analyzed by using techniques such as one-way ANOVA (p-value < 0.01) and PCA allowed for positive identification of three compounds of interest for *A. diaperinus*: 2-methy-1,4-benzoquinone, 2-ethyl-1,4-benzoquinone and 2-ethyl-1,4-hydroquinone. The methodology applied herein could be extended to further analytical data, allowing for reliable identification of compounds of interest.

Additionally, a reasonable synthetic route for one of the selected compounds of interest is proposed, to ease further studies towards more accurate validation of the action of these compounds and the possibility of using them in biotraps.

Future work may consist in using additional data sets, which may include different beetle groups. The use of multivariate statistical analysis will allow to identify common semiochemicals that can be used for developing a biotrap for multiple insect species simultaneously.

7. REFERENCES

1. Integrated Taxonomic Information System. Alphitobius diaperinus (Panzer, 1797). *Taxonomic Serial No. 187197* https://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=187197#null. Date Accessed: 16/11/2021.
2. Guimas, A. Control de plagas de Alphitobius diaperinus en la industria avícola mediante el uso de biotrampas en base a feromonas. (Pontificia Universidad Católica del Perú, 2020). URL: https://repositorio.pucp.edu.pe/index/handle/123456789/179150
3. Chernaki, A. M. & De Almeida, L. M. Exigências Térmicas , Período de Desenvolvimento e Sobrevivência de Imaturos de Alphitobius diaperinus (Panzer) (Coleoptera : Tenebrionidae). *Neotrop. Entomol.* **30**, 365–368 (2001).
4. Segabinazi, S. D., Flôres, M. L., Barcelos, S., Jacobsen, G. & Eltz, R. D. Bactérias da família Enterobacteriaceae em Alphitobius diaperinus oriundos de granjas avícolas dos Estados do Rio Grande do Sul e Santa Catarina, Brasil. *Acta Sci. Vet.* **33**, 51–55 (2005).
5. Templeton, J. M., De Jong, A. J., Blackall, P. J. & Miflin, J. K. Survival of Campylobacter spp. in darkling beetles (Alphitobius diaperinus) and their larvae in Australia. *Appl. Environ. Microbiol.* **72**, 7909–7911 (2006).
6. De Las Casas, E., Harein, P. K., Deshmukh, D. R. & Pomeroy, B. S. Relationship between the lesser mealworm, fowl pox, and Newcastle disease virus in poultry. *J. Econ. Entomol.* **69**, 775–779 (1976).
7. De Las Casas, E. D. Las, Harein, P. K. & Pomeroy, B. S. Bacteria and Fungi within the Lesser Mealworm 1 Collected from Poultry Brooder Houses 2. *Environ. Entomol.* **1**, 27–30 (1972).
8. Goodwin, M. A. & Waltman, W. D. Transmission of Eimeria, Viruses, and Bacteria To Chicks : Darkling Beetles As Vectors of Pathogens. *J. Appl. Poult. Res.* **5**, 51–55 (1996).
9. Watson, D. W., Guy, J. S. & Stringham, S. M. Limited transmission of turkey coronavirus in young turkeys by adult Alphitobius diaperinus (Coleoptera: Tenebrionidae). *J. Med. Entomol.* **37**, 480–483 (2000).
10. El Sitio Avícola. http://www.elsitioavicola.com/articles/2256/control-del-escarabajo-de-la-cama-en-avicultura/ (2012). Date Accessed: 16/11/2021.
11. Jakuboski, S. The Dangers of Pesticides. *Scitable by Nature Education. Blog: Green Science* https://www.nature.com/scitable/blog/green-science/the_dangers_of_pesticides/ (2011). Date Accessed: 16/11/2021.
12. Wolf, J., Potrich, M., Lozano, E. R., Gouvea, A. & Pegorini, C. S. Combined physical and chemical methods to control lesser mealworm beetles under laboratory conditions. *Poult. Sci.* **94**, 1145–1149 (2015).
13. Lara, L. J. & Rostagno, M. H. Impact of heat stress on poultry production. *Animals* **3**, 356–369 (2013).

14. Hassemer, M. J., Borges, M., Withall, D. M., Pickett, J. A., Laumann, R. A., Birkett, M. A. & Blassioli-Moraes, M. C.. Development of pull and push – pull systems for management of lesser mealworm , Alphitobius diaperinus , in poultry houses using alarm and aggregation pheromones. *Pest. Manag. Sci.* **75**, 1107–1114 (2019).

15. Hassemer, M. J., Sant'ana, J., Withall, D., Pickett, J. A., de Oliveira, M. W. M., Laumann, R. A., Birkett, M. A. & Blassioli-Moraes, M. C.. Revisiting the Male-Produced Aggregation Pheromone of the Lesser Mealworm, Alphitobius diaperinus (Coleoptera, Tenebrionidae): Identification of a Six-Component Pheromone from a Brazilian Population. *J. Agric. Food Chem.* **64**, 6809–6818 (2016).

16. Baran, B., Krzyzowski, M., Cup, M. Janiec, J, Grabowski, M. & Francikowski, J. Repellent effect of volatile fatty acids on lesser mealworm (Alphitobius diaperinus). *Insects* **9**, (2018).

17. Nielsen, K. F. & Larsen, T. O. The importance of mass spectrometric dereplication in fungal secondary metabolite analysis. *Front. Microbiol.* **6**, 1–15 (2015).

18. Creative Proteomics. Electrion Ionization. https://www.creative-proteomics.com/support/electron-ionization.htm.

19. Chemistry LibreTexts. Gas Chromatography. https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_(Analytical_Chemistry)/Instrumental_Analysis/Chromatography/Gas_Chromatography. Date Accessed: 16/11/2021.

20. Wikimedia Commons. Electron Ionization. https://upload.wikimedia.org/wikipedia/commons/0/0c/Electron_Ionization.svg. Date Accessed: 16/11/2021.

21. Helal, A. I. State of the Art of Hard and Soft Ionization Mass Spectrometry. *At. Energy Auth.*

22. Munson, B. Chemical Ionization Mass Spectrometry: Theory and Applications. *Encycl. Anal. Chem.* 1–22 (2000).

23. Wikimedia Commons. Chemical Ionization. https://upload.wikimedia.org/wikipedia/commons/7/7b/Chemical_Ionization.png. Date Accessed: 16/11/2021.

24. TOF WERK. Advantages of Time-of-Flight Mass Spectrometry Over Quadrupole MS. https://www.tofwerk.com/advantages-time-of-flight-mass-spectrometry-over-quadrupole-ms/. Date Accessed: 16/11/2021.

25. Böcker, S. & Kaltenbach, H. M. Mass spectra alignments and their significance. *J. Discret. Algorithms* **5**, 714–728 (2007).

26. RIKEN. MS-DIAL Software. http://prime.psc.riken.jp/compms/msdial/main.html.

27. Builtin. A Step-by-Step Explanation of Principal Component Analysis (PCA). https://builtin.com/data-science/step-step-explanation-principal-component-analysis. Date Accessed: 16/11/2021.

28. Romero, P. E. *et al.* From the Andes to the desert: 16S rRNA metabarcoding

characterization of aquatic bacterial communities in the Rimac river, the main source of water for Lima, Peru. *PLoS One* **16**, e0250401 (2021).

29. El-Sayed, A. The Pherobase: Database of Pheromones and Semiochemicals. (2020).

30. Group, W. U.-B. Mass calculations: mass error and m/z from formula. https://warwick.ac.uk/fac/sci/chemistry/research/barrow/barrowgroup/calculators/mass_errors/. Date Accessed: 16/11/2021.

31. Bartelt, R. J., Zilkowski, B. W., Cossé, A. A., Steelman, C. D. & Singh, N. Male-produced aggregation pheromone of the lesser mealworm beetle, alphitobius diaperinus. *J. Chem. Ecol.* **35**, 422–434 (2009).

32. Hassemer, M. J. *et al.* Chemical Composition of Alphitobius diaperinus (Coleoptera: Tenebrionidae) Abdominal Glands and the Influence of 1,4-Benzoquinones on Its Behavior. *J. Econ. Entomol.* **108**, 2107–2116 (2015).

33. University of Michigan. Friedel-Crafts Acylations. http://www.umich.edu/~chemh215/CHEM216/Experiment1/experiment 1.pdf. Date Accessed: 16/11/2021.

34. Soriano, D. S. Example of the Wolff-Kishner reduction procedure suitable for an undergraduate organic lab experiment: Preparation of oxindole. *J. Chem. Educ.* **70**, 332 (1993).

35. PrepChem. Prepation of p-benzoquinone. https://prepchem.com/synthesis-of-p-benzoquinone/. Date Accessed: 16/11/2021.