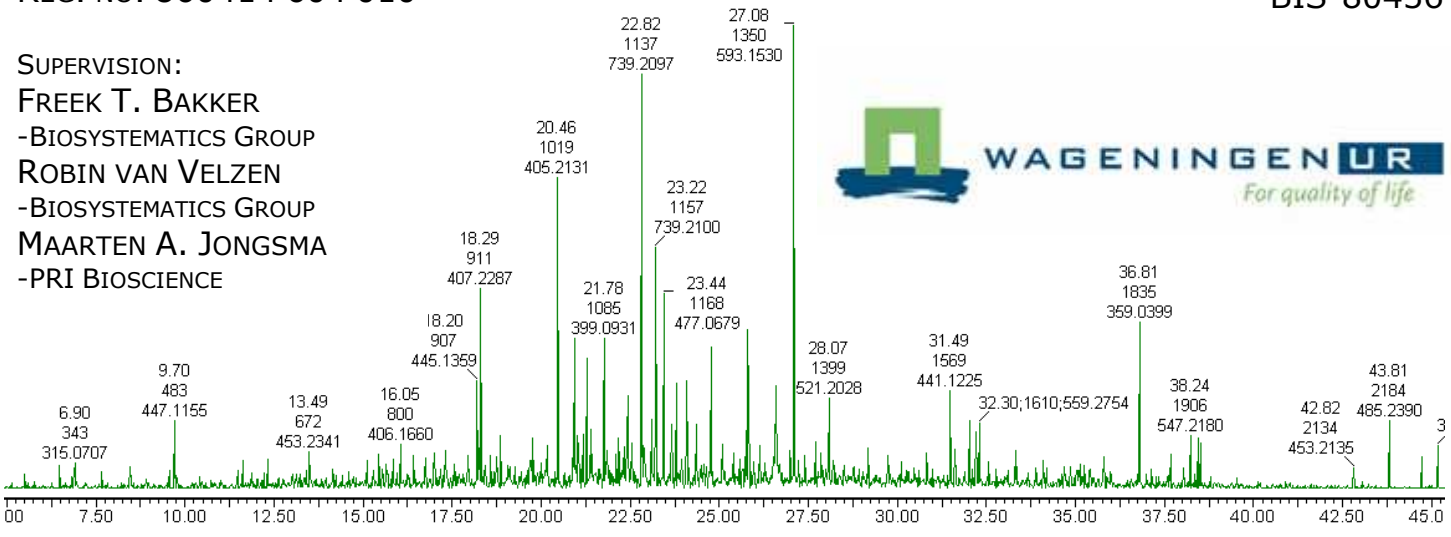
THE CHEMISTRY OF SPECIFICITY IN THE *RINOREA-CYMOTHOE* HOST-HERBIVORE ASSOCIATION

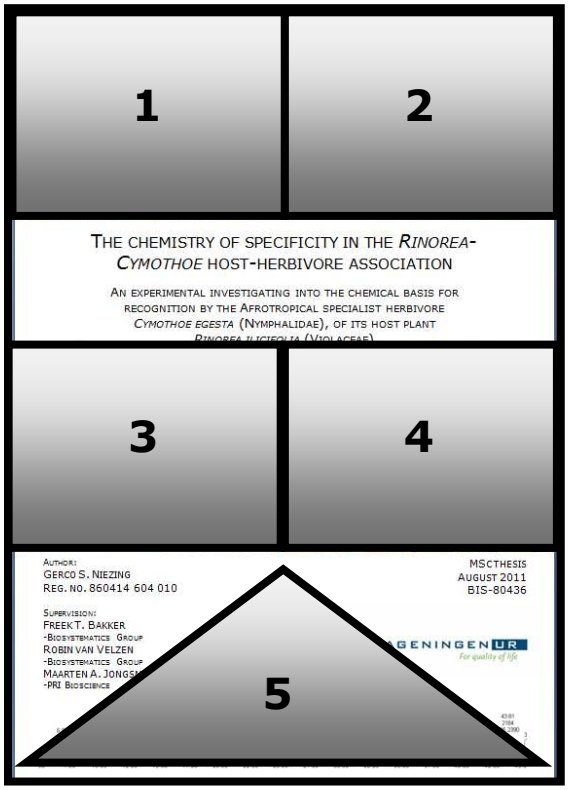
AN EXPERIMENTAL INVESTIGATING INTO THE CHEMICAL BASIS FOR RECOGNITION BY THE AFROTROPICAL SPECIALIST HERBIVORE *CYMOTHOE EGESTA* (NYMPHALIDAE), OF ITS HOST PLANT *RINOREA ILICIFOLIA* (VIOLACEAE)

AUTHOR:
GERCO S. NIEZING
REG. NO. 860414 604 010

SUPERVISION:
FREEK T. BAKKER
-BIOSYSTEMATICS GROUP
ROBIN VAN VELZEN
-BIOSYSTEMATICS GROUP
MAARTEN A. JONGSMA
-PRI BIOSCIENCE

MSC THESIS
AUGUST 2011
BIS-80436





Title Page:
1: View of Kakum National Park from the Canopy Walkway

2: *Cymothoe egesta* eggs on a *Rinorea ilicifolia* leaf

3: Cage experiment with a female *Cymothoe egesta* feeding on coconut

4: *Cymothoe egesta* caterpillar on a young *Rinorea ilicifolia* leaf

5: LC-MS chromatogram, i.e. "the chemistry underlying host specificity"

THE CHEMISTRY OF SPECIFICITY IN THE *RINOREA-CYMOTHOE* HOST-HERBIVORE ASSOCIATION

AN EXPERIMENTAL INVESTIGATING INTO THE CHEMICAL BASIS FOR RECOGNITION BY THE AFROTROPICAL SPECIALIST HERBIVORE *CYMOTHOE EGESTA* (NYMPHALIDAE), OF ITS HOST PLANT *RINOREA ILICIFOLIA* (VIOLACEAE)

AUGUST 2011 MSC THESIS BIS-80436

Author:
Gerco S. Niezing, BSc Reg. No. 860414 604 010 gerco.niezing@wur.nl

Supervisors:
Dr. Freek T. Bakker Biosystematics Group freek.bakker@wur.nl
Ir. Robin van Velzen Biosystematics Group robin.vanvelzen@wur.nl
Dr. Ir. Maarten A. Jongsma PRI Bioscience maarten.jongsma@wur.nl

PREFACE

It was my ambition to have this thesis do two things for me personally: allow me to go somewhere far, warm and adventurous, as well as let me have a taste of scientific research. The first goal was met ten times over. I lived with a local family in a Ghanaian village where I connected intimately with the local culture and – every biologist's dream – I could run around in a tropical rainforest.

With regard to the second goal, I have explored some very diverse scientific aspects: field work, behavioural experimentation as well as lab work with cutting edge chemical analyses. At the top-level, this thesis deals with macroevolution, the science of speciation and the basis of biodiversity; this fascinates me more than anything. The methods I have used have never been performed in this manner before to my knowledge and that in my opinion already warrants this study.

I did however often wonder whether I had not bitten off more than I could chew in constructing this project. Whilst actually being an animal biologist, I have herewith performed an investigation in the fields of ecology and phytochemistry using experimental methods that have never been done in this way before, at a chair group that normally deals with phylogenetic studies.

Still, I am very well aware of the fact that the MSc thesis investigation is a training period. I have learned a great deal about science, evolution, the world and myself in the past year. It is my only hope that the esteemed reader will also learn something about the interesting world of insect-plant biology and the role of chemistry in its evolution.

Gerco Niezing
Wageningen, August 2011

TABLE OF CONTENTS

SUMMARY

Plants and phytophagous insects together show an enormous diversity. Studying the driving factors behind the diversification of these groups is a major topic in ecology and evolutionary biology. One of the most striking observations regarding this topic is the high level of host specialization by herbivorous insects. There may be a process of reciprocal co-evolution at the basis of the observed patterns. Most evidence for this however, is gathered with clear cases of plant defence and insect adaptation to these defences. Whether plants radiate because of selective pressure from insects, and what causes insects to be host specific even on apparently undefended plants is still not quite understood.

This MSc thesis deals with these issues using the model host-herbivore association *Rinorea-Cymothoe*. This is an "extreme case" of host specificity involving a large number of congeneric species, and without clear co-adaptation regarding (anti-)herbivory. Within the WU biosystematics group, this association is being study by PhD student Robin van Velzen, who is performing phylogenetic studies of both genera to gain insight into the evolutionary patterns. In order to make this a more integrative study on different levels, I have attempted to uncover the chemical basis for the proximate cause of host specificity: female oviposition preference.

To this end, I have conducted field work in Kakum National Park, Ghana, during which I observed and experimented on the model association *Rinorea ilicifolia-Cymothoe egesta*. I had two main approaches: 1) untargeted metabolomics to select hypothetical oviposition stimulants to *Cymothoe egesta*, and 2) testing of the hypothesis that 2-(3'-cyclopentenyl)glycine (CPG) is a chemical used for host recognition by *Cymothoe egesta*. CPG has been found by an earlier study to be present in *R. ilicifolia* and speculated to be a host recognition chemical, on the basis of similar chemicals in related taxa in the Malphigiales clade.

For the first approach, I gathered leaves of which I had observed acceptance by the female butterflies – meaning they laid an egg after having alighted upon and evaluated the leaf – as well as rejected leaves. These I dried and brought back to the laboratory, where I performed LC-MS on these samples and compared the metabolic profiles of accepted and rejected leaves. I selected several metabolites that correlate with butterfly behaviour in this manner, however I found that *C. egesta* will almost invariably accept young leaves and reject old leaves. The correlation I found is therefore likely to have more to do with leaf ontogeny than bioactivity. I did attempt to identify several metabolites based on their molecular weight, as proof of principle for this method.

The second approach required the capture of butterflies, in order to test whether they responded to CPG by laying an egg after they taste it, i.e. whether it is an oviposition stimulant. The experimental method I tried involved agar plates as artificial substrate. An objective of this approach was to use it as a pilot for behavioural experimentation involving *Cymothoe.* No eggs were laid on these agar plates, but there were on branches I had placed inside the cage, which became further proof of the young leaf preference. Another approach to this hypothesis was to also use LC-MS to measure the levels of CPG in accepted and rejected leaves and see whether this chemical correlates to bioactivity. No correlation was found for this whatsoever and although the sample size of was smaller, it does not appear that CPG in *R. ilicifolia* is an oviposition stimulant to *C. egesta*, provided this butterfly selects for leaves with the highest levels of oviposition stimulant.

Results were inconclusive of this investigation, but recommendations are given to improve upon a hypothetical follow-up experiment or alternative, related experiments.

Chapter 1 Introduction

1.1 GENERAL INTRODUCTION

It has been estimated that, together with the plants they feed on, herbivorous insects make up about half of all multicellular taxa (excluding fungi, algae and microbes) found on this planet (Strong, Lawton & Southwood, 1984). Other estimates indicate that this figure may even be an underestimation (Ødegaard, 2000; Novotny, 2002). Their sheer diversity makes herbivorous insects a major conduit of solar energy from plants to higher trophic levels (Futuyma & Agrawal, 2009). Apart from providing a good model system for the study of the mechanisms of speciation and the driving forces of evolution in general, a greater understanding of insect-plant relationships will allow for a more effective conservation of biodiversity (especially in tropical areas). Moreover, it can provide insight in phytophagous insect's capabilities for adaptation to pesticides as well as assist in predicting the behavior of invasive species and their development into agricultural pests, with respect to the global warming model.

Elucidating the driving factors behind the remarkable diversification in herbivorous insect and host-plant clades, is therefore a major topic in ecology and evolutionary biology (Schoonhoven, Van Loon & Dicke, 2005; Futuyma & Agrawal, 2009; Nyman, 2010). A striking feature common throughout herbivorous insect clades that was recognized early on, is that many species have specialized associations with plant hosts, and are mono- or oligophagous, i.e. feeding on a single, or only on a few plant species or lineages (Thompson, 1994; Bernays & Chapman, 1994). Furthermore, it was also observed that related butterflies generally feed on related plants (Ehrlich & Raven, 1964; Janz & Nylin, 1998; Nyman, 2010). In this respect, the importance of secondary plant chemicals ("secondary" because such chemicals do not appear to have a function in "primary" metabolism) has long been acknowledged, as compounds present in plants may be toxic to some herbivores, and harmless to others, thereby restricting herbivore diets (Dethier, 1954).

In an effort to explain the observations listed above, it has been proposed by Ehrlich and Raven (1964) that plants and herbivorous insects drive reciprocal co-speciation due to a state of constant arms race. The theory states that plants develop novel (chemical) defenses in response to herbivory, and – having temporarily escaped herbivory – radiate into a lineage of species sharing this novel defense. Subsequently, insect herbivores will undergo adaptation to increase their tolerance of these defenses and colonize this herbivore free lineage, thus driving co-speciation (Ehrlich and Raven, 1964). According to this theory, herbivorous insects have thus developed a tolerance to the toxic and repellent chemical properties of a limited number of plants species, which is the basis of host specificity.

This "escape and radiation" theory has been criticized, especially on the notion that phytophagous insects drive plant speciation in the manner described above. For example, Jermy (1993) has stated that there is little evidence that insect herbivores have a negative effect on plant fitness and that the insect-plant relationships are the result of unconstrained selection. Other authors have shared such reservations, as reviewed by Cornell & Hawkins (2003), and have opted that selective pressure placed on plants by insect herbivores is too weak and diffuse to drive plant speciation, plant defenses may be aimed against microorganisms rather than insects and that the existence of generalists does not support the prediction of increased specialization (Cornell & Hawkins, 2003). Nevertheless, evidence for many of the predictions of the "escape and radiation" theory of Ehrlich and Raven (1964) has been gathered and there is little doubt that plants have developed defenses against herbivory, and that many of the toxic plant compounds have a defensive function aimed against insect herbivores (Cornell & Hawkins, 2003; Futuyma & Agrawall, 2009).

A strong indication that insects have certainly influenced the evolution of defensive traits in plants is the convergence of such traits in unrelated lineages, which is the case with for example latex formation, resin canals, cyanogenesis and induction of

indirect defense involving parasitoids (Farell et al., 1991; Agrawal, 2011). The escalation (increase) of defensive chemicals over time is another indicator of selection, an example of which is provided by Agrawal, Salminen & Fishbein (2009), who used a combination of phylogenetic tools and chemical analyses to show that the production of phenolic compounds showed an increase over time in milkweeds (*Asclepias* spp.). Conversely, cardenolides, which are sequestered by specialist insect herbivores of *Asclepias* actually show a trend towards decline as the lineage radiates (Agrawal, Salminen & Fishbein, 2009).

A well-documented example of what is generally thought to be the result of a co-evolutionary arms race, is the case of the tropical trees of the genus *Bursera*, the main herbivores of which belong to the chrysomelid beetle genus *Blepharida*, an association that has been extensively studied by Becerra (Becerra & Venable, 1999; Becerra, 2003; Becerra, Noge & Venable, 2009). *Bursera* displays clear chemical defenses such as an array of terpenes and other compounds toxic to insect herbivores. Additionally, some species have evolved a manner of mechanically squirting resin, mechanisms which act to repel all herbivores except for specialized species of *Blepharida* (Becerra & Venable, 1999). This association was found to be very old (approx. 112 myr) and more or less phylogenetically congruent (Becerra, 2003). Also, chemical diversity of *Bursera* has been found to have increased in time, which indicates an escalation in chemical defenses, presumably as a result of herbivore induced selective pressure (Becerra, Noge & Venable, 2009).

Plant defenses such as those described above place obvious selective pressure on herbivorous insects (Schoonhoven, Van Loon & Dicke, 2005; Futuyma & Agrawall, 2009) because adaptation to a well defended hosts reduce an insect's ability with a different host's defenses. Radiation across a plant lineage with a shared anti-herbivory defense can occur quite rapidly once an evolutionary innovation has occurred in herbivores. This was shown for the tribe Pierinae (Pieridae), which is much more diverse than its sister clade, due to the evolution of a nitrile-specifier protein that detoxifies glucosinolate defenses prevalent in the Brassicaceae, which allowed the Pierinae to specialize on this plant lineage and diversify (Wheat et al, 2007). Overall, it is clear that plant diversity has increased the diversity of their associated herbivores (Janz, Nylin & Wahlberg, 2006). An explanation for speciation may be the "Oscillation hypothesis" by Janz & Nylin (2008) which states that host specific insects retain the ability to feed on their ancestral hosts, allowing them to jump hosts during a brief periods of polyphagy. A new host can have different geographical distribution range, allowing an insect to spread to an area where only the new hosts are present, causing vicariance and subsequent speciation.

Yet, plants may not always have defensive systems as clear as those described above, but still be host to highly specific herbivores. Of course, host quality is in the eye of the beholder, and what appears non-toxic to us, may well be influencing acceptability by insects (Nyman, 2010), however rather than use secondary metabolites as toxins and repellent, plants may have developed tolerance to loss of tissue to withstand, or a relationship with insect parasitoids to defend against herbivores (Futuyma & Agrawall, 2009). The mechanisms governing host specificity in such cases are not always clear, as generalist feeding on such plans seems rather a good solution. Some explanations have been proposed. Even though plants may not be actually defending themselves chemically, plants can be very variable in the resources they have to offer and – to an insect – represent moving island in resource space (Nyman, 2010). Nyman (2010) states that ecological speciation can follow an "island hopping" event by an herbivore.

More mechanistic drivers of host specificity that have been proposed are also worth mentioning. Quental, Patten & Pierce (2007) propose that plant derived chemicals can be sequestered by herbivores and subsequently used as male pheromone which attracts a female. The females that are attracted to this pheromone will subsequently place their offspring on those plants containing the chemical basis for the diet-derived pheromone, thereby restricting the number of possible host species (Quental, Patten & Pierce, 2007). Alternatively, it has been proposed by Bernays (2001) that the neurological simplicity of insects constrains the number of variable sensory signals they can distinguish between, when choosing a host. Innately determined preferences for a

limited number of (chemical) host specific cues greatly favor specialization in this manner (Bernays, 2001).

It is clear that the underlying principles of host specificity and the radiation of herbivorous insects and their host plants are still not fully understood. The study of these general principles is best performed using extreme model systems, with high specificity and a large number of congeneric species. The *Rinorea-Cymothoe* host-herbivore association is such a system.

1.2 THE CASE OF RINOREA AND CYMOTHOE

The genus *Rinorea* Aublet (1775) (Violaceae: Violoidea) contains around 200 species of woody shrubs found in tropical areas in both the Old and New World (Hekking 1988). 110-120 of these species occur in Africa, with the centre of distribution found in Cameroon and Gabon (Achoundong, 1996). Species of *Rinorea* are host to the larvae of several species of the genus *Cymothoe* Hübner (1819) (Nymphalidae, Limenitidinae, Neptini). This butterfly genus consists of around 72 species, all occurring in the African tropics (Larsen, 2005). 28 species of *Cymothoe* feed on *Rinorea*, of which 18 are strictly monophagous, 6 are oligophagous and 3 feed on up to six species of *Rinorea* (Amiet & Achoundong, 1996). The high level of host specificity observed among the large numbers of congeneric hosts and associated herbivores, and the fact that *Rinorea* does not appear to have an obvious defensive system aimed at insect herbivores, makes this system unique, especially in the Tropics.

Within the WU Biosystematics Group, PhD student Robin van Velzen is currently investigating the evolution of insect-host plant associations among *Cymothoe* and *Rinorea*. Robin attempts to identify the driving factors behind this pattern, using the *Rinorea-Cymothoe* case as a model system for the study of speciation and evolution of host plant-insect associations in general.

Robin's research aims to test for congruence between dated phylogenetic reconstructions of *Cymothoe* and *Rinorea*, which would co-speciation and long-term association of *Cymothoe* and *Rinorea* i.e. "host tracking". Alternatively, a mechanism where *Cymothoe* have colonized and radiated sequentially across long diversified *Rinoreae*, would imply more functional tracking, see figure 1.1. To this end, phylogenetic studies have been performed by Onstein (2010) and Van Velzen (2006, ongoing). Preliminary results seem to favor the latter hypothesis (Van Velzen, personal communication), which is consistent with other such studies (Futuyma & Agrawal, 2009).

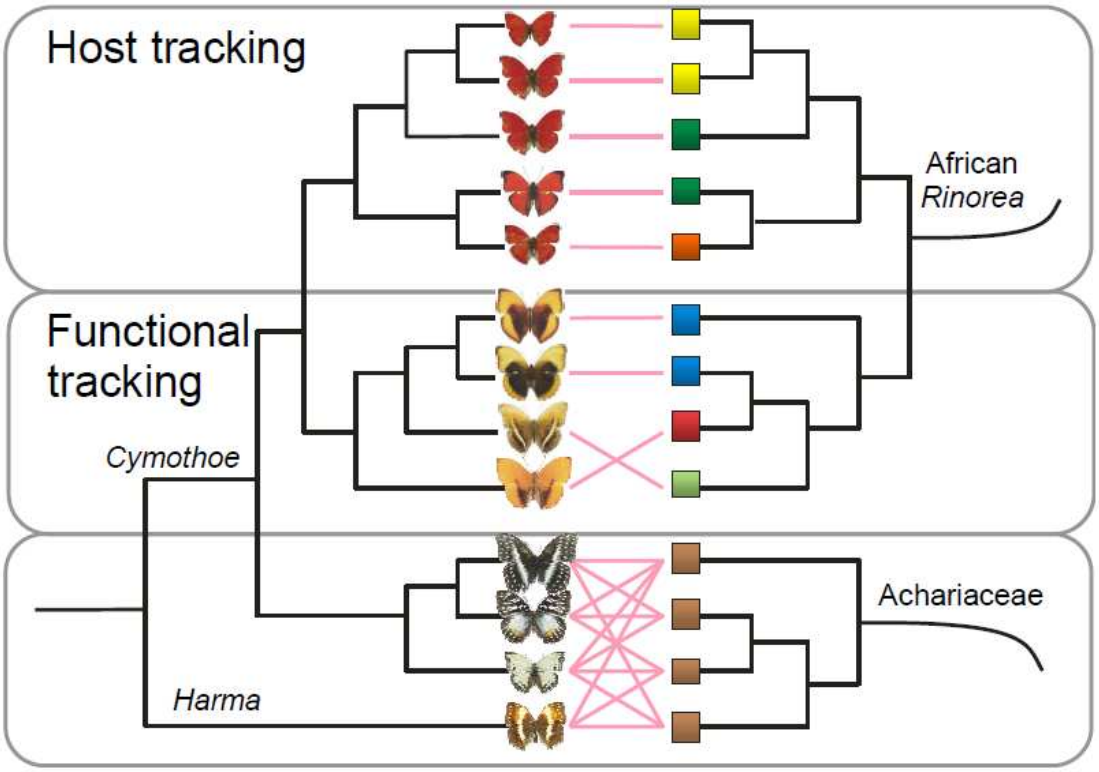


Figure 1.1 Schematic tanglegram of *Cymothoe-Rinorea* associations. *Cymothoe* (left) and *Rinorea* (right) cladograms are purely hypothetical. Associations are shown in pink. Coloured squares represent host phytochemical properties. Top: phylogenetic congruence, confirming the *host tracking* hypothesis. Middle: host shifts mediated by host phytochemistry indicating *functional tracking*. Bottom: generalistic associations with ancestral Achariaceae hosts. Adopted from Robin van Velzen.

Concerning the mechanisms driving *Cymothoe* speciation, Van Velzen also addresses whether related *Cymothoe* feed on related *Rinorea* hosts, whether *Cymothoe* host use is evolutionarily stable and whether *Cymothoe* diversification rates have correlated with host range expansion (i.e. whether colonization of *Rinorea* has increased the diversification rate of *Rinorea* feeding *Cymothoe* relative to the ancestral state of feeding on Achariaceae). Preliminary results suggest that related *Cymothoe* generally feed on related *Rinorea*, but that host jumps have occurred and *Rinorea* feeding may not be evolutionarily stable (Van Velzen, project progress presentation 2011). These results would be in accordance with the "escape and radiation" hypothesis of Ehrlich & Raven (1964) as host jumps are mediated by plant chemical properties, which usually do not match plant phylogeny (Futuyma & Agrawall, 2009). However, preliminary results also suggest that *Cymothoe* diversification has increased dramatically at the basal node, but not at the moment of *Rinorea* colonization (Van Velzen, personal communication).

Apart from these phylogenetic analyses, investigating the chemical basis underlying the observed patterns will lead to a greater understanding of the processes governing *Cymothoe* radiation. This will allow for a true integrative study of the macroevolutionary mechanisms that are involved in this remarkable host-herbivore association. Attempting to clarify the mechanisms that proximately cause *Cymothoe* host specificity will therefore be the main aim of this MSc thesis.

1.3 OVIPOSITION PREFERENCE AS A PROXIMATE CAUSE OF HOST SPECIFICITY

In order to gain insight in the evolutionary mechanisms that have ultimately driven the host-specific radiation observed in the *Rinorea-Cymothoe* association, I want to attempt to elucidate the proximate cause this host specificity, i.e. answer the question: "what makes a *Cymothoe* species select its specific *Rinorea* host, rather than a closely related *Rinorea* (growing in sympatry)?"

For many phytophagous insects, butterflies in particular, host specificity has two distinct premises: firstly, a larval insect needs to recognize its host and be stimulated to feed on it, while simultaneously its digestive system needs to cope with the chemical make-up of this host; and secondly, a fertilized female searching for a host needs to recognize its specific host and be stimulated to oviposit (i.e. lay an egg) on it (Schoonhoven, Van Loon & Dicke, 2005). Both of these premises are genetically determined and can therefore hypothetically be selected upon (Wiklund, 1975; Schoonhoven, Van Loon & Dicke, 2005).

An insect larva in general and a butterfly caterpillar in particular, has limited motility in comparison to the adult. Initial distribution of larvae is therefore mediated by the oviposition preference of the female insect, a butterfly in this case, causing host recognition by the female to have a greater influence on host specificity than larval performance (Lederhouse et al., 1992). In fact, it appears that larval diet breath can be even wider than the host range selected by the ovipositing female, as has first been found by Wiklund (1975), who showed that *Papilio machaon* caterpillars perform well on plants that are not selected by females. Conversely, some plants were selected by females that were unsuitable as food for the caterpillars (Wiklund, 1975). Similar examples of incongruence between oviposition preference and larval performance have been found in the moths *Spodoptera exigua* (Berdegué, Reitz & Trumble, 1998) and *Trichoplusia ni* (Shikano, Akhtar & Isman, 2010) and the gall midge *Rabdophaga rosaria* (Nyman et al., 2011). In the *Nymphalidae* tribe *Nymphalini* it was found that the larvae of related butterflies could survive on plants on which females laid no eggs (Janz, Nyblom & Nylin, 2001). Female host recognition and oviposition preference, rather than larval performance are therefore thought to be responsible for host specificity (Lederhouse et al., 1991; Honda, 1995; Schoonhoven, Van Loon & Dicke 2005). In the case of *Rinorea* and *Cymothoe*, no incongruence between female preference and larval performance has been measured. Caterpillars have been found to conform to the choice of the females (Van Velzen, 2006; Onstein, 2010) but given the above, I will attempt to elucidate the mechanism behind host selection and subsequent oviposition by the female.

Table 1.1 List of Nymphalid butterflies (tribe mentioned in brackets) with host plants, confirmed oviposition stimulants and publications.

Nymphalid	Host plant	Oviposition Stimulant	Reference
Junonia coenia (Nymphalinae)	Plantago lanceolata (Plantaginaceae)	Iridoid glycosides: catalpol & aucubin	Pereyra & Bowers , 1988
Melitaea cinxia (Nymphalinae)	Plantago laceolata (Plantaginaceae)	Iridoid glycosides: catalpol & aucubin	Reudler Talsma et al., 2008
Danaus plexippus (Danainae)	Asclepias curassavica (Apocynaceae)	Flavonoid glycosides	Haribal & Renwick , 1996
Idea leuconae (Danainae)	Parsonsia laevigata (Apocynaceae)	Pyrrolizidine alkaloids	Honda et al., 1997
Ideopsis similis (Danainae)	*Tylophora tanakae* (Asclepiadaceae)	Phenanthroindolizidine alkaloids	Honda et al., 2001
Parantica sita (Danainae)	Marsdenia tomentosa (Asclepidaceae)	Conduritols	Honda et al., 2004

1.4 IDENTIFYING CHEMICALS INVOLVED IN HOST RECOGNITION BY CYMOTHOE

Before oviposition occurs, a female butterfly will have identified a leaf as belonging to (one of) its host plant(s). This recognition can occur in two stages leading up to oviposition. First, visual and/or olfactory cues are needed for the butterfly to alight (land) on the leaf for further investigation. Second, surface evaluation and tasting of the chemical make-up of the leaf will stimulate the female to oviposit (Renwick & Chew, 1994; Schoonhoven, Van Loon & Dicke, 2005). This surface evaluation is performed during a characteristic "drumming" behavior of the fore-tarsi by the butterfly, during which it detects chemicals present in the leaf (Renwick & Chew 1994; Baur, Haribal & Renwick, 1998). Renske Onstein (2010) and Robin van Velzen (2011) have reported this "drumming" behavior to be clearly audible for some *Cymothoe*.

Plant secondary chemicals that elicit a response from insects were designated etymologically by Dethier, Barton Brown & Smith (1960), and correct use of terminology is required. For host searching and oviposition the following types of compounds can be of significance: "attractant" – a volatile chemical that causes an insect to orient its movement towards the source of the chemical; "repellent" – a volatile chemical that causes an insect to orient its movements away from the source if the chemical; "stimulant" – a non-volatile chemical that elicits a behavior (e.g. oviposition) upon physical contact; "deterrent" a non volatile chemical that inhibits a behavior (e.g. oviposition) upon physical contact (Dethier, Barton Brown & Smith, 1960; Renwick & Chew, 1994). Of these types of chemicals, I will focus on the oviposition stimulants, which – in contrast to moths which are found to be greatly influenced by volatile attractants (Renwick & Chew, 1994) – are generally thought to have the greatest influence on host specificity for Papilionid butterflies (Renwick & Chew, 1994, Honda 1995). Such chemicals are usually polar and present as tightly bound to the leaf cuticle, rather than present in the wax layer (Renwick & Chew, 1994).

The mechanism of host recognition and subsequent oviposition by contact chemoreception has been extensively studied in Papilionid butterflies. The most frequently used model system herein is the polyphagous crucifer feeding Cabbage White *Pieris rapae* (Pieridae) (Schoonhoven, Van Loon & Dicke, 2005). Chemicals found to be stimulating its oviposition are the glucosinolate sinigrin (Renwick & Radke, 1983) and glucobrassicin (Traynier, 1991) that can both individually stimulate oviposition, although the latter shows effect in lower concentrations. Comparative research on other members of the genus *Pieris* has revealed that deterrents also play a role in the differential oviposition preference exhibited by these butterflies (Honda, 1995). Another family that has been studied in detail is the Swallowtail butterflies (Papilionidae). These exhibit a different mechanism of host recognition, generally requiring several chemicals acting in complex synergy to be present in a host plant in order to elicit oviposition (Honda, 1995; Haribal & Feeny, 1998; Nakayama et al., 2003; Murphy & Feeny, 2006). For Papilionidae, plant secondary metabolites that act as deterrents have also been identified (Honda, 1995; Nishida, 2005).

Cymothoe belong to the tribe Limenitidinae of the family Nymphalidae (Larsen, 2005). Although this is a highly speciose family, less is known about the chemical basis for Nymphalidae host specificity than for the Papilionidae and Pieridae. Oviposition stimulants have been identified for some genera, however, see table 1.1. Unfortunately, for the tribe Limenitidinae no oviposition stimulant from a host plant has been positively identified. From the behavioral experimentations performed on other Nymphalidae however, it becomes clear that single secondary plant metabolites usually stimulate oviposition in this family (as in Pieridae, see above). Therefore, I expect a similar mechanism to exist for *Cymothoe*. In this investigation I will therefore assume that the chemical basis for host recognition by *Cymothoe* consists of non-volatile secondary plant metabolites which might be tested with a behavioral experiment.

1.5 IDENTIFYING CHEMICALS INVOLVED IN HOST RECOGNITION BY CYMOTHOE

Before testing the effect on female oviposition behavior of any chemical in a behavioral bioassay, candidate oviposition stimulants first need to be identified. The list in table 1.1 shows that, even within one family of butterflies, a great variety of substances can have this biological significance.

The traditional approach to finding bioactive plant compounds from among the vast variation of plant secondary chemicals is by iterative, bioassay-guided fractionation experiments such as used by Haribal and Renwick (1996) (Van Loon, personal communication), see also figure 1.6A for an overview. Metabolites in table 1.1 have also been found using mainly this method. This method extracts all metabolites from a bioactive plant – e.g. one that is oviposited upon – and divides this extract into fractions, afterwards testing each for bioactivity in a bioassay, e.g. an oviposition experiment. The extract fraction displaying bioactivity is then fractionated again and a subsequent bioassay yields the fraction containing the bioactive compound. This final fraction is then analyzed chemically with structure elucidation methods such as H NMR to identify the bioactive compound. Although this method can yield compounds of which ecological relevance is assured, it is time consuming and costly. Also, the fractionation process creates the risk of the chemicals deteriorating. Moreover, there is a possibility that no single chemical fraction will show a biological function, due to several chemicals acting in synergy, which can make the search difficult (Prince & Pohnert 2010). This lengthy and complicated process is expected to be unfeasible *in situ* (Africa) and impossible within the confines of an MSc thesis.

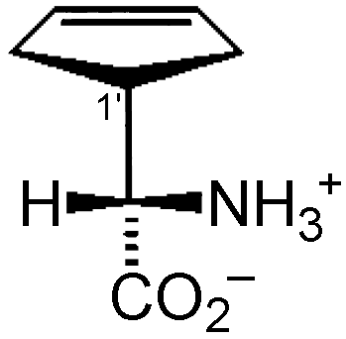


Figure 1.2 Structural formula of 2-(3'-cyclopentenyl)glycine (CPG). Adopted from Clausen et a., 2002.

In this particular case, however, one investigation by Clausen and coworkers (2002) has revealed a candidate for an oviposition stimulant in *Rinorea* for *Cymothoe*, as they discovered 2-(3'-cyclopentenyl)glycine (CPG, my abbreviation, see figure 1.2), a non-proteogenic amino acid, in leaf material of *R. ilicifolia* from Ghana. Achariaceae, Passifloraceae and Turneraceae, which are related to *Rinorea* (family Violaceae) within the order Malpigiales, (Tree of Life Web Project, 2002; Stevens, 2011) were reported to contain chemically related cyanogenic cyclopentenoids (Clausen et al., 2002).

Interestingly, species of the tribe *Heliconiinae* (*Acraea* most notably) feed on these cyanogenic Malpigiales genera, have been found to be cyanogenic themselves (Nahrstedt & Davis, 1981) and contain cyclopentenoid compounds after having fed on Passifloraceae (Engler et al., 2000). Indeed, Heliconiinae and Limenitidinae (to which *Cymothoe* belongs), being sister clades (Wahlberg & Brower, 2009), appear to be closely associated to the Violaceae, Passifloraceae, Turneraceae and Achariaceae, see table 1.2. This is in accordance with the notion that related insects feed on related plants (see chapter 1.1) and it led Clausen and coworkers (2002) to speculate there may be a common chemical character prevalent throughout these plant clades that might be used by these butterflies as common chemical basis for host recognition. Because the cyanogenic cyclopentenoids found in the Passifloraceaous group are similar to CPG, discovered by Clausen et al (2002) in *R. ilicifolia,* although the latter is not cyanogenic, Clausen and coworkers hypothesized that such cyclopentenoids can act as "host recognition templates", i.e. oviposition stimulants to Nymphalid butterflies (Clausen et al., 2002).

Table 1.2 Host plant relationships of *Cymothoe*, *Harma* and Heliconiinae genera feeding on *Rinorea* and families belonging to the Passifloraceous group, according to the online HOSTS database (http://www.nhm.ac.uk/research-curation/projects/hostplants/) Adopted from Van Velzen, 2006

		Host Plant Clades			
Tribe ↓	Genus ↓	*Rinorea*	Achariaceae	Passifloraceae	Turneraceae
Heliconiinae	*Acraea*	×	×	×	×
	Phalantha	×	×		
	Terinos	×	×		
Limenitidinae	*Cymothoe*	×	×		
	Harma		×		

It can be questioned whether the similarities in cyclopentenoid chemicals can be a basis for differential host preference. For example, *Caloncoba echinata*, a host species for several oligophagous butterflies such as *C. jodutta*, *C. druyri* and *Harma theobene* (Van Velzen, 2006), was found to contain mostly 2-2'-CPG and hardly any 2-3'-CPG (Clausen et al., 2002), the only chemical difference between the two being the double bond having shifted one position. This plant occurs in sympatry with *R. ilicifolia* in Ghana and the slight difference between these chemicals might cause *Cymothoe* to distinguish between them. However, continuing this parallel within the wide radiation of *Rinoreae*, this would imply that each *Rinorea* species with a specialist *Cymothoe* would have a slightly different cyclopentenoid. This seems unlikely considering the relative simplicity of the molecule compared to oviposition stimulants known for different butterfly genera (e.g. those listed in table 1.1). Still, CPG may be an oviposition stimulant, in a system where host specificity is determined by differential deterrents rather than oviposition stimulants. Such a system has been discovered in Pieridae and Papilionidae, but would be new in Nymphalidae (Honda, 1995; Nishida, 2005).

The hypothesis of Clausen and coworkers was never tested by means of a behavioral experiment. Proving or disproving that CPG has a stimulatory effect on *Cymothoe* oviposition, i.e. testing whether this chemical has biological significance, is therefore warranted. This is will be one of the main aims of the present investigation, and provides the starting point in the process of elucidating the mechanism of host recognition by *Cymothoe.* Onstein (2010) discovered that in Ghana, *Rinorea ilicifolia* is host to *Cymothoe egesta*. The biological significance of CPG should logically in the first place be tested with Ghanaian *C. egesta* females. I have therefore attempted a choice experiment using wild-caught butterflies, in order to test whether CPG functions as an oviposition stimulant to this butterfly, see chapter 1.8.

Apart from testing the significance of CPG, the present study also aims to identify alternative compounds, i.e. hypothetical oviposition stimulants to *Cymothoe*, using untargeted metabolomics, see chapter 1.7. For convenience, I use the host-herbivore association of *Rinorea ilicifolia-Cymothoe egesta*, as a model association during this study.

1.6 MODEL ASSOCIATION RINOREA ILICIFOLIA-CYMOTHOE EGESTA

Rinorea ilicifolia (Welw. ex Oliv.) Kuntze, 1891 (Malpigiales, Violaceae) is a rainforest shrub with alternate leaf arrangement and simple, dentate leaves. The species has small, yellow quinary flowers (see figure 1.3) and fruits are 3-locular (Hawthorne & Jongkind, 2006). The species is widespread in tropical Africa from Guinea east to Kenya, and south to Angola and Mozambique (Hawthorne & Jongkind, 2006). Clausen and coworkers (2002) found 2-(3'-cyclopentenyl)glycine in leaf material of *Rinorea ilicifolia* from Ghana.

Figure 1.3 Photographs of *Rinorea ilicifolia* I took in Kakum National Park, Ghana. Left: habitus. Top right: flowers. Bottom right: detail of leaf base showing characteristic cuspis and edge spines, as well as two *Cymothoe egesta* eggs.

A striking feature of *R. ilicifolia* are the spines it has along the edges of its leaves, which are not unlike the spines of holly leaves (*Ilex* spp., hence "*ilicifolia*"). Merz (1959) has found some evidence that such spines protect *Ilex* from edge feeding insect herbivores, and similarly these spines could be an anti-herbivory adaptation in *R. ilicifolia*. The anti-herbivory function of *Ilex* spines has since been disputed however as

Figure 1.4 Photographs of *Cymothoe egesta* I took in Kakum NP. Top left: caterpillar on *R. ilicifolia* leaf; bottom left: male; right: female.

Potter and Kimmerer (1988) found a thick cuticle and tough leaf margin to have a greater deterring effect on larvae of generalist herbivore *Hypantria cunea*.

Cymothoe egesta Cramer, 1775 (Lepidoptera, Nymphalidae, Limenitidinae), commonly known as the Yellow Glider, is a rainforest frugivorous butterfly with a forewing of 43 mm (Larsen, 2005; personal observation). Like all *Cymothoe,* this species shows clear sexual dimorphism with males having an ochreous yellow ground colour, and the females being dark brown with a white band across all four wings (Larsen, 2005; personal observation), see figure 1.4. Caterpillars have a green back and red sides with a yellow line in between (Amiet, 1997; personal observation). See figure 1.3 for eggs. The distribution range of the species is West African under "the armpit" until Cameroon and Gabon (McBride, Van Velzen & Larsen, 2009), see figure 1.5. Renske Onstein informed me this species is particularly abundant in Kakum National Park, which is why I chose this location for my field work.

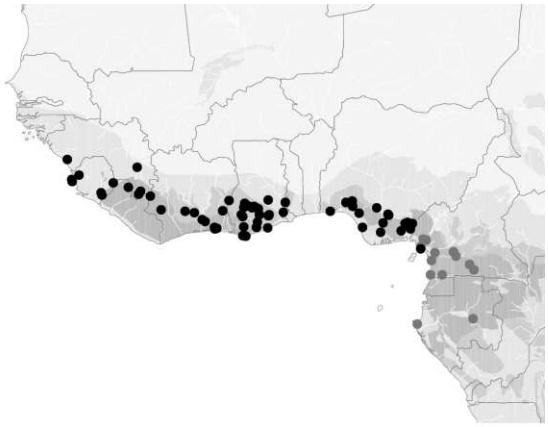


Figure 1.5 Overview of *Cymothoe egesta* collections providing an estimate of distribution range. Adopted from Robin van Velzen. Cryptic species *Cymothoe confusa* has a range further east (McBride, Van Velzen & Larsen, 2009)

It was thought that *C. egesta* was a specialist feeder on *R. lepidobotrys* (Amiet and Achoundong, 1996; Van Velzen, 2006) and *R. braviracemosa* (McBride, Van Velzen & Larsen, 2009), whereas cryptic sister species *C. confusa* fed on *R. ilicifolia,* among other *Rinoreae* not native to Ghana (Van Velzen, 2006; McBride, Van Velzen & Larsen, 2009). Because these two groups of host plants are not very closely related phylogenetically (Bakker et al., 2006; McBride, Van Velzen & Larsen, 2009; Onstein, 2010) it was thought this may be evidence of a "host jump" event (Onstein, 2010). Interestingly, Onstein (2010) rather discovered *C. egesta* to feed on *R. ilicifolia* in Ghana with high specificity. She speculated this may be either due to a misjudgment of *C. egesta* host plant associations in Cameroon, a misidentification of *C. confusa* as *C. egesta* in Ghana or a "host jump" event from *R. ilicifolia* to *R. lepidobotrys* and *R. braviracemosa* by *C. egesta* East of the Dahomey Gap (Onstein, 2010). The possibility that a study into the mechanisms responsible for the observed host-herbivore association patterns of this cryptic species pair, may shed some light on this matter, also makes the *C. egesta-R. ilicifolia* association a very appropriate model association.

1.7 UNTARGETED METABOLOMICS APPROACH

As stated in chapter 1.4, I have also attempted untargeted metabolomics in order to generate novel hypotheses regarding candidate chemicals that may be used by Cymothoe for host recognition (Prince & Pohnert 2010). This method includes chemical analysis, e.g. LC- or GC-MS, of extracts from biologically active and inactive plants and subsequently comparing their metabolic profiles (see figure 1.6). The method that was chosen in this case is LC-MS, which uses Liquid Chromatography to separate, the individual metabolites in a plant extract, and subsequent Mass Spectrometry to visualize, quantify and determine accurate mass of metabolites in an extract. From accurate mass, the molecular formula may be inferred. The metabolic profile of a sample is composed of these data put together. This method is called "untargeted" because it is not directed towards a single, or a class of compounds, unlike the investigation by Clausen and coworkers (2002), which specifically extracted cyclopentanoids. More analytical details are discussed in chapter 4.

This approach requires biologically active and inactive plant material to find chemical differences between them that may correlate with, and therefore possibly cause, the observed biological difference. Biological activity in this specific case can be defined as a *Rinorea* specimen or leaf that is accepted as host or oviposition site by a *Cymothoe* female. Conversely, a biologically inactive specimen or leaf is one that is rejected. It is known that plants display a great variation in secondary metabolites between clades, species, (cultivars,) populations, individuals, developmental stages or even times of day (Schoonhoven, Van Loon & Dicke 2005, Moço, 2007). All of these causes of chemical difference may or may not affect host choice by butterflies, but are highly likely to affect the metabolic profiles of the sampled plants (Ric de Vos, personal communication). Ideally in metabolomics research, comparisons are made between plants – usually of the same species – that have been grown in a controlled setting, and only varying in the experimental variable. Growing *Rinoreae* was not feasible during this MSc thesis, so I had to use wild specimens on which I had observed oviposition/acceptance (bioactivity) or rejection (bio-inactivity) by *Cymothoe* females *in situ*, i.e. the Ghanaian jungle. Because I am interested in the difference between bioactive and bio-inactive plants, I had to make choices as to the levels of comparison in order to be sure of the minimum bias in the data.

Comparison between clades/families/genera - In Clausen et al (2002) a comparison is made between Achariaceae, Passifloraceae, Turneraceae and *Rinorea*. They found a similar class of compounds, namely cyclopentanoids which in many cases are cyanogenic (though not in *Rinorea*), in all of these plant clades. Certain groups of Nymphalid butterflies have colonized these plant clades and on this basis they hypothesized these cyclopentanoids to act as host recognition chemicals to the butterfly females. A key point however, is the high level of specificity that *Cymothoe* display when selecting their host. Many will accept a single species of Rinorea, but not another that occurs in sympatry. When studying *Cymothoe* speciation and the development of host specialization in this genus by investigating the specific host recognition displayed by *Cymothoe* species, it would therefore be more informative to investigate the differences in chemistry between *Rinorea* species, rather than the similarities between *Rinorea* and related plant taxa.

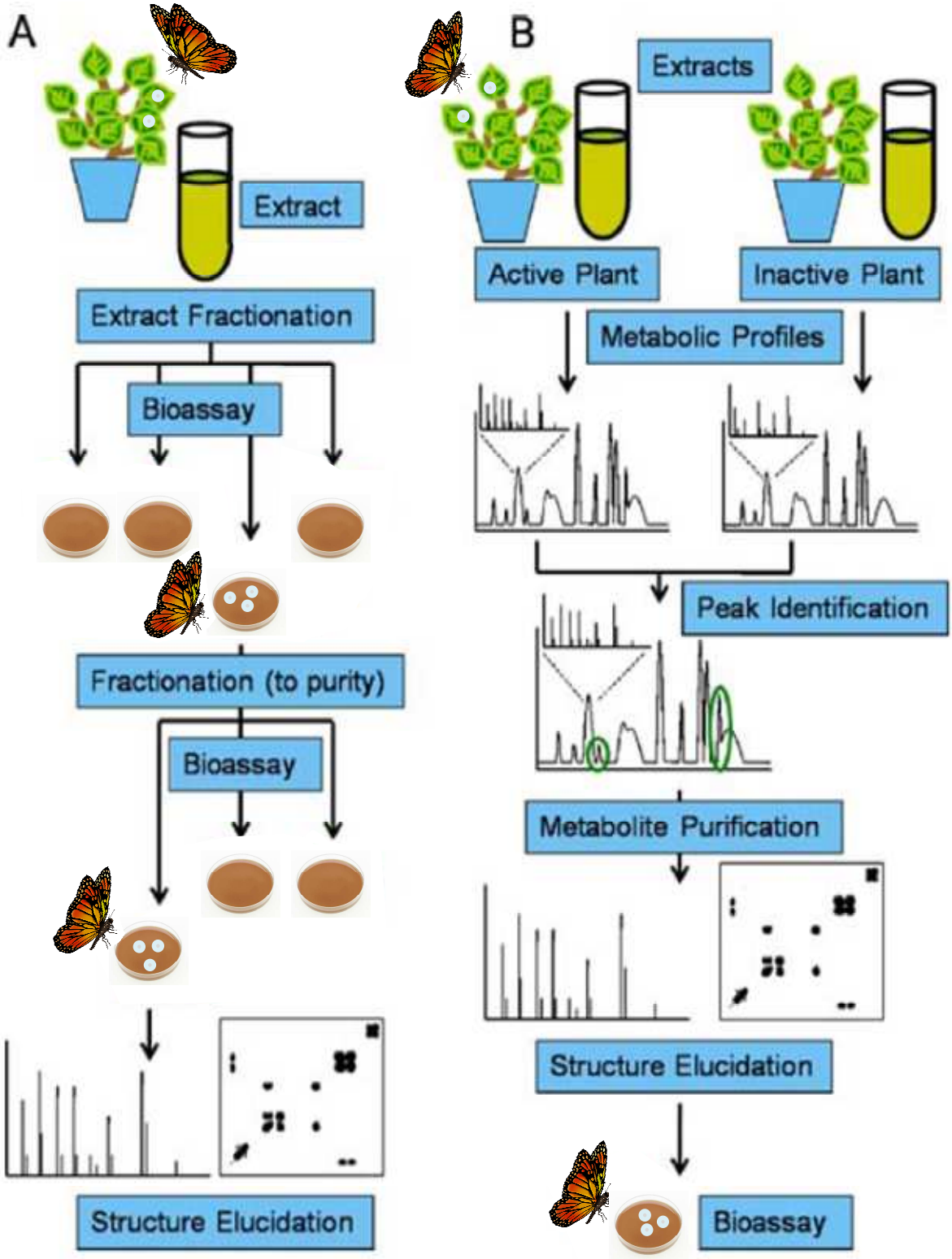


Figure 1.6 Schematic representation of two methods for identifying ecologically relevant plant compounds. Bioassays are shown here as choice experiments, using agar as an artificial "leaf" surface, such as I have attempted in this study. **A:** An iterative, bioassay-guided fractionation approach. **B:** A metabolic profiling approach, chemically comparing extracts from bioactive and bio-inactive plants. Above Peak Identification, a schematic view of a metabolic profile is depicted, above Structure Elucidation a schematic view of H NMR peaks and a 2D H NMR spectrum. This study has attempted to follow B up to Metabolite Purification, as well as perform the Bioassay for CPG. Figure adopted from Prince & Pohnert (2010) and adapted to depict the search for butterfly oviposition stimulants.

Comparison between Rinorea species (interspecific variation) - In the case of a bioactive *Rinorea* that is used by a *Cymothoe* species as a specific host, an inactive variant could either be a related *Rinorea* that is not a host to this particular butterfly, or a specimen of the host species that is rejected. The high specificity in host recognition by Cymothoe butterflies implies a difference in chemical make-up of these related *Rinorea* host species which can be minor, but is certainly consistent, otherwise the butterflies could not distinguish between hosts and display such consistent host preferences. This level of comparison can therefore in principle be used to identify secondary metabolites that have some biological function to *Cymothoe*. However, because this indiscriminate method targets the entire plant metabolome, directly comparing metabolic profiles of different *Rinorea* species will undoubtedly show numerous chemical differences, most of which will probably not be involved in host recognition (Maarten Jongsma en Ric de Vos, personal communication). If interspecific variation is to yield likely candidates for oviposition stimulants, a more phylogeny-guided approach involving oligophagous *Cymothoe* and multiple *Rinorea* species may be an option to narrow down the number of these candidate molecules.

Because Nymphalid butterflies are reported to be stimulated to oviposit by a single chemical, sometimes occurring in different plant genera (Honda, 1995), I do not expect an oligophagous *Cymothoe* species to be attracted by different chemicals in different *Rinorea* hosts. Because of the high specificity of the host-insect relationships, any identical peak found in the metabolic profiles of both bioactive *Rinorea* species will therefore represent a compound that could theoretically be used for host recognition. Furthermore, because we have knowledge of the *Rinorea* phylogeny, a comparison can be made with the metabolic profile of a related *Rinorea* species that occurs in sympatry, but is not oviposited upon (ecologically inactive). Peaks present in the metabolic profiles of the ecologically active, but not in those of inactive *Rinorea* species represent likely candidates for oviposition stimulants.

Unfortunately, no such oligophagous associations have been observed in Ghana by Renske Onstein during her field study in 2010, see figure 1.7, and I do not expect to be able to study this using only material gathered during this study in Ghana. The fact that Renske Onstein did not observe for instance *C. coccinata* or its hosts does not mean that these do not occur in Ghana at all, as they have been reported there (Larsen, 2005, Hawthorne & Jongkind, 2006). During my fieldwork, I kept on the lookout therefore, but I did not find these associations.

In Cameroon, Robin van Velzen reported several such oligophagous associations and he collected a large amount of *Rinorea* material in Cameroon, namely dried leaf material and alcohol surface swabs which may yield valuable information. Because there is greater species richness of both *Rinorea* and *Cymothoe* in Cameroon than in Ghana (Larsen 2005, Van Velzen 2007, Onstein 2010), more of the associations from table 1 occur there. For many of these samples, Robin van Velzen has observed oviposition. Appropriate oligophagous associations together with a related non-host can be inferred from the phylogenetic reconstruction of *Rinorea* as shown in figure 4. One such association is *R. welwitschii*, *R. rubrotincta* and *R. longicuspis*, all of which are host to *C. sangaris*. Metabolic profiles of these samples can be compared to those of related non-hosts such as *R. kamerunensis* and *R. leiophylla*. Another such association is the *C. fumana* hosts *R. oblongifolia*, *R. amietti* and *R. longisepala* coupled with non-hosts *R. verrucosa* and *R. dewildei*. The dried samples from Cameroon may still be used for such comparisons. **Recommendation:** analyze these samples and investigate the possibility of performing interspecific comparison in Cameroon.

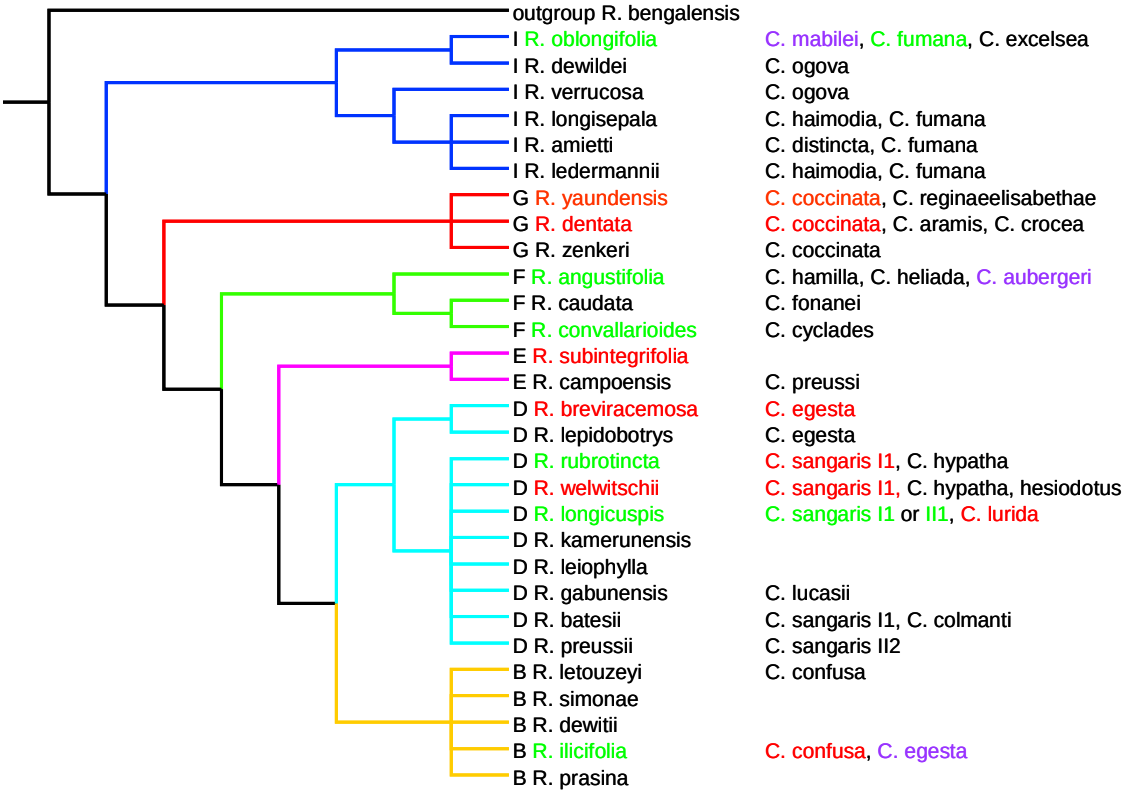


Figure 1.7 Schematic cladogram of *Rinorea*, based on the nrDNA ITS +indel +PRANK sequences 50% majority rule consensus tree resulting from Bayesian analysis, from the thesis report of Onstein (2010). Indicated are the clades (B, D, E, F, G and I) which are color coded, and the *Cymothoe* which feed on these *Rinorea* (Onstein, 2010). Coloured taxa indicate the Ghanaian insect-host associations with green being the expected and found *Rinorea* and associations, red the expected but not found *Rinorea* species and associations and purple the unexpected observations of species and associations from Onstein 2010. Of the *Rinorea* listed, Robin van Velzen has gathered a sample from Cameroon which may be used in a chemical analysis.

Comparison between populations - There appears to be some level of difference between populations of *Rinorea*, with respect to oviposition preference of the *Cymothoe* that feed on them, judging by the reports of the Cameroonian by Amiet & Achoundong (1996) and Van Velzen (2006), which are somewhat in conflict with the report by Onstein (2010) of the Ghanaian population, see chapter 1.6. In theory, this difference may be caused by a chemical difference between the populations that can be tested, however it may also be caused by a difference in the butterfly's preference. Also, differences found from comparisons between populations are likely to be affected by sampling time and growth conditions in different areas. Moreover, different populations are likely to have a different genotype, which is the case for *Rinorea ilicifolia* (Van Velzen, personal communication), which is very likely to affect the chemical contents of a plant. These considerations, as well as practical problems faced when trying to sample plants in different countries, caused me not to choose this level of comparison for metabolomics. Onstein (2010) reported host specificity to be consistent among different populations within Ghana, which is therefore also not a basis for finding a basis for differential host recognition.

Comparison between individuals of the same species in sympatry - Intraspecific variation in secondary metabolites could be responsible for differential host selection by specialist herbivores (Poelman et al., 2009). For example, Stermitz and coworkers (1989) found that ovipositing females of *Euphydryas editha* (Nymphalidae) can discriminate between individuals of their host, the hemiparasite *Pedicularis semibarbata* (Lamiales, Orobranchaceae), individuals of which were consistently accepted or rejected. This accepted or rejected state of *P. semibarbata* individuals persisted from year to year, showing clear intraspecific variation (Stermitz et al., 1989). Similarly, there are anecdotal

reports by Van Velzen (2007) and Onstein (2010) of individuals of the same *Rinorea* species, growing in close proximity, where one was frequently visited and oviposited upon by *Cymothoe* females, whereas its neighbor was not. If such within-population differences in oviposition behavior by female butterflies can be observed in the field, untargeted metabolomics might offer the opportunity to compare the secondary metabolites of these specimens directly, and identify compounds that correlate to the observed butterfly behavior.

I used the *R. ilicifolia-C. egesta* association for this analysis, because I also use it for testing the stimulatory effect of CPG and Onstein reported these species to be especially common in Kakum National Park, making this a convenient model association for testing for intraspecific chemical differences. Because this level of comparison uses individuals of the same species of the same population, the main variable will be observed butterfly behavior. Although there will still be variation in day to day conditions and time between sampling and analysis, this is unavoidable when sampling plants in the field. I have therefore used untargeted metabolomics on *R. ilicifolia* leaf samples for intraspecific comparison, to select hypothetical candidate compounds for oviposition stimulants to *Cymothoe egesta*.

1.8 TESTING OF CPG AS OVIPOSITION STIMULANT

Once candidate compounds have been identified, using methods such as the untargeted metabolomics method described above, a behavioral experiment to actually test the biological effect of the purified chemical must be set-up (Prince & Pohnert, 2010). Because Clausen et al. (2002) proposed CPG as such a candidate compound, I devised a behavioral experiment to test its stimulatory effect on *Cymothoe egesta* oviposition. The method of experimentation I used was modified from Hovanitz & Chang (1964) and Pereyra & Bowers (1988) and involves a choice experiment using agar plates as artificial substrates (i.e. artificial leaves). These can contain either dissolved CPG, ground *R. ilicifolia* leaf material as positive control or nothing (except green colorant) as a negative control. See chapter 4.3 for this experiment.

Oviposition experiments are normally performed with reared female butterflies that have mated in captivity (for examples, references of table 1.1). This ensures reproducibility of the experiment and allows for controlling the age of the female butterflies, which relates to the number of eggs that are carried. Also, exposure to a wild environment may have influenced the butterfly by conditioning to natural hosts, which may prevent them from ovipositing on artificial substrates (Schoonhoven, Van Loon & Dicke, 2005). Rearing *C. egesta* is not expected to be feasible within this project, however, since it will take a lot of time. Despite the problems mentioned above, using wild-caught females for a behavioral experiment is therefore the most realistic approach to this experiment (Van Loon, personal communication). Although age difference affects the number of eggs a female can lay, age differences should be randomly distributed among different treatments. Also, the effect of learning in butterfly oviposition behavior has been disputed. Although associative learning has been observed in Pieridae (Traynier, 1984), host preference appears to be completely innate in Papilionidae (Heinz & Feeny, 2005) and Nymphalidae (Parmesan, Singer & Harris, 1995; Kerpel & Moreira, 2005).

Using wild-caught females does require the presence and capture of ovipositing *C. egesta* females in the field. Fortunately, Renske Onstein has observed many of these in Kakum National Park during her fieldwork in Ghana in October and November of 2009, and she informed me *Cymothoe* numbers appeared to increase towards the end of her expedition. For this reason, I went to this location during November and December of 2010, to maximize my chances of catching ovipositing *C. egesta* females.

Apart from testing this hypothesis by means of a behavioral experiment, I performed an indirect test of CPG's effect by measuring levels of this chemical in the samples used for the untargeted metabolomics. Pereyra & Bowers (1988) showed that the Buckeye Butterfly *Junonia coenia* (Nymphalidae) was able to distinguish between slight differences in concentration of its oviposition stimulant, iridoid glycosides, ovipositing on the higher levels. Higher levels of CPG in accepted leaves than in rejected leaves of *R. ilicifolia* indicates this chemical may be involved in host recognition. If no such correlation is found, CPG is less likely to be an oviposition stimulant to *Cymothoe egesta*.

1.9 AIMS AND HYPOTHESES

The main aim of this investigation is to uncover the mechanism behind, i.e. the proximate cause of host specificity in the *Rinorea-Cymothoe* host-herbivore association. It is assumed that differences in secondary metabolites between *Rinorea* species or individuals are the basis for *Cymothoe* host discrimination. Ghanaian populations of *Rinorea ilicifolia* and *Cymothoe egesta* are used as a model association in this study, for two distinct, yet intimately linked experimental approaches:

1. Untargeted metabolomics
   Aim: to compare metabolic profiles of *Rinorea ilicifolia* leaves, in order to identify plant metabolites that correlate with observed oviposition behaviour of *C. egesta* females.
   Premises: chemical differences between *Rinorea ilicifolia* individuals cause differential host selection by *Cymothoe egesta*. Metabolites present in greater abundance in accepted than in rejected leaves are potential oviposition stimulants to *C. egesta.*
2. Testing of 2-(3'-cyclopentenyl)glycine (CPG) as oviposition stimulant
   Hypothesis: CPG is an oviposition stimulant to *Cymothoe egesta*
   Experiment A: to conduct a behavioral experiment testing for the stimulatory effect of pure CPG on oviposition by *C. egesta* females.
   Experiment B: to test whether levels of CPG are higher in *R. ilicifolia* leaves accepted by *C. egesta* females than in those rejected.

For a visual overview of the present study, see the conceptual and experimental framework in figure 1.8.

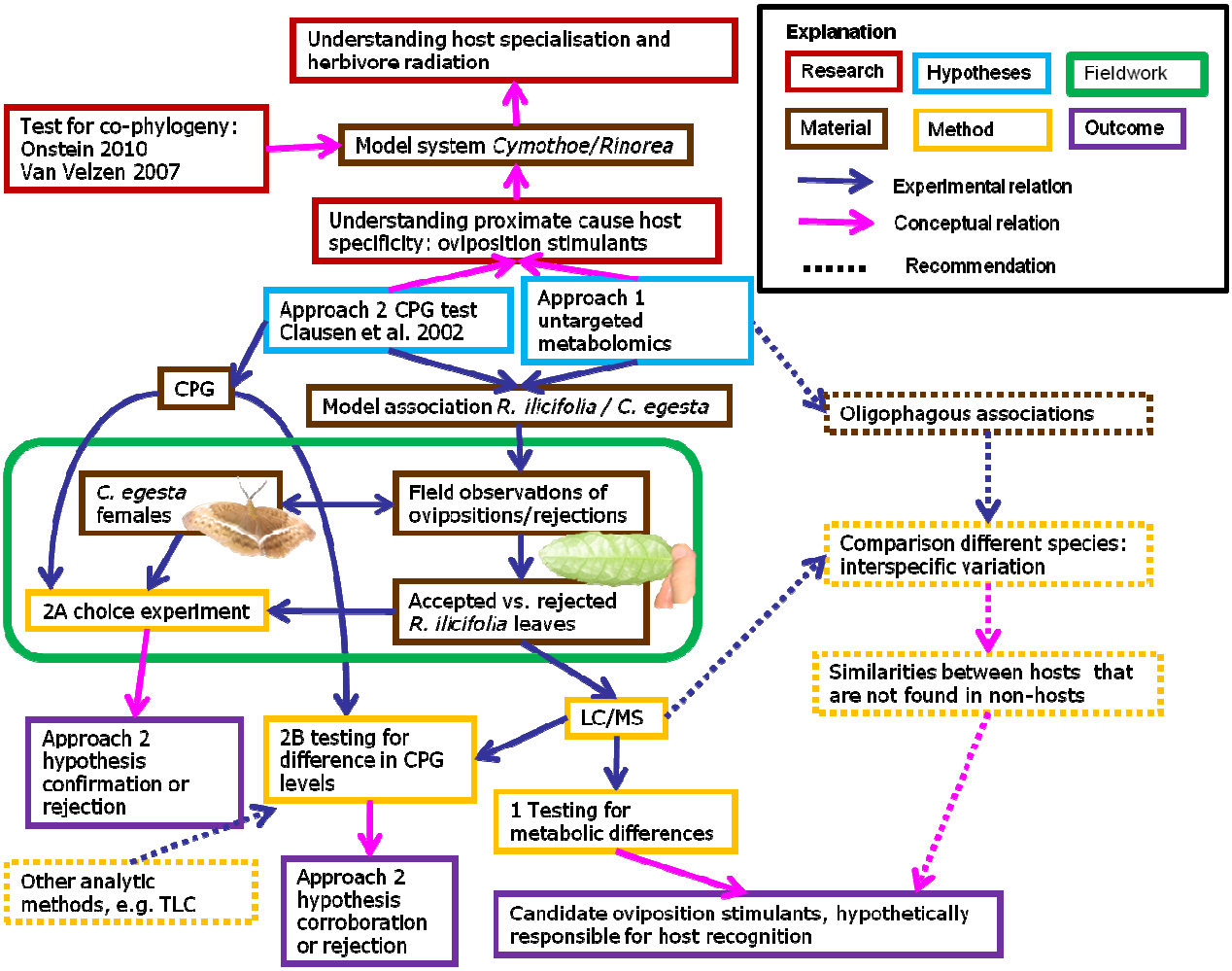


Figure 1.8 Conceptual and experimental framework of this thesis

Chapter 2 Field work: set-up and observations

2.1 SET-UP AND LOCATION

Fieldwork took place in Kakum National Park, Central Region, Ghana, see figure 2.1. There were several reasons for this choice of locality. First, Clausen and coworkers discovered 2,3-CPG in *R. ilicifolia* leaf material from Ghana and hypothesized on its effect as oviposition stimulant based on the presence of similar chemicals in related plants that are host to butterflies related to *Cymothoe*. As Renske Onstein discovered during her field study in Ghana in 2009, the *Cymothoe* species that feeds on *R. ilicifolia* in Ghana is *C. egesta*. In order to test the hypothesis of Clausen and coworkers using a behavioral assay, this assay should be performed using *C. egesta*. A second reason for this field work to take place in Kakum NP is therefore that Renske Onstein reported this species, as well as *R. ilicifolia*, to be particularly abundant in Kakum. Also, she made several local contacts, who were willing to help me as well. She had collaborated with the Ghanaian Butterfly Conservation Society (BCGhana), which allowed me to contact Charles Owusu, a student associated with BCGhana. After arriving in Accra, I stayed there for several days preparing the journey to Kakum, with the help of Charles, who introduced me to the

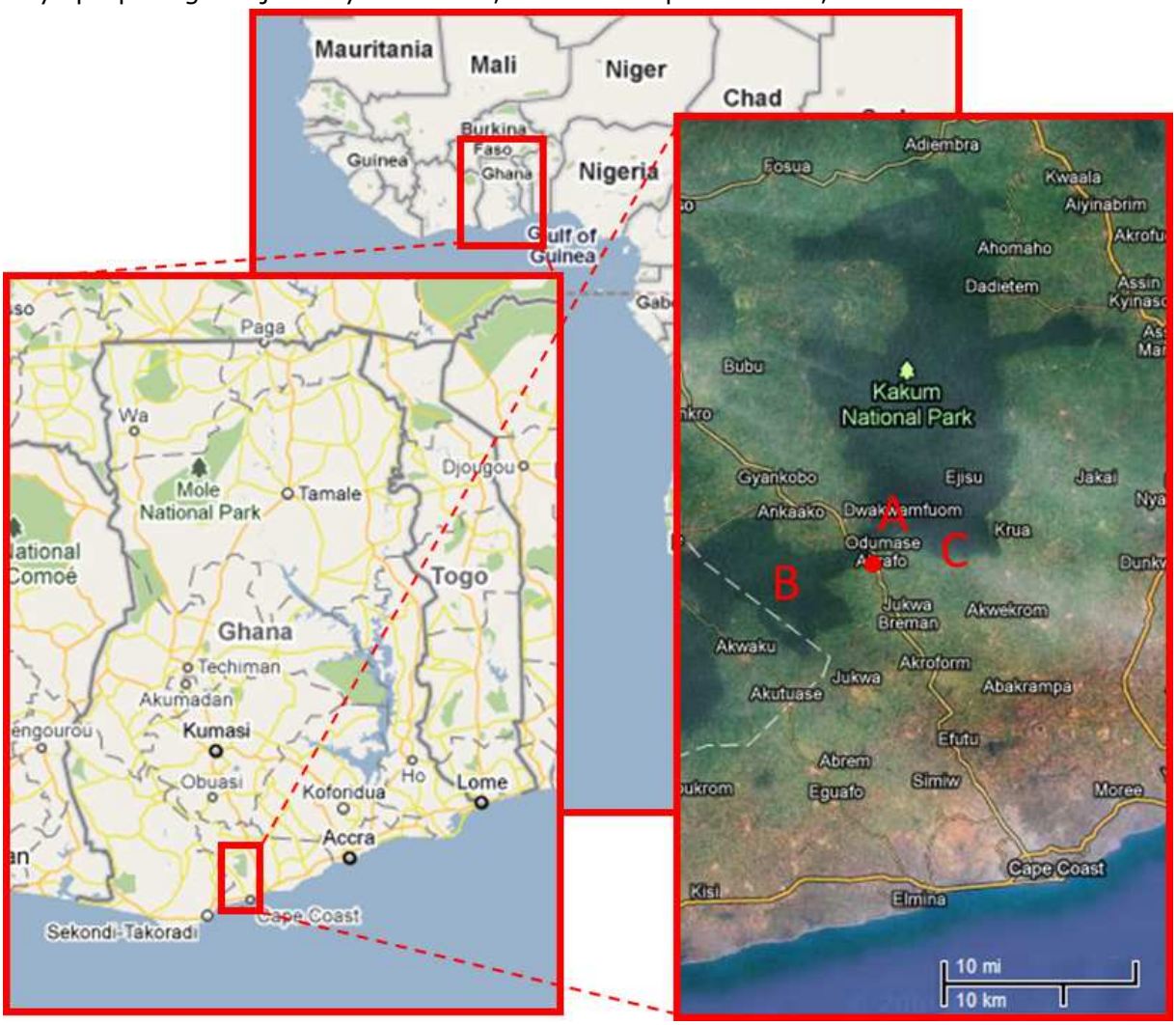


Figure 2.1 Overview of research area. Ghana is located in West Africa under the "armpit" of the continent. It is bordered by Ivory Coast in the West, Burkina Faso in the North and Togo in the East, with the Gulf of Guinea to the South. Kakum National Park is in the Central Region, west of the country's capital Accra, with Cape Coast as the nearest sizable town. The red dot indicates (Odumase) Abrafo, where I stayed in a guestroom with the familiy of my guide Andrews. **A** indicates the "Canopy Area" of Kakum NP where most research was performed, **B** indicates the Forest Reserve and **C** indicates "Simon's Place", which borders Kakum.

head of BCGhana, Safian Szabolc, who was kind enough to lend me his research permit.

At Kakum NP, Renske had been assisted by the wildlife guide Andrews Kankam, whom she recommended to me for his skill as a guide and knowledge of the *Rinorea* and *Cymothoe* present at the locality and who became my personal guide as well. I had initially expected to stay at the park lodge, but was persuaded to stay in the nearby town of Abrafo, in the home of Andrews' father, Richard Kankam. This guesthouse was thereafter used as my base of operations and had a garden that contained several trees where I could conveniently hang the butterfly cages for experimentation. Clean drinking water from a borehole pump was readily available, as were electricity and a gas heater needed for behavioral experimentation (see chapter 4.3).

I established a routine in which I went on regular field expeditions together with Andrews, gathered leaves that were accepted or rejected as host plants by ovipositing *C. egesta* females for further chemical analysis, and tried to catch these butterflies. I noted down my observations as detailed as possible. Leaf samples were carried to my guesthouse in plastic bags and subsequently dried in a plant press and subsequently stored (see chapter 3.1). Butterflies were caught and initially taken to the guesthouse in a pop-up cage for experimentation, a method that was later adapted to experimentation inside the forest itself, see chapter 4.3.

Most of the field observations, sample gathering and butterfly catching was performed around the "canopy area" (see figure 2.1). This section of secondary rainforest is part of the park near the visitor's centre that is open to the public and has the famous Canopy Walkway as a major tourist attraction. In spite of the frequent disturbance by visitors travelling the paths in this area, I found *Rinorea ilicifolia to be* abundant here, as well as *Cymothoe egesta*. An advantage was that this area could be accessed without having an armed guide present as opposed to the rest of the park, and it was at walking distance from my guestroom in Abrafo. During expeditions to "Simon's place" or deeper into Kakum, I did not find any *R. ilicifolia* and *C. egesta*. In the Forest Reserve (B in figure 2.1), which is opposite Kakum and separated from the park by the Cape Coast to Twifo-Praso road, I did find several species of both *Rinorea* and *Cymothoe* (including *C. egesta* and *R. ilicifolia*) which I expect belong to the same Kakum population. This area was however less accessible than the "canopy area" of Kakum NP, as well as freely open to local people who regularly hunted and gathered food here, which is why I did not use this area for experimentation.

2.2 OBSERVATIONS OF CYMOTHOE EGESTA (OVIPOSITION) BEHAVIOR

Overall, numbers of *Cymothoe* were lower than expected, presumably because of the frequent rain. The timing of the fieldwork was chosen because of seasonality and Renske Onstein's advice. She experienced an increase in *Cymothoe* by the end of her field study halfway November as the dry season started and rains became less frequent. For this reason, I chose to visit during November and December, as the weather was expected to be mostly dry during this period. Unfortunately, rains persisted throughout the month of November which reduced observed butterfly numbers. The weather therefore limited the number of observations of *Cymothoe* females ovipositing on *Rinorea* and the number of butterflies caught. Some field days were so rainy that no *Cymothoe* were observed in ovipositing behaviour at all. By December, rains had mostly stopped and this meant an increase in the observed number of ovipositing *Cymothoe*. This increase persisted until the end of my fieldwork towards the end of December. **Recommendation:** For a follow-up investigation, I would advise any field work such as this in (southern) Ghana not to be performed until mid-November.

Typical observations were made in patches that had a relatively high density of *R. ilicifolia* specimens. Several such patches were present in the "canopy area" near the park entrance (figure 2.1 A). Once in such an area, my guide Andrews and I would oftentimes simply wait for the female *C. egesta* to appear. Next, I assessed whether they were displaying one of three behaviors: feeding, sunning or host searching, the latter including rejection and acceptance/oviposition. Recordings of such behaviors are listed in table 2.1, with emphasis on the latter behavioral type.

Butterfly feeding

Cymothoe egesta feed mostly on fallen fruit. This can be observed in areas with a ripe fruit bearing tree or epiphyte, such as a strangler fig *Ficus sp.* These will litter the ground with fallen, rotting fruit attracting many types of frugivorous butterflies. I observed that during *Cymothoe egesta* food searching behavior in such a patch with an abundance of fallen fruit, the butterfly will take off from the ground and land a short distance away, then probing with its tongue for a fruit and repeat this sequentially until it happens to land next to a bit of food. Speculatively, smell is enough to attract the butterfly to a general area of fallen fruit, but contact chemoreception is necessary to find the actual fruit. In that sense food searching behavior is not unlike host searching behavior in the sense that alighting appears to be a fairly random process, speculatively guided by a sensory cue like smell or vision that is repeated until acceptance of the foodstuff or host-plant respectively.

Sunning

Butterflies will often fly from leaf to leaf in a manner similar to host searching behaviour, in order to find a sunny place to warm up, especially after a rainy period. However, in contrast to host searching behaviour, butterflies will stay perched on leaves for longer periods (many minutes if left undisturbed), flexing their wings in the sunlight, rather than quickly taking off and alighting.

Host searching behavior including rejection and acceptance/oviposition

Host searching behavior was quite distinct, as the female would quickly flutter from leaf to leaf in the undergrowth, as described earlier by Renske Onstein for Ghana and Robin van Velzen for Cameroon. Usually, this behavior occurred more in the afternoon, presumably as most moisture (from dew or rain) had evaporated from leaf surfaces and the butterflies had sufficiently increased their body temperature. In Renwick & Chew (1994), the sequence of events leading up to oviposition in Lepidoptera is subdivided in searching, orientation and encounter of a possibly suitable host, followed by landing (alighting), contact evaluation and subsequently acceptance or rejection.

Alighting on a leaf by a female butterfly is apparently guided by visual stimuli (green leaves) rather than host specific stimuli (e.g. volatiles), as also non-*Rinorea* leaves were visited and clearly tested by *C. egesta* females by means of tarsal drumming. Drumming was also clearly audible, and possibly even more clearly so on non-*Rinorea*, in which case drumming was invariably followed by rejection. Most *Rinorea ilicifolia* leaves visited by *Cymothoe egesta* females were observed to be rejected. As table 2.1 lists, I recorded a total of 18 ovipositions and 16 rejections, however, by far not all rejections were recorded. Such statistics were very hard to gather as I was doing several things at once, namely trying to remember which leaves were alighted on whilst observing where the female was going and whether she was laying an egg or not and then trying to catch the female if she did, for use in the behavioral assay. Also, if no oviposition has occurred before a female has flown out of sight, it is difficult to be certain she had been fertilized and had actually been searching for a host, making any observed rejections of limited credibility. Moreover, the behavioral sequence leading up to oviposition was found to be somewhat more complicated than was originally anticipated.

A sequence of several observations led to me to determine the actual behavioral sequence of *C. egesta* oviposition behavior. Oftentimes, a female would oviposit not on or under a leaf, but rather on the stem or branch of the accepted host plant. Renske Onstein had observed this behavior once during her field work in Kakum NP, and photographed a *C. egesta* female ovipositing on the stem of an *R. ilicifolia*. This behavior was also described by Robin van Velzen, but only for *Harma theobene* in Cameroon, which he reported as a striking difference between the generalist *H. theobene* and *Cymothoe*.

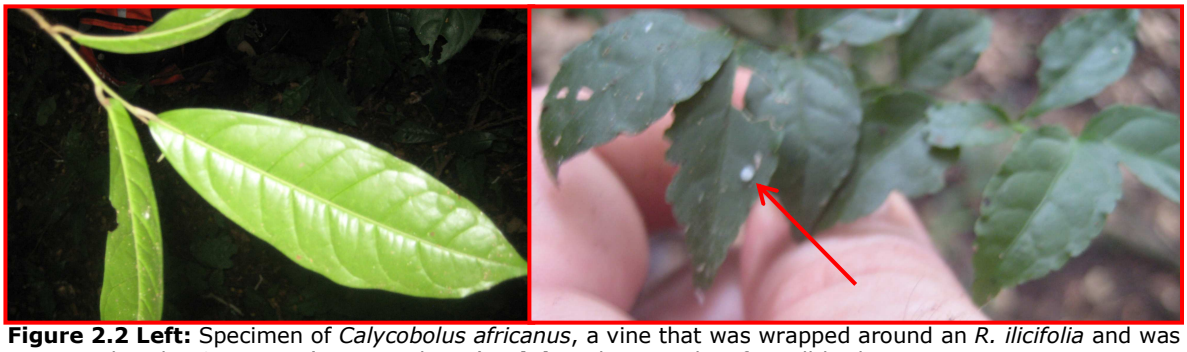


Figure 2.2 Left: Specimen of *Calycobolus africanus*, a vine that was wrapped around an *R. ilicifolia* and was oviposited on by *C. egesta* (egg not shown). **Right**: Photographs of small herbaceous non-Rinorea growing under *R. ilicifolia* that was oviposited by *C. egesta*. The red arrow indicates the egg. Picture is somewhat out of focus due to bad lighting and quality of camera.

One of the most striking observations, was the large number of apparent "mistakes" made by the ovipositing females. As I have witnessed three times, an egg was deposited by a *C. egesta* female on a non-*Rinorea* plant that grew in the vicinity of *R. ilicifolia* specimens. The first time involved a small herbaceous plant, see figure 2.2 right, growing underneath an *R. ilicifolia* specimen, the latter being approximately 40 cm tall. Clearly, the *Rinorea* had been accepted by the female, which subsequently took flight again and landed on this different plant where she oviposited. A second case involved a vine that was wrapped around an *R. ilicifolia* specimen, see figure 2.2 left. My guide Andrews observed that an *R. ilicifolia* individual was alighted upon by the female prior to making the ovipositing "mistake". The vine was later identified, with the aid of Carel Jongkind, as *Calycobolus africanus* (Convolvulaceae) a species that is unrelated to Rinorea and is no host to *Cymothoe*. A third incident was somewhat similar to Renske Onstein's observation of an oviposition at the base of a stem of *R. ilicfolia*, although I observed a female oviposit at the base of the stem of another very different shrub. Although I was tracking this female through the undergrowth at the time, I did not clearly observe the actual acceptance of an *R. ilicifolia* leaf prior to oviposition, however this was presumably the case because this plant did grow in a patch of vegetation where *R. ilicifolia* grew in abundance.

Upon closer inspection, the sequence of events leading up to oviposition by *C. egesta* females – at least in the Kakum/Forest Reserve population – follows the sequence

described in Renwick & Chew (1994) up to the point between acceptance and oviposition. Almost invariably, oviposition by *C. egesta* is preceded by alighting and drumming on a young, apical *R. ilicifolia* leaf, which – should the young leaf be accepted – is followed by the butterfly taking off and ovipositing somewhere nearby. Usually the oviposition itself would occur on an older leaf of the *R. ilicifolia* individual of which the young leaf was accepted, although non-*Rinorea* species growing below- or vines entwining an *R. ilicifolia* specimen could also be oviposited. Oviposition on the stem, which Renske Onstein reported and photographed in Kakum NP was also observed several times in this study. Acceptance is therefore clearly very host specific, whereas the actual oviposition that follows it does not appear to be guided.

For the rest of the experimentation, the significance of this "extra step" in the oviposition sequence is that a clear distinction has to be made between a site that is accepted by a butterfly and one that is oviposited upon. For the metabolomics approach, for example, not the oviposited leaves, but rather the leaf visited just prior to oviposition, i.e. the leaf that was drummed upon, should be gathered. This distinction can only be made by closely observing butterfly behavior, as this sequence of events takes place within fractions of seconds. This phenomenon has also caused the discrepancy between the numbers of recorded ovipositions (18) and acceptances (15), because only the former leaves clear, lingering evidence, i.e. an egg.

Recommendation: Observe the behavioral sequence of host selection very carefully and look very close to the actual leaf that is accepted, it may not be the same leaf that was oviposited.

Tabel 2.1 Assorted field observations, butterfly catches and leaf samplings

MM/DD	Observation	Sampled/ caught	Remarks
11/06	Found many *R. ilicifolia* and one *C. egesta* ♀ displaying egg laying behavior		First time into Kakum NP. Also observed oviposition* by *C. mabilei* on *R. angustifolia* and caught a *C. aubergeri**
11/07	No *C. egesta* or *R. ilicifolia*		Into privately owned forest adjacent to Kakum, C in figure 2.1. Observed *R. oblongifolia* but not *R. ilicifolia*. Saw one *Cymothoe*
11/08	Observed *C. egesta* and *R. ilicifolia*	♂1	Coordinates: N05° 21.164' W01° 23,131'
		♀1	Coordinates: N05° 21.164' W01° 23.014'
11/11	6 rejections on older leaves*	1_rej_o	Filmed this observation
	2 rejections on older leaves*	4_rej_o	Filmed this observation. Coor: N05° 21.210' W01°22.996'
	C. egesta caterpillars*		
	Acceptance and oviposition	2_acc_y ♀2	Oviposition on non-Rinorea*, stored on silica. Also collected egg, stored on alcohol (ploof). Coor: N05° 21.205' W01° 23.001'
	Acceptance and oviposition	3_acc_y	Egg on 3_acc_y, already egg on it*. Collected both eggs.
11/13	Males and females feeding on fallen fruit, no oviposition		I dug out a small *R. ilicifolia* for use in behavioral experiment. Not observed any oviposition on it.
11/16	1 rejection*		Filmed this observation
	C. egesta caterpillar from 11-11*		Had eaten a leaf
11/18			Took longer hike into Kakum, no *R. ilicifolia* found deeper into the forest.
	C. egesta caterpillar from 16-11*		Had eaten another leaf
11/19	No oviposition, another caterpillar		Not many butterflies, too much rain for the time of year
11/20	Oviposition not acceptance by Andrews		Went to the Forest Reserve to the southwest of Kakum NP, many *R. ilicifolia*. Egg on non-*Rinorea* vine wrapped around *R. ilicifolia*. Stored vine on silica and collected egg as well. Coordinates: N05° 20.722' W01° 23.597'
11/24	Acceptance and oviposition	5_acc_y	Egg on older leaf, collected this as well. Coor: N05° 21.210' W01° 22.992'
		6_acc_y	Oviposition on this accepted, young leaf itself
11/25	Acceptance, oviposition and rejection by ♀3	7_acc_y 8_rej_o ♀3	Acceptance of younger leaf but oviposition on older leaf
11/27	Observed *C. egesta* females feeding on fallen fruit	♀4 ♀5	Dug out live *R. ilicifolia* with young leaves. Apparently vegetative reproduction
11/30	Oviposition, not acceptance		Near ground on stem of non-Rinorea
	Oviposition not acceptance		On *R. ilicifolia*
	Acceptance, oviposition and rejection	9_acc_y 10_rej_y	Oviposition on older leaf, accepted young +6. Rejected young -4 but unsure whether drumming had occurred
12/03	Feeding on coconuts by ♀6	♀6	
12/06	Acceptance and oviposition	11_acc_y♀7	Oviposition on older leaf
	Acceptance and oviposition	12_acc_y♀8	12_acc_y had two eggs, three more eggs were found on another young leaf, an old leaf and a branch of the same R. ilicifolia specimen
12/07	Observed only feeding on fruit		
12/10	Feeding on fallen fruit	♀9	

Tabel 2.1 continuation

12/12	Rejection, acceptance and oviposition	17A_rej_o 14_acc_y♀ 10	♀10 rejected 17A_rej_o old leaf, then accepted young leaf 14_acc_y and oviposited on older leaf from plant 14
	Rejection	18_rej_y	Young leaf, spotted by Andrews
	Acceptance and oviposition	15_acc_y	Young leaf, egg on older leaf
	Rejection, acceptance and oviposition	19_rej_o 16_acc_y	Rejected 19_rej_o old leaf, then accepted 16_acc_y on nearby *R. ilicifolia*. Egg on old leaf
12/13	Acceptance and oviposition	20_acc_y	Egg on stem of same plant
	Acceptance and oviposition	21_acc_y	Egg on older leaf
	Acceptance and oviposition	22_acc_y	Young leaf, but big enough so egg on 22_acc_y
12/15	Rejection, acceptance and oviposition	23_acc_y2 4_rej_o 25_rej_y	Several older leaves and one younger leaf were rejected by the same ♀. She then accepted young 23_acc_y, oviposited on an older leaf of the same Rinorea
12/16	Rejection, acceptance and ovipositions	26_acc_y 27_acc_y 28_acc_o 29_rej_o	Observed several rejections, then acceptance of 26_acc_y, which had approx 20 eggs already from another insect (possibly a butterfly), oviposition by ♀ C. egesta on older leaf. Then ♀ continued egg laying behavior, rejecting several including 29_rej_o, then accepting 27_acc_y and ovipositing on the stem, then continuing again and accepting 28_acc_o ovipositing on older leaf

2.3 MISCELLANEOUS OBSERVATIONS

I made some observations that are not directly related to oviposition behavior, but which I feel are worth mentioning.

Caterpillars

I observed several *C. egesta* caterpillars during fieldwork, invariably on *R. ilicifolia*. I observed and photographed one individual in particular several times over the course of 8 days, during which time it stayed on a single one *R. ilicifolia* individual. After this time it went missing, either having migrated to another *R. ilicifolia* to continue feeding or to a place to pupate, or having been eaten by a predators like driver ants (*Dorylus*) that regularly patrolled the area. During this time, two of the *R. ilicifolia* individual's leaves disappeared, see figure 2.3, presumably eaten by this individual caterpillar. Strikingly, these were older leaves rather than young leaves. In the case of newly hatched caterpillars, Renske Onstein found first instars mostly on the very young apical leaves but I made no observations of these.

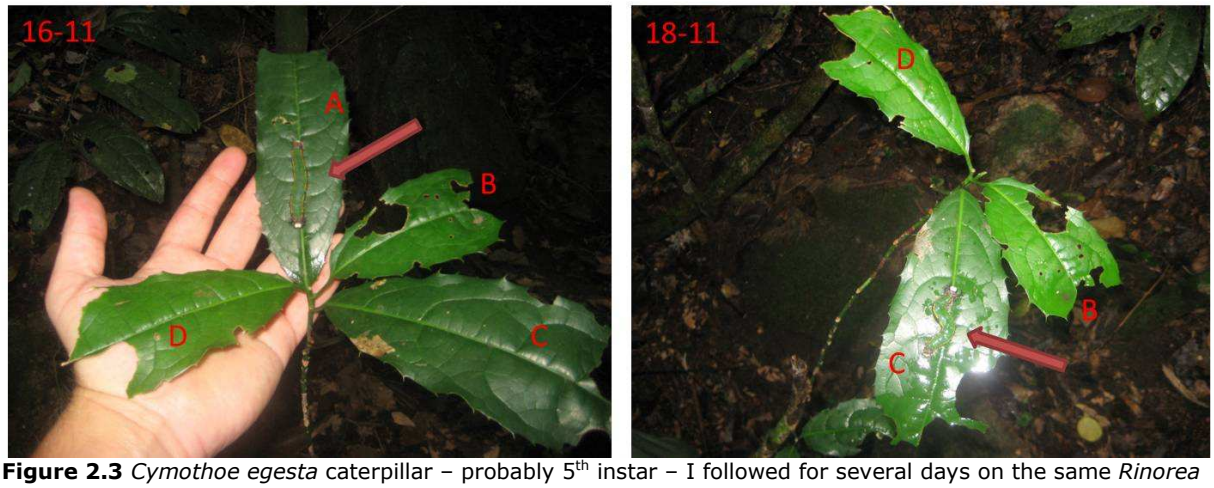


Figure 2.3 *Cymothoe egesta* caterpillar – probably 5th instar – I followed for several days on the same *Rinorea ilicifolia* individual. Leaves are indicated A-D and the position of the caterpillar is indicated by the red arrow. Note that the mature leaf A, on which the caterpillar sits on November 16, is gone on November 18, presumably eaten by the caterpillar. The times at which these pictures were taken are 9:06 and 9:02 respectively.

Rinorea ilicifolia vegetative reproduction

On November 27th, I tried to dig out a small *R. ilicifolia* specimen for use in the behavioral experiment, see table 2.1. It had young, yet fairly large leaves which I thought to be ideal for testing whether a captive butterfly would oviposit on it, see chapter 4.3. I noticed however, this was not an individual specimen but rather an offshoot shoot of a larger shrub, connected with a tough rhizome, see figure 2.3.

Figure 2.4 *Rinorea ilicifolia* shoot showing vegetative reproduction by rhizome formation.

Other herbivores on R. ilicifolia

Figure 2.4 Caterpillars, presumably *Acraea spp.* Found feeding on *R. ilicifolia*

Caterpillars other than *C. egesta* were found on *R. ilicifolia* quite abundantly, see figure 2.4. These presumably belonged to a species of *Acraea* and were often found to feed on *R. ilicifolia* with multiple individuals at once.

R. ilicifolia fruit parasites

I found a *R. ilicifolia* fruit that had been attacked by a grub, which had bored its way into the fruit and was feeding on the seeds, see figure 2.5.

Figure 2.5 *Rinorea ilicifolia* fruit with a parasite, presumably a beetle grub, having bored into it.

CHAPTER 3 UNTARGETED METABOLOMICS

Using untargeted metabolomics – approach 1 in chapter 1.8 – I have tried to essentially "start from scratch" in trying to pinpoint chemicals in *Rinorea ilicifolia* that are used by *Cymothoe egesta* to recognize its host. By comparing metabolic profiles of individual *R. ilicifolia* leaves that have elicited either an acceptance or a rejection response by a *C. egesta* female in ovipositing behavior (see chapter 2), I have looked for metabolites the presence of which in these metabolic profiles correlates with bioactivity, conceptually following figure 1.6B (Prince & Pohnert, 2010).

3.1 METHOD AND MATERIALS

PROCEDURE IN FIELD WORK AND SAMPLING PROTOCOL

I gathered leaf samples by picking individual leaves after either acceptance or rejection of this leaf by a female *C. egesta* had been observed on these leaves, see chapter 2.2. I picked the leaves using scissors or a knife. A label was attached to each leaf indicating the sample code and acceptance/rejection. Subsequently, I took these leaves back to my guesthouse in a plastic zip bag.

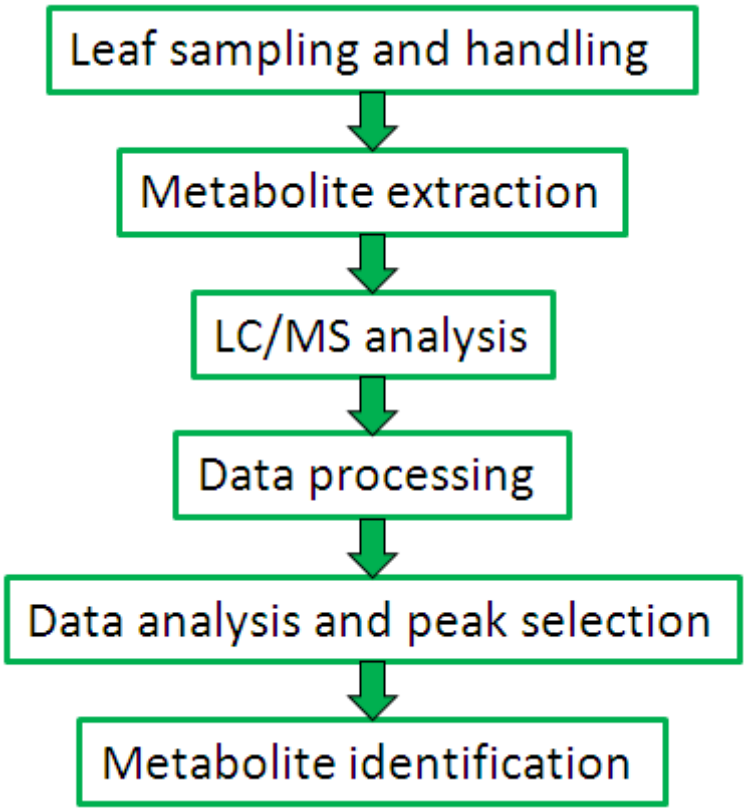


Figure 3.1 Flowscheme summarizing the sequence of events for untargeted metabolomics.

Ideally for plant metabolomics research, a (leaf) sample is flash frozen in liquid nitrogen in order to quench metabolism, before grinding, storage at -80°C and metabolite extraction in the laboratory (Moço, 2007; Maier, Kuhn & Müller, 2010). Unfortunately it was not possible to use liquid nitrogen in the field and store samples at -80°C, nor was I able to perform the LC-MS analysis shortly after the samples were taken. Important for generating a metabolic profile that represents the *in vivo* situation is the quenching of the metabolic processes that alter the chemical make-up of plant material after it has been picked. A method for field sampling followed by a delayed laboratory processing of plant material for metabolic fingerprinting was therefore required. The only such I found in the literature, is described by Maier, Kuhn & Müller (2010). Their method utilizes fresh leaves that are shredded in a falcon tube with a mixture of methanol and dichloromethane using a handheld disperser. However, because methanol and dichloromethane are both volatile and toxic, these solvents were not expected to be either suitable or readily available for use in Ghana. Instead, because all metabolic processes require water and drying therefore of leaf material effectively quenches metabolism, I dried leaves using a plant press coupled to a hairdryer, as was also used by Renske Onstein (2010) to make herbarium specimens, see figure 3.2. This method was chosen after performing a preliminary protocol optimization using a

specimen of *Rinorea yaundensis* I took from Burger's Bush prior to the field work. For more information on this, see Appendix.

After leaves were picked in the forest, they were carried to my place of residence in a damp plastic bag so as to limit damage prior to drying. Leaves were pressed in a plant press between sheets of (news)paper and corrugated metal or cardboard ensuring channels for airflow, see figure 3.2. A large plastic bag was pulled over the press, making sure of an air tight seal using duct tape. At the other end of the plastic bag, a hair dryer was attached which provided a constant flow of unheated air through the plant press. No heating was applied by the hair dryer so as not to melt or burn the plastic bag, but also because of the increase in chemical degradation products in the LC-MS spectrum of the oven dried samples, presumably caused by heat. Drying took place over the course of at least two whole days. During the protocol testing using *R. yaundensis* material, this timeframe was determined by comparison of dry weight of the oven dried material to that of the hair dryer/plant press material. Oven drying overnight left roughly 30% of dry weight. This level of evaporation had been reached in the hair dryer/plant press method after two days. Once dry, leaves were stored in plastic zip bags and kept in a closed box. **Recommendation:** although I noticed no adverse effects of this storage method (e.g. fungal growth), in hindsight it may have been better to store the dry leaves in paper bags and with silica gel to prevent rehydration.

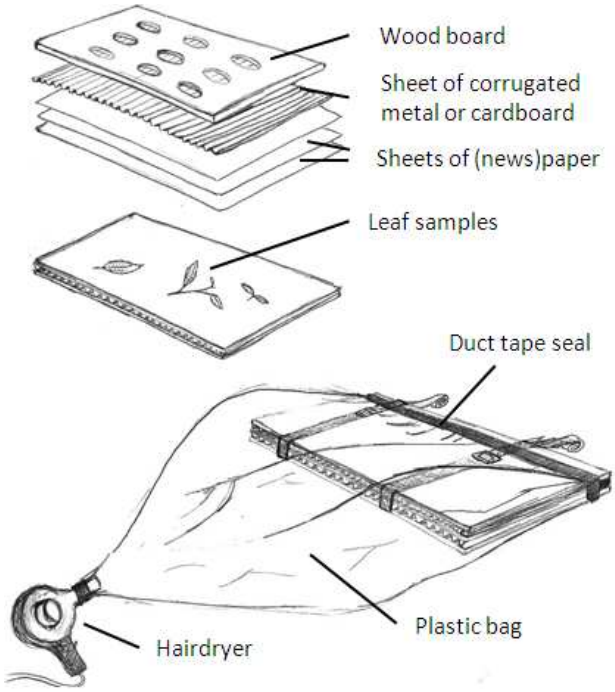


Figure 3.2 Schematic of the plant press/hairdryer set-up.

METABOLITE EXTRACTION

After dried leaves had been taken back to the laboratory in Wageningen, samples selected for LC-MS analysis were ground in a pestle and mortar, using liquid nitrogen to make the leaves more brittle so grinding was easier. Powdered leaves were kept frozen and stored in eppendorff tubes at -80°C until further processing, as humidity from the air will deposit on ground material due to low temperature. The following day, lyophilisation (freeze-drying) was performed overnight to remove all water from the samples. Lyophilisation will allow for direct comparison of metabolic profiles, as any residual water content dilutes the dissolved chemicals unevenly across samples. Also, once ground material is thoroughly dry, it can be stored (in the dark) at room temperature. **Recommendation:** pierce lids of eppendorff tubes prior to lyophilisation; do not open the lids completely. Initially, I had performed the latter which somehow caused some of the material to be thrown out of the tubes and scatter through the freeze-drying machine. Possible causes for this are that the material had started to boil due to the vacuum. This was unexpected since the samples had been air dried prior to freezing and had been stored at -80°C prior to placement in the freeze-dryer. Another cause could have been a sudden release of the vacuum, which would be due to human error on behalf of an unknown culprit, but this is speculation. Piercing the lids would have prevented the spilling out of material. Because of this unforeseen setback, several other samples were ground and lyophilized, this time piercing the eppendorff lids, saving these extra samples from a similar fate.

After lyophilisation, leaf material was weighed in a new eppendorff tube used for extraction. Per 2 mg leaf material, 35 µL of solvent was added for extraction. This dilution was advised by PRI LC-MS operator Bert Schipper, after we had performed a test run using a 7 fold higher dilution. At this point, I included four technical repeats or quality controls, from a sample that had yielded sufficient ground leaf material, that are used to assess the technical variation and quality of the subsequent analytical steps. Solvent consisted of 75% Methanol, 25% H_2O and 0.1% Formic Acid. This solvent composition is reported to be suitable for efficient extraction of a wide range of plant compounds (De Vos et al, 2007).

For extraction, I placed the eppendorff tubes containing the mix of sampled leaf material and solvent in a sonication bath for 15 minutes. Between 100 µL and 1 mL of extract was then filtered using an Anotop 10 membrane filter tipped syringe injected into a 1 mL autosampler vial with aluminium crimp cap. **Recommendation:** apply pressure to the filter tipped syringe carefully and gradually, as the filter tip may be pressed off, causing extract to leak which is both a waste of sample and potentially hazardous, as the sample is extracted in MeOH.

The exact amount of extract poured into the vial is not important, as the LC-MS device injects a small, standardized amount each for each LC-MS run. For samples that had only a small amount of dry leaf material and subsequently a low volume of extract, a 700 µL insert was placed in the vial to allow for submersion of the auto-injector. A minimum amount of extract that can be analysed in this way is approximately 100 µL. Vials containing extract were placed into the LC-MS autosampler and placed in the LC-MS device.

LC PDA QTOF MS ANALYSIS

Liquid Chromatography Photo Diode Array Quadrupole Time Of Flight Mass Spectrometry (LC-PDA-QTOF-MS) is an analytical technique that combines the physical separation of chemicals in a sample using liquid chromatography, with ionization and subsequent detection of ions in the mass spectrometer (Moço, 2007; De Vos, 2007). During ionization, metabolites are given either a positive or a negative charge, which is measured in positive and negative detection mode respectively. In this type of LC-MS, only one of these detection modes is used at once, and I used negative detection mode, as it is used most often at PRI Bioscience (De Vos, 2007). Not every metabolite receives either charge equally readily, which had consequences for the testing of the CPG hypothesis, see chapter 4.

A schematic overview of this technique is given in figure 3.3. Operational settings of the device were as described in De Vos et al (2007). The HPLC Q-TOF MS was guided by MassLynx software, which converted and stored the signals from HPLC, PDA and TOF MS as ".raw" data files. Several software packages were subsequently used in the further processing and analysis of these data, as described in the next chapters.

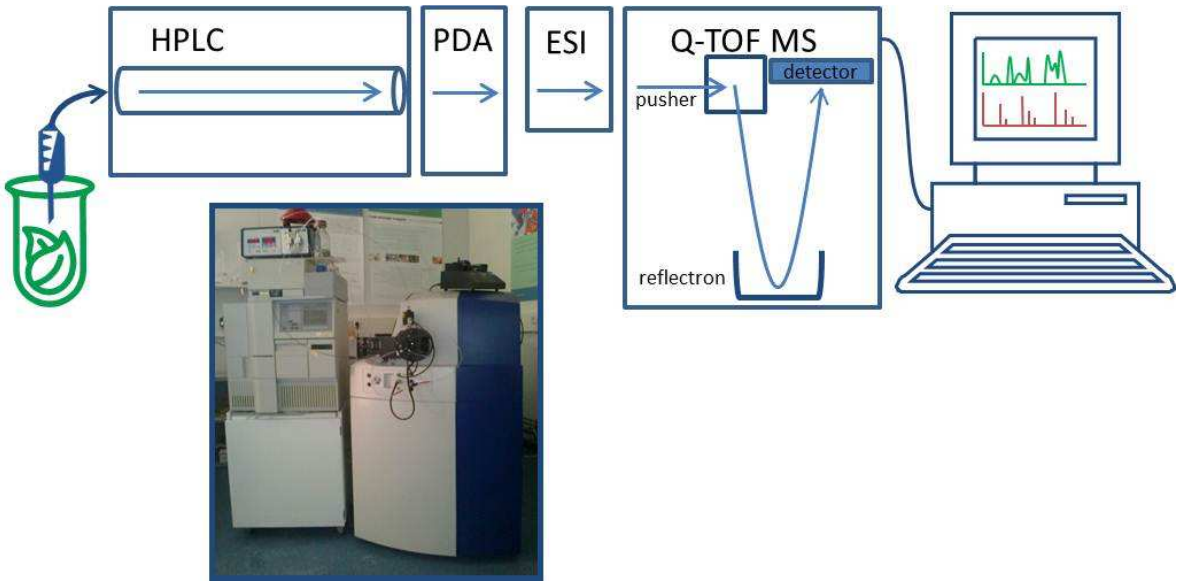


Figure 3.3 Schematic flow-scheme of the processes from plant extract to raw data, that take place inside the LC PDA TOF MS like the one on the photograph, which is the device in the laboratory of PRI Bioscience on which my analyses were performed. The LC-MS vials containing plant extracts, symbolized by the leaves in a test tube, were auto injected by the device into the High Performance Liquid Chromatography (HPLC) column, which separates metabolites based on polar/apolar interactions between the solid phase and liquid (acetonitryl) phase. This causes metabolites to reach the MS apparatus at a different Retention Time (RT). Next, a Photo Diode Array (PDA) measures the UV absorbance spectrum of the eluting substances (UV-vis). Subsequently the chemicals in the extract are given a charge by Electrospray Ionization (ESI). Ions are accelerated in the 'pusher' by means of an electric field, after which ions have reached a certain velocity. The magnitude of the velocity depends on the mass and the charge of the ion. The time it takes an ion to reach the detector, the 'Time Of Fligt' (TOF), will vary according to velocity. The charge given during ESI is constant, thus making the TOF proportional to the mass of the ion. TOF is lengthened for higher resolution by means of a "reflectron", which is another static electric field that reverses the direction of the ion, functioning in effect as a mirror. Also, the reflectron compensates for slight differences in kinetic energy and focuses ions of the same mass to reach the detector at the same time. All ions are detected and TOF is used to determine the accurate mass of the ions, which can now be called "mass signals". For every extract, a metabolic spectrum is thus generated as ions eluting generate a mass spectrum that varies along the retention time (Moço, 2007; De Vos, 2007).

DATA PROCESSING

Raw metabolic profiles generated by MassLynx data management software 4.0 (Waters, Manchester), can be viewed and visually compared using the same software package. An example of such a chromatogram is given in figures 3.5A and B. This raw data cannot be used for statistical analyses, and first needs to be converted into a dataset that allows for comparisons between metabolic profiles of different extracts. The first steps in this data processing were taken using the software package MetAlign (A. Lommen, see De Vos et al, 2007 for details) which performs the following processing steps, as quoted from De Vos et al (2007): "(i) mass data smoothing using a digital filter related to average peak width; (ii) local noise calculation as a function of retention time and ion trace; (iii) baseline correction of all ion traces and introduction of a threshold to obtain noise reduction; (iv) scaling and calculation and storage of peak maximum amplitudes; (v) between-chromatogram alignment using high signal-to-noise peaks common to all chromatograms; (vi) iterative fine alignment by including an increasing number of low signal peaks; (vii) output of aligned data into a csv-file compatible with

Microsoft Excel and most multivariate programs". This output file provides a list of mass peaks with calculated accurate mass and retention time for each of these peaks, as well as the measured intensity (TIC) of these peaks for each extract's LC-MS run (see figure 3.5). To reduce the quantity of redundant mass peaks, these were filtered using MetAlign Output Transformer (METOT) (H. van der Geest, PRI), which removed all peaks with intensity consistently lower than four times noise, and those that were present in less than 3 extracts. This program also performs a quality assessment using the quality control extracts.

Chemicals will have isotopes due to for example the natural presence of C-13 carbon and other atomic isotopes. These isotopes represent the same metabolite, but register as a distinct mass peak. An example of this can be seen in figure 3.5D, which shows a distinct "step" pattern in the mass spectrum. Also, metabolites can fracture during ionisation and some atoms or groups are lost whilst still eluting at the same retention time. These atoms and groups have known masses, for example oxygen which has a mass of approximately 16 Da. To compact the dataset, these mass peaks belonging to the same metabolite were grouped together, using a software package called MsClust (Y. Tikunov, PRI) This software combines chromatographic distance and pattern similarity distance and uses Multivariate Mass Spectra Reconstruction (MMSR) to cluster mass signals around centres of density. Each of these clusters should essentially represent a metabolite. The program assigns a dimensionless membership value called the cent-factor (between 0 and 1) to each mass signal, depending on the distance to this centre of density, with the closest signal receiving the highest membership value. A metabolite cluster is represented in the resulting dataset by a centrotype, the intensity of which is calculated by adding the intensities of all signals, weighed by multiplying each by its cent-factor, which further reduces noise.

Data analysis and peak selection

Two main strategies of data analysis were used. Firstly, differential analyses between classes of samples (old versus young, and accepted versus rejected) by means of t-tests in both SPSS using the MSClust output and MetAlign, which provides such a function on raw data files while these undergo the processing steps i-vii (Vorst et al 2005). Secondly, multivariate statistics were applied to the datasets using GeneMaths and Simca-P.

SPSS t-tests: after importing the MSClust output data into SPSS, each metabolic profile I wanted to analyze in this manner was assigned a class, e.g. accepted or rejected, and Student's t-tests were performed on every metabolite. For each individual metabolite in these metabolic profiles, mean intensity and standard deviation were determined and compared between classes. Per metabolite, it then becomes clear whether it differs significantly in intensity between the assigned classes.

MetAlign t-tests: has the option of performing differential analysis on the raw data file (Vorst et al*.,* 2005). For this the metabolic profiles were divided into two classes that were to be compared, which were aligned and baseline corrected (MetAlign steps i-vii) and analyzed in two ways. MetAlign performed t-tests on each aligned mass peak to asses which are either 1: present at two fold amount in one group over the other ($p<0.01$); and 2: are present significantly ($p<0.01$) in one class but are absent (do not elute above the background noise threshold) in the other class. Depending on the class division (e.g. accepted vs. rejected), this yielded metabolites of interest. MetAlign produces differential chromatograms which can be used to easily visualize these mass peaks. For more details, see Vorst et al (2005).

GeneMaths clustering and PCA: MSClust output was loaded into GeneMaths (Applied-Maths) (Vorst et al, 2005; De Vos, 2007), which was used to cluster metabolic profiles to give an overall indication of similarity between metabolic profiles, or samples in effect. The GeneMaths software provides several options with regard to clustering methods, with the Pearson correlation distance measure clustering with a UPGMA tree being mostly used with this sort of analysis (Vorst et al, 2005). I also used a Jaccard similarity clustering, which ignores shared absences and only takes shared presences

into account, with Neighbor joining tree for comparison. These clusterings of samples were only performed visualize the similarity between samples and to assess for example whether samples would cluster according to ontogenetic stage (young or old leaf) or if they belonged to the same species. Using GeneMaths, a Principal Component Analysis (PCA) was performed on the dataset, one including. Metabolites were selected that were the most correlated with young leaves, and were located centrally within the variation between the young leaf runs.

Simca-P OPLS-DA: Using a different package, Simca-P (Umetrics) an Orthogonal Partial Least-Squares Discriminate Analysis (OPLS-DA) was performed (Wiklund et al., 2008) to distinguish between classes of samples using multivariate statistics, and assess which metabolites have the most impact on the principal components. This method is used to select metabolites using multivariate statistics to distinguish between groups of samples that cluster together in a PCA, but may still have chemical differences that correlate to a biological difference.

Using methods listed above, some centrotypes or mass peaks were selected that correlate to butterfly behaviour, and thereby possibly have a biological function for the butterfly. Some of these metabolites were subsequently subjected to attempts at structure elucidation.

METABOLITE IDENTIFICATION

Usually, in metabolomics, peak identification is the most time consuming step (Moço, 2007). Using untargeted LC-MS, one obtains several characteristics of the metabolites/mass peaks detected, namely HPLC retention time, UV absorbance spectrum (if PDA coupled MS is used) and accurate mass (Moço, 2007). Of these, accurate mass is used to provide the molecular formula of a chemical, so the amount and character of different elements in the chemical. In the process of peak identification from LC-MS I was instructed by Ric de Vos. The first step is to turn to the MSClust output dataset and pinpoint the "parent ion". I then compared the accurate mass of the parent ion to a list of metabolites of Masanori Arita (www.metabolome.jp), which Ric de Vos had at his disposal. A measure of similarity between the masses from this list and accurate measured mass is calculated by the absolute value of (measured mass – calculated mass) / calculated mass * 10^6. As a rule of thumb, if this measure falls within 5 ppm, these masses are likely to be composed of the same atoms. This method is dependent on the metabolites already reported and present in a database, which is not necessarily the case for metabolites extracted from a plant that has not been subjected to metabolomics before, such as *R. ilicifolia*, or any *Rinorea* for that matter. LC-MS is therefore of limited power in structure elucidation, however it can direct targeted analysis.

An obvious subsequent analysis would be tandem mass spectrometry (LC-MS/MS), during which a metabolite is selected for based on RT and mass so only the metabolite of interest is ionized. After ionization, this target metabolite is further fractionated, providing a characteristic fractionation pattern (Moço, 2005). This pattern would yield both structural information based on the masses of the fragments, which could be distinct molecular groups, and also in comparison to molecular libraries. Again, the power of this latter structure elucidation method is limited by the metabolites already present in the library. Another method towards structure elucidation could be H NMR, such as performed in this study on a CPG sample (see chapter 4). The major drawback of this method is that it requires a purified chemical in a relatively large amount in order to provide a clear H NMR spectrum.

3.2 RESULTS

SAMPLING OF RINOREA ILICIFOLIA PLANT MATERIAL

A list of the samples taken is provided in table 3.1. It must be noted that, although there are fewer samples from rejected leaves, actually many more rejections were observed than acceptances. The causes of this are mainly the difficulties involving confirmation of rejection and the following of a butterfly as mentioned in observations.

I expected to find *R. ilicifolia* specimens that were invariably rejected, but I soon realized that mature leaves were usually rejected and young, apical leaves were usually accepted, even if these belonged to the same *R. ilicifolia* specimen. An invariably rejected specimen is therefore presumably simply a specimen without young leaves. Without an apparent basis for intraspecific variation, and with rejected leaves being a rule rather than an exception (since there are many more old/mature leaves per individual *R. ilicifolia* specimen than ontogenetically young ones), I decided to focus on gathering the accepted leaves that will certainly contain the chemical by which *C. egesta* recognizes its host.

I now expected the difference between accepted and rejected plants to be mostly due to a difference in leaf development and show this by also comparing young and old leaves within individuals and trying to minimize between individuals variation. This means that any correlation found between a metabolite and biological activity, is likely to be caused by leaf ontogeny, and such a metabolite probably does not possess any biological significance in host recognition by *C. egesta*. Having stated this, I will continue down the path of this untargeted metabolomics approach as shown in figure 3, in order to provide an example and "proof of principle" of this method.

I had not clearly observed and sampled cases where a young leaf had been accepted but one or more older leaves of the same individual plant had been rejected. I therefore sampled cage experiment specimens as well, counting the young oviposited leaves as accepted and unoviposited old leaves as rejected. I did include three young rejected leaves in the analysis, although I was not entirely certain of the observation. For leaf 10 (young rejected), alighting was clearly observed but not "drumming". For young leaf 18, I did not personally observe the rejection, but my guide Andrews did.

I also took young and old leaves from a "random" *R. ilicifolia* that had many young leaves as well as older leaves, but on which I had observed neither acceptance nor rejection. Of these, sample Old_5 was used for technical repeats and extracted four times to assess my technical variation.

Table 3.1 list of samples gathered. Samples marked with an X under LC-MS were analysed. The number under Sample/Leaf indicates a *Rinorea ilicifolia* individual, a letter indicates an individual leaf.

LC-MS	Sample/ Leaf	Activity	Leaf age	Date gathered MM/DD	Remarks
	1	rejected	Old	12/16	Rejection observed in November
	2	accepted	Young	11/11	
	3	accepted	Young	11/11	
	4	Rejected	Old	11/11	
	5	Accepted	Young	11/24	
	6	Accepted	Young	11/24	
X	**7A**	Accepted	Young	11/25	
X	**7B**	Unknown	Old	11/25	Old leaves on the same branch as 7A, to compare within individual
X	**7C**	Unknown	Old	11/25	
	8	Rejected	Old	11/25	
	9	Accepted	Young	11/30	
X	**10**	Rejected	Young	11/30	Unsure of "drumming" observation
	11	Accepted	Young	12/06	
X	**12**	Accepted	Young	12/06	
X	**13A**	Accepted	Young	12/10	From the 12/06 cage experiment, see table 4.1, taken the two young, oviposited leaves as accepted (13A & B) and two older unoviposited leaves as rejected (13C & D)
X	**13B**	Accepted	Young	12/10	
X	**13C**	Rejected	Old	12/10	
X	**13D**	Rejected	Old	12/10	
X	**14**	Accepted	Young	12/12	
	15	Accepted	Young	12/12	
	16	Accepted	Young	12/12	
X	**17A**	Rejected	Old	12/12	See tables 2.1 & 4.1: observed rejection of 17A by ♀10. From subsequent cage experiment with ♀10 and same *R. ilicifolia* came young oviposited leaf 17B as accepted, and unoviposited leaves 17C-E as rejected
X	**17B**	Accepted	Young	12/15	
X	**17C**	Rejected	Old	12/15	
X	**17D**	Rejected	Old	12/15	
	17E	Rejected	Old	12/15	
X	**18**	Rejected	Young	12/12	Andrews observed rejection
	19	Rejected	Old	12/12	
X	**20**	Accepted	Young	12/13	
X	**21**	Accepted	Young	12/13	
X	**22**	Accepted	Young	12/13	
X	**23**	Accepted	Young	12/15	
X	**24**	Rejected	Old	12/15	
X	**25**	Rejected	Young	12/15	
	26	Accepted	Young	12/16	
X	**27**	Accepted	Young	12/16	
	28	Accepted	Old	12/16	
	29	Rejected	Old	12/16	
X	Old 4	Unknown	Old	12/16	These samples belong to the same *R. ilicifolia* individual that had many young and old leaves. Sample "Old 5" was extracted four times as QC
X	Old 5	Unknown	Old	12/16	
X	Young 4	Unknown	Young	12/16	
X	Young 5	Unknown	Young	12/16	

LC PDA TOF MS ANALYSIS AND DATA PROCESSING

Examples of MassLynx chromatograms are provided in figure 3.5. Given here are the Retention time on the horizontal axis and on the vertical axis Base Peak Intensity (BPI) of two extracts, 10A and 17A respectively. Visually comparing the two chromatograms shows many differences, which are to be expected since 10A is an extract of a young leaf, as opposed to 17A being an older leaf. BPI shows only the count of the ion eluting in greatest abundance at a particular RT, which provides a "smoother" looking chromatogram than when the Total Ion Count or TIC was viewed, which shows the cumulative ion count of all mass peaks as a function of RT. A view of the mass spectrum at RT = 21.79 minutes (at the vertical line) is given in figure 3.5C, showing the many different mass peaks at this particular RT.

MetAlign data processing yielded a dataset with 7993 aligned mass peaks. Using METOT this figure was reduced to 3294 mass signals Quality assessments were made by comparing the quality controls, the technical repeats made from sample Old_5. Results

of this can be seen in figure 3.4. Unfortunately, one of the four quality controls (QCs) had, due to an error in LC-MS settings, not been injected into the device, leaving only three technical repeats. This processing step using MSClust provides a dataset with far fewer centrotypes than mass peaks present in the MetAlign output, roughly 400 centrotypes in my dataset, which were used for further data analysis.

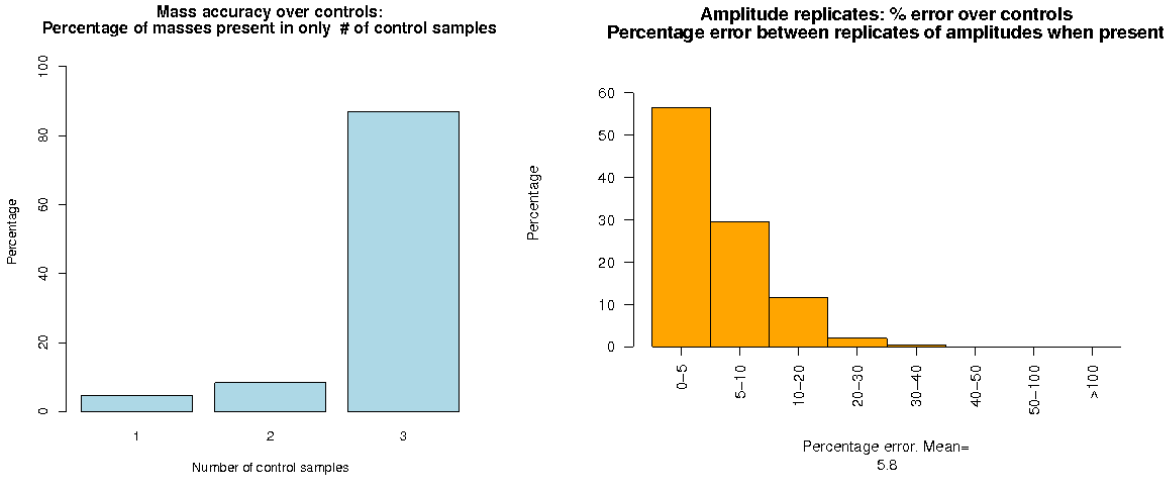


Figure 3.4 Quality assessments produced by METOT. The left bar chart shows the percentage of mass peaks present in 1, 2 or 3 quality controls, indicating about 90% of mass peaks are consistently found in all three QCs. After METOT filtering, any mass peak present in less than three extracts was filtered out, indicating that any mass peak found in only one or two QCs must be present in at least one or two other extracts. The right bar chart shows the variation in amplitude (intensity) of mass peaks between QCs, showing nearly 60% of peaks have only 0-5% variation in amplitude. Average error is 5.8% indicating a consistent technical handling and processing has occurred. Any variation found in presence and intensity of mass peaks in extracts is therefore unlikely to be caused by the extraction and LC-MS procedures, but will rather be due to biological variation or earlier sample handling up to the point of extraction.

DATA ANALYSIS AND PEAK SELECTION

I used different methods of dataset analysis, in order to select several metabolites and/or mass peaks that correlate to leaf age and/or butterfly behaviour: differential analysis using SPSS on the METOT output, differential analysis using MetAlign on the raw LC-MS run and multivariate analyses. I four of these which I subjected to further identification, again mainly as proof of principle and to illustrate how this process is performed, with the inclusion of some caveats one must bear in mind when using this method of metabolite identification. Since I have little indication that any metabolites/mass signals found in these analyses, including the individual ones discussed here, correlate more to butterfly behaviour than to leaf age, I do not recommend these particular ones to be subjected to further methods of structure elucidation, i.e. LC-MS/MS and/or H NMR.

SPSS t-tests: for the SPSS analysis, samples with unknown bioactivity such as Old_4 and _5 were omitted from the t-test. Independent Samples t-tests were performed for every metabolite by comparing a profiles of accepted (N=11) with rejected (N=9) leaves. The results of this is given in table 3.2, listing only those metabolites that are more abundant in accepted than rejected leaves with $p<0.01$. It should be noted that there were a greater number of metabolites more abundant in rejected than accepted leaves, presumably because the accepted (young) leaves were far more variable in chemical make-up than the rejected (old) leaves. From these results, metabolites 1150 and 550 were selected to serve as examples of the process of metabolite identification on the basis of accurate mass as determined by LC-MS, as these will show to be illustrative for this process (see next step: metabolite identification).

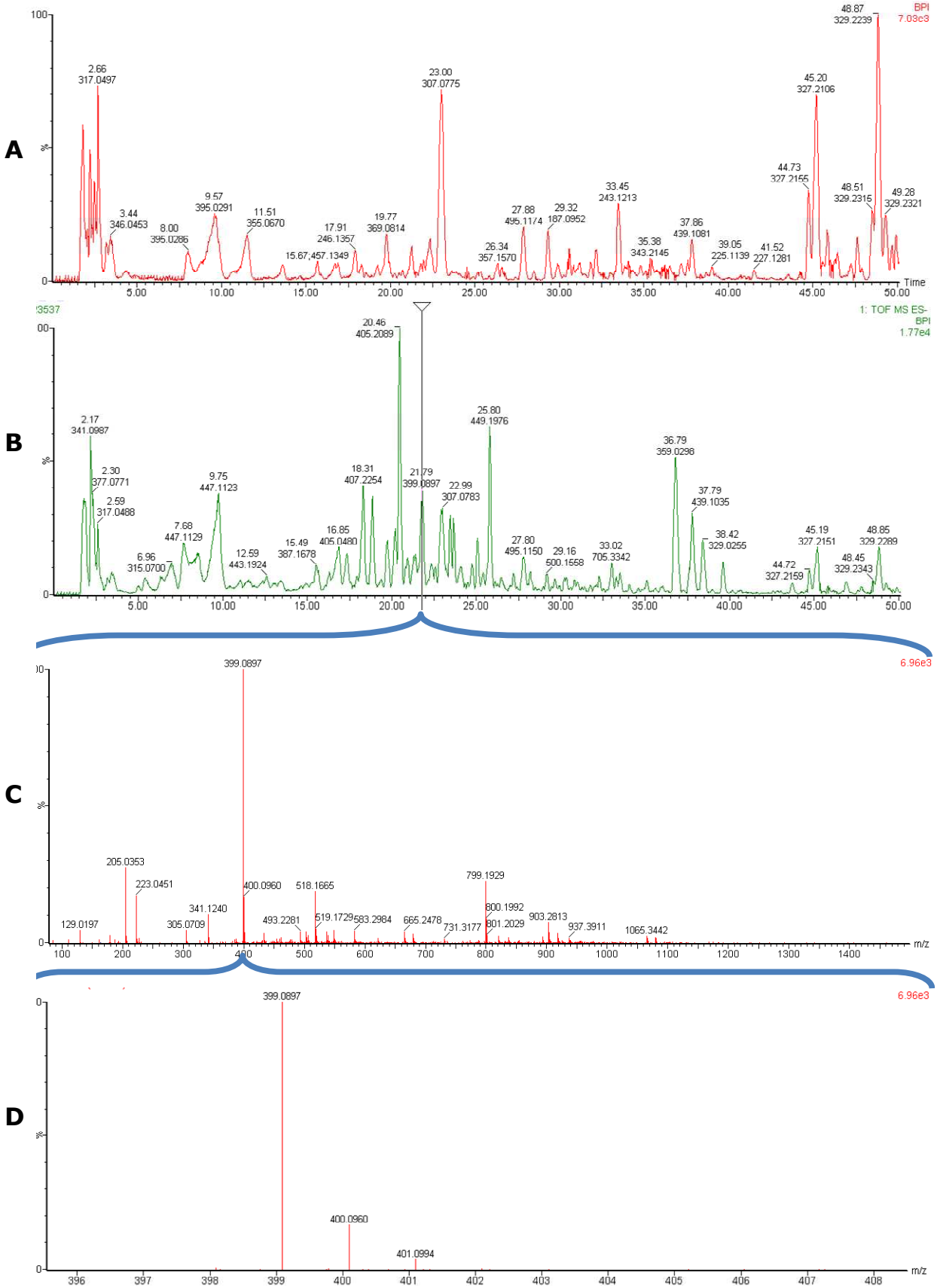


Figure 3.5 Examples of MassLynx output of raw data. **A** shows the HPLC run of extract 10A, an accepted young leaf for an RT range of 0 to 50 minutes. **B** shows the chromatogram of extract 17A, an older, rejected leaf in the same range. Vertical axes show the range of intensity from 0 to 100% of the highest intensity of each chromatogram/mass spectrum, which is given in the top right corners. For some chromatogram peaks, the program provides the RT and mass of the most abundant chemical eluting at that RT. At the vertical line at RT = 21.79 minutes, **C** gives the mass spectrum of all mass signals eluting at that RT. **D** zooms in around m/z = 400D, which shows the characteristic "step" pattern of the ion with m/z = 399.0897 D with its isotopes m/z = 400.0960D and 401.0994D respectively.

Table 3.2 List of metabolites present in greater abundance in extracts of accepted leaves than in extracts of rejected leaves (rejected group including three rejected young leaf extracts) as determined by SPSS Independent Samples t-tests at 2-tailed sig. p<0.01

	Metabolite Descriptives (MSClust)			SPSS Independent Samples t-test			
Meta-bolite	Cent-factor	Ret(µmin)	Mass (D)	activity	Mean intensity	Std. Dev.	Sig. (2-tailed)
98	0.70	2065850	215.01442	accept reject	413.08 226.54	131.34 62.58	,001
113	1	2155133	114.01855	accept reject	942.34 497.83	325.40 136.67	,001
474	0.39	8996083	749.09967	accept reject	80.72 56.22	22.92 11.18	,007
479	1	9069417	711.14093	accept reject	663.80 439.31	189.01 102.43	,004
486	0.78	9176583	147.02974	accept reject	155.61 102.31	42.97 23.86	,003
492	0.94	9646617	395.02911	accept reject	4228.62 2252.08	1537.61 535.08	,002
505	1	9809250	191.01930	accept reject	8251.18 4366.61	2434.58 1463.07	,000
512	0.25	9844983	749.09753	accept reject	112.62 57.95	34.05 21.22	,000
550	1	11505150	355.06690	accept reject	17171.39 10409.50	4236.84 2460.57	,000
660	1	13617517	355.06766	accept reject	3385.53 2001.17	974.88 420.13	,001
1144	1	19736467	129.01968	accept reject	917.90 505.05	318.13 163.57	,002
1150	1	19736467	405.05008	accept reject	581.63 276.79	256.33 53.27	,003
2057	1	26287867	401.14667	accept reject	538.92 231.23	244.23 126.38	,002
2806	1	33471931	244.12769	accept reject	1328.12 684.92	407.53 353.27	,001
3259	1	47893665	265.08554	accept reject	470.45 248.07	176.40 37.96	,002

MetAlign t-tests: Using the raw LC-MS data, I performed differential analyses in MetAlign, the result of which is visualized in figure 3.6. T-tests were performed by the software on individual mass peaks (not centrotypes as such as provided by the MSClust output used for SPSS t-tests) after alignment by MetAlign, to see which peaks are 1: present in over two fold higher abundance ($p<0.01$) in class 1 over class 2, and 2: are present significantly ($p<0.01$) in class 1 and absent in class 2; class 1 being young, accepted leaves and class 2 being old, rejected leaves. Metabolic profiles of young rejected leaves were omitted in this analysis, because including these in class 2 no longer yielded any peaks of interest. For further mass peak identification, I selected the peak with m/z≈749.1 D, eluting at RT≈11.5 min, which has the highest signal in analysis 2. This signal is equally high in analysis 1, but appears to be lower because some other peaks not occurring in analysis 2 have a relatively higher TIC.

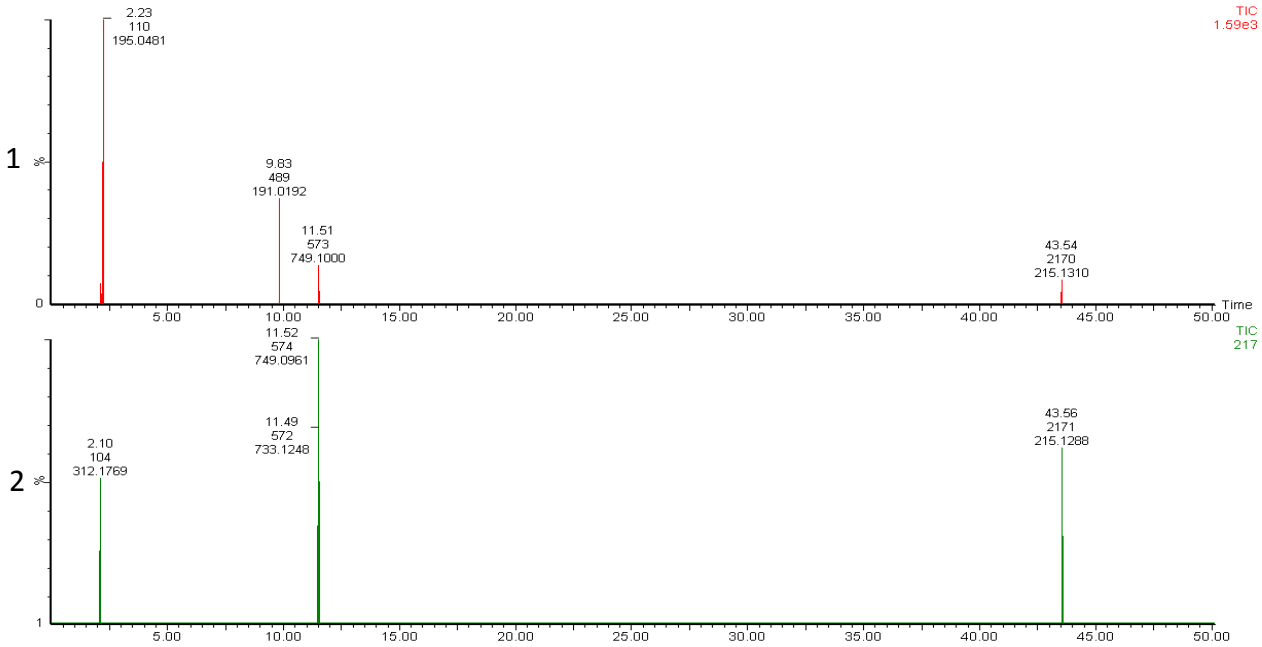


Figure 3.6 Output of the differential analyses of MetAlign output viewed in MassLynx. **1** shows which mass peaks have a twofold higher intensity in young accepted leaves than in old rejected leaves. **2** shows which mass peaks are present in significant amounts in young accepted leaves, but do not elute above the noise threshold in old rejected leaves (for both $p<0.01$). Peaks in **2** are also present in 1, but appear lower because **1** also contains peaks with much higher intensities. Total Ion Count (TIC) is given in top right corners. For each peak the RT, scan number (corresponds with RT) and mass are given (top to bottom).

GeneMaths clustering and PCA: GeneMaths clustering can be viewed in figure 3.7. The lower figure, with UPGMA tree with Pearson correlation clustering shows a consistent split between young and old leaves, not only indicating that between individuals there is a consistent chemical difference between young and old leaves, but also within species. For example, old rejected leaves 13C and 13D group inside the "old leaf" cluster, whereas young, accepted leaves 13A and 13B group together in the "young leaf" cluster, although these belong to the same individual *R. ilicifolia* individual. Similar examples are provided by individuals 7, 17 and the specimen from which Young_4 & Young _5 and Old_4 & Old_5 were taken (see figure 3.7). The Jaccard distance Neighbor Joining tree shows an overall similar picture, although the old leaf samples form a distinct clade whereas the young leaf samples do not, presumably because young leaves are generally more variable in metabolite content.

Also using GeneMaths, I performed principal component analyses on the metabolic profiles of the leaf extracts, both including samples of unknown bioactivity, e.g. the quality control individual (see figure 3.8) and only including rejected or accepted leaves (see figure 3.9).

Figure 3.10 shows some metabolites that were selected, and these included 550 and 1150 as were selected in the t-tests and from this analysis no attempt was made to identify other metabolites.

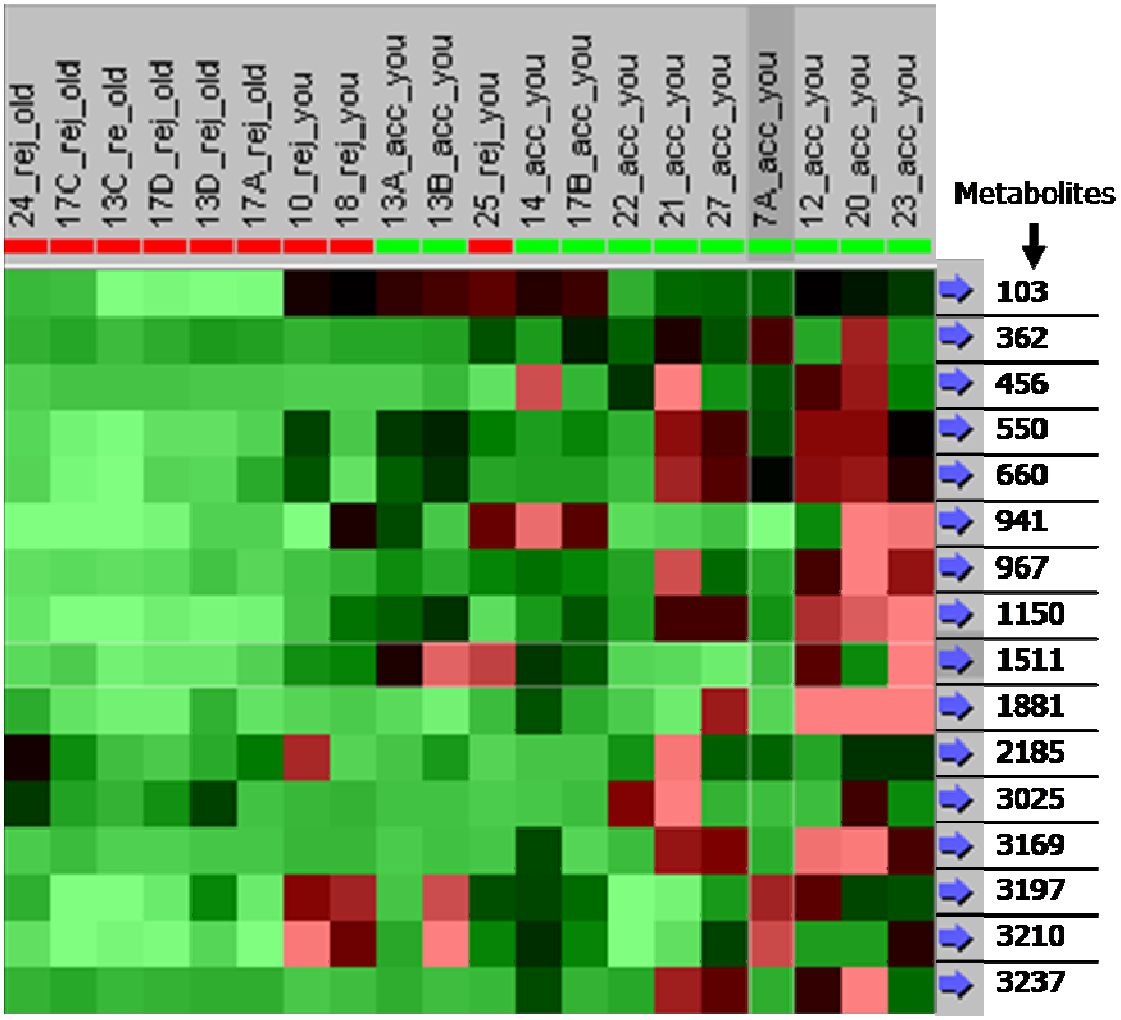


Figure 3.10 Metabolites selected from the PCA of figure 3.9, left. Colors represent relative abundance (red: high, green: low).

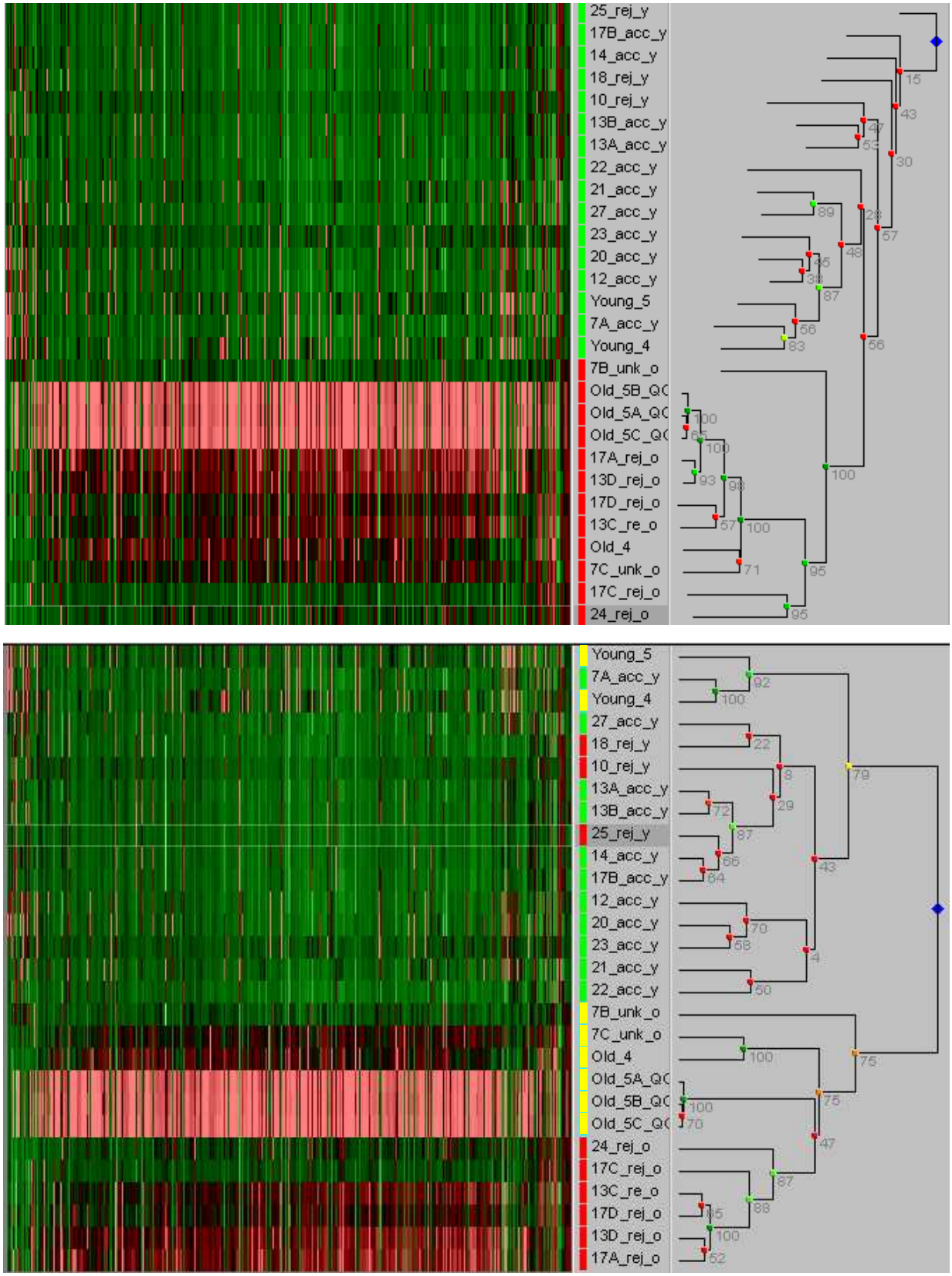


Figure 3.7 Cluster analyses of all runs. Top is a a Neighbour joining tree using a Jaccard distance measure, colours in front of sample names represent leaf age with green: young, red: old. Bottom is a Pearson correlation clustering with UPGMA tree using the same dataset. Colours in front of samples represent bioactivity with green: accepted, red: rejected, yellow: unknown. Numbers at nodes are bootstrap values in both clusterings. The alignment to the left of both clusterings shows all aligned metabolites arranged by retention time, and coloured according to relative abundance with red: high and green: low.

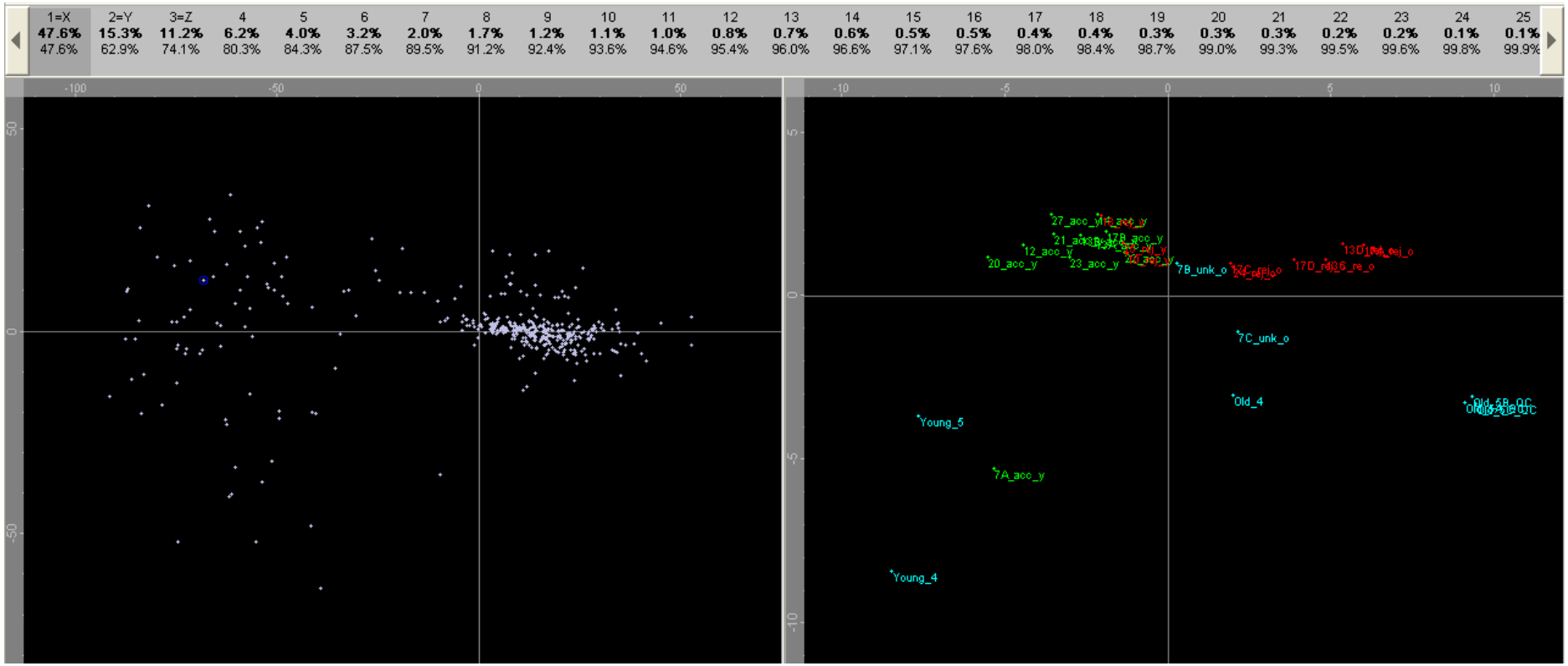


Figure 3.8 Principle component analysis of all extract LC-MS runs. Left shows a PCA of individual metabolites, right the corresponding PCA of extracts. Colouring of extract names, green and red are the same as in Figure 3.7, cyan corresponds to yellow. Axes are the same component in both the left and right PCA, with 47.6% of variation represented by the horizontal axis and 15.3% by the vertical. The horizontal axis appears to correlate with leaf age, but the vertical axis seems to be caused by the deviant Young_4/5 and Old_4/5ABC samples and the samples of *R. ilicifolia* individual 7.

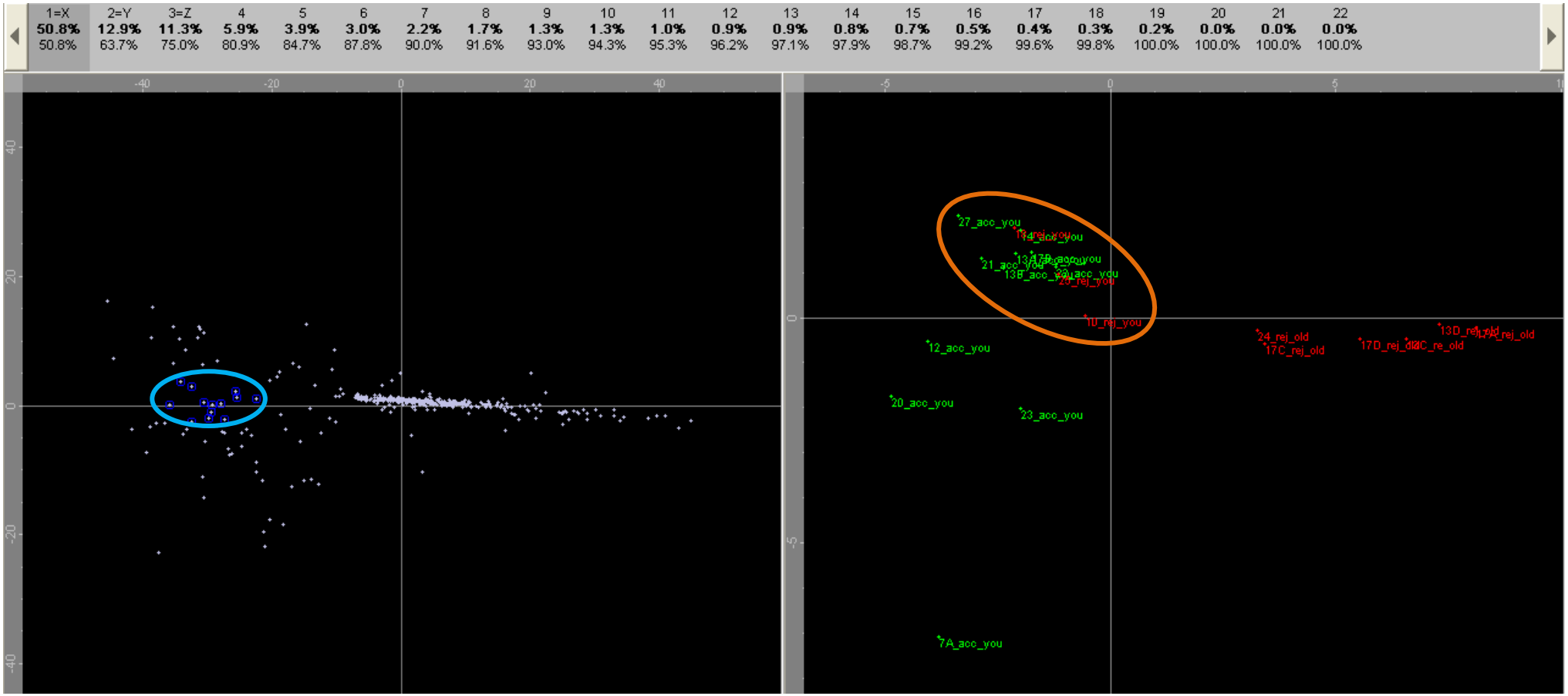


Figure 3.9 Principal component analysis of the metabolic profiles, as in figure 3.8, now of a selection of extracts, having excluded the ones of unknown bioactivity (the cyan samples in figure 3.8) but using the same alignment and MSClust output. The extracts from the old, rejected samples (on the right) mainly show variation along the horizontal component axis (50.8% of variation), whereas the samples from young leaves are also show variation in the vertical direction (12.9% of variation). Encircled in blue are some metabolites that are generally more abundant in the young leaves than in the older leaves but do not add much to the variation within the young leaves. Encircled in orange are a cluster of young accepted and young rejected leaves that were focused on using Simca-P, see figure 3.11 and 3.12.

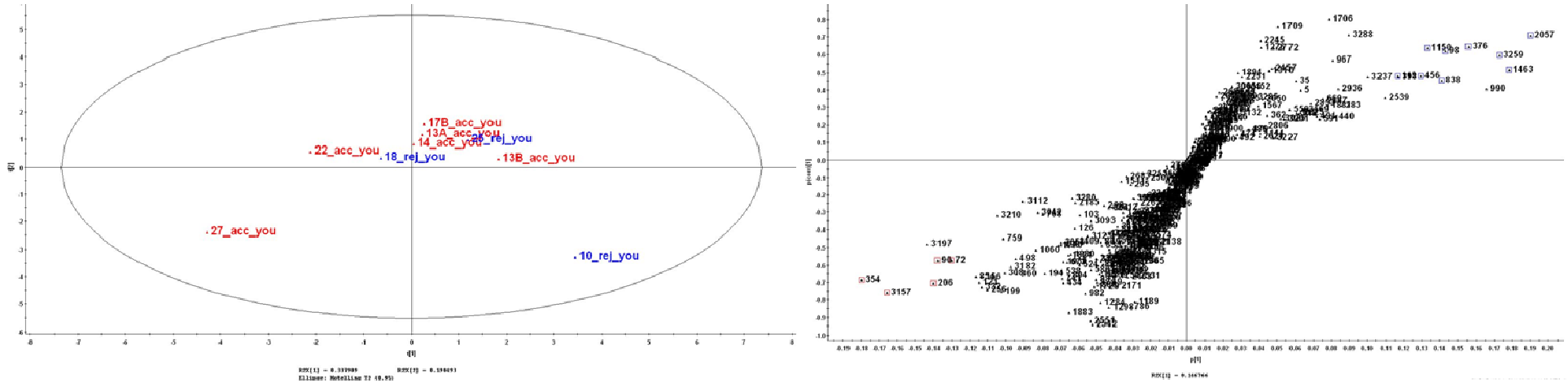


Figure 3.11 Left: PCA generated by SIMCA-P using only the accepted young and rejected young samples that clustered together in the GeneMaths output (see figure 3.9, right). **Right:** S-plot generated by SIMCA-P from an OPLS-DA, of the spread of metabolites. Outliers on the right hand side marked in blue are present in greater abundance in young accepted leaves, those on the left marked in red in the young rejected.

Simca-P OPLS-DA: Several metabolic profiles of young accepted leaves clustered together with young rejected ones (see figure 3.9), which is unsurprising because all are young leaves. From a comparison between young leaves may however circumvent the factor of leaf age and allow us to find metabolites that may be hypothetically responsible for the observed difference in bioactivity. With the help of Roland Mumm of PRI Bioscience, I therefore performed an OPLS-DA on the metabolic profiles in this cluster using SIMCA-P to assess whether there any metabolites could be found that would explain a difference between these two groups, even though these cluster together. Results of this analysis are given in with a PCA of this cluster and an S-plot in figure 3.11, and metabolites of interest in figure 3.12. Only those metabolites with a VIP greater than the confidence interval are relevant, however values of these are so low, they are hardly significant. This means that among young samples, no metabolite convincingly correlated with differential butterfly behavior. Still, metabolite 2057 was chosen from this analysis, again as "proof of principle".

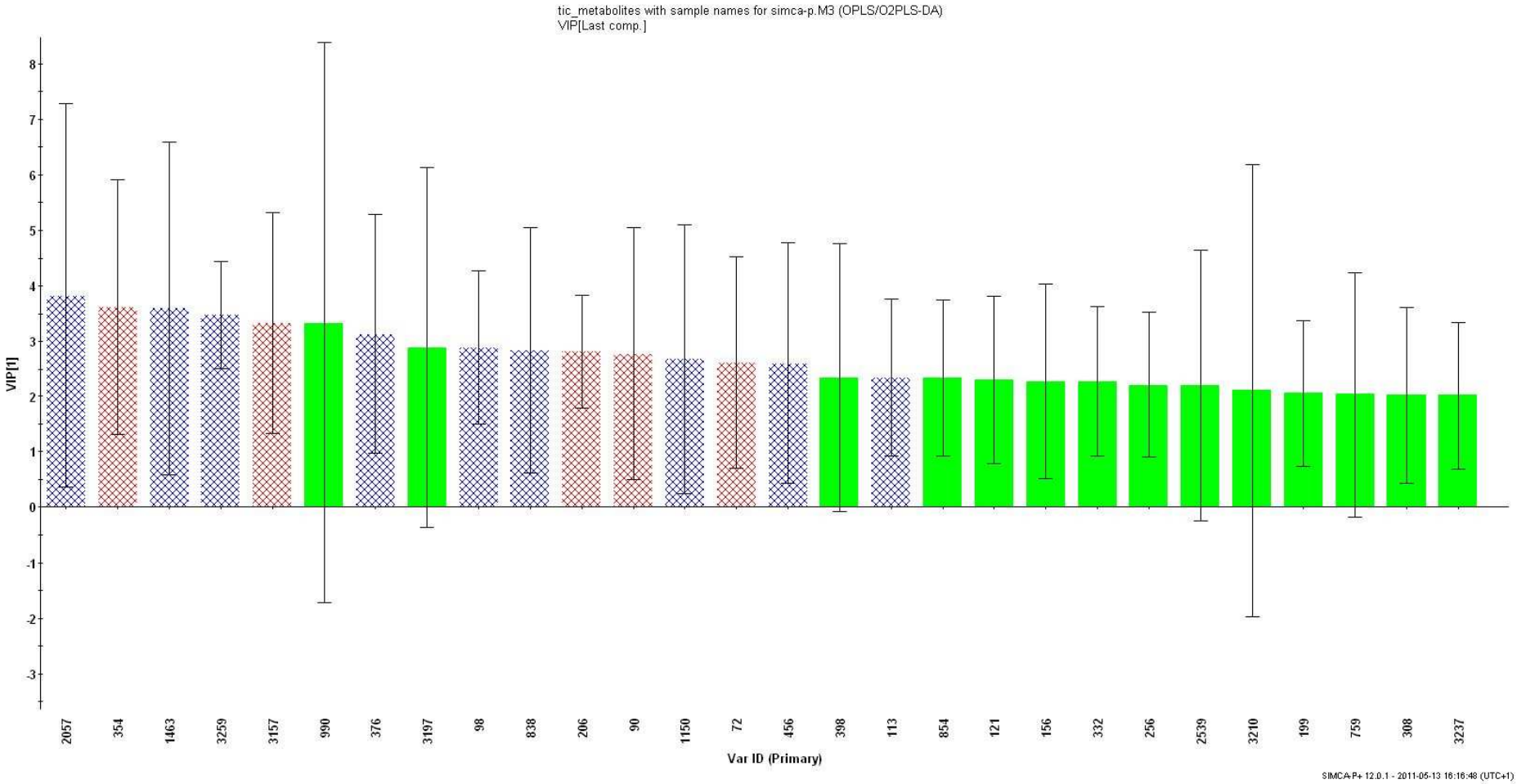


Figure 3.12 Plot of Variable Importance in Projection, plotting the relation between the correlations versus the covariance. The brackets are confidence intervals, and give an indication of the significance of the metabolites selected in the S-plot of figure 3.11. Blue and red correspond to colors in the S-plot, and have a VIP value greater than the confidence interval. Metabolite 2057 was selected for further identification.

METABOLITE IDENTIFICATION

Using the different methods of data analysis, I selected several metabolites and/or mass peaks which I attempted to identify: metabolite/centrotypes 550, 1150 and 2057, as well as the mass peak of m/z ≈ 749.10 D eluting at RT = 11.5 min.

Table 3.3 Descriptives of metabolite 1150. Listed are membership determined by MSClust, RT from the LC-MS, detected mass of each mass peak as determined by MetAlign and the maximum intensity of these peaks measured.

metabolite	membership	RT (min)	detected mass (D)	max intensity
1150	0.927671	19.7365	405.050079	798
1150	0.827251	19.7186	407.047302	127
1150	0.119184	19.7365	506.186066	188

For metabolite 1150, three mass peaks were clustered together into metabolite or centrotype 1150, see table 3.3. Comparison of membership and maximum intensity shows both are highest for the mass of 405.050079 D, indicating this is the most likely parent ion. The two chemicals in this list with masses closest to the 1150 parent ion are $C_{15}H_{18}O_{11}S$ (m/z=405.0497056D) and $C_{22}H_{14}O_8$ (m/z=405.0615906D), which have a measure of similarity of 0.9 ppm and 28.4 ppm respectively. Further indication that 1150 contains a sulphur atom is provided by looking to the raw LC-MS output, see figure 3.14. The list of metabolites provides one putative identity for $C_{15}H_{18}O_{11}S$: 1-O-p-Coumaroylglucose; ?-D-form, (2 or 4 or 6)-O-Sulfate. This is merely the only recorded chemical with this molecular formula recorded (in the library I used) and seeing as many

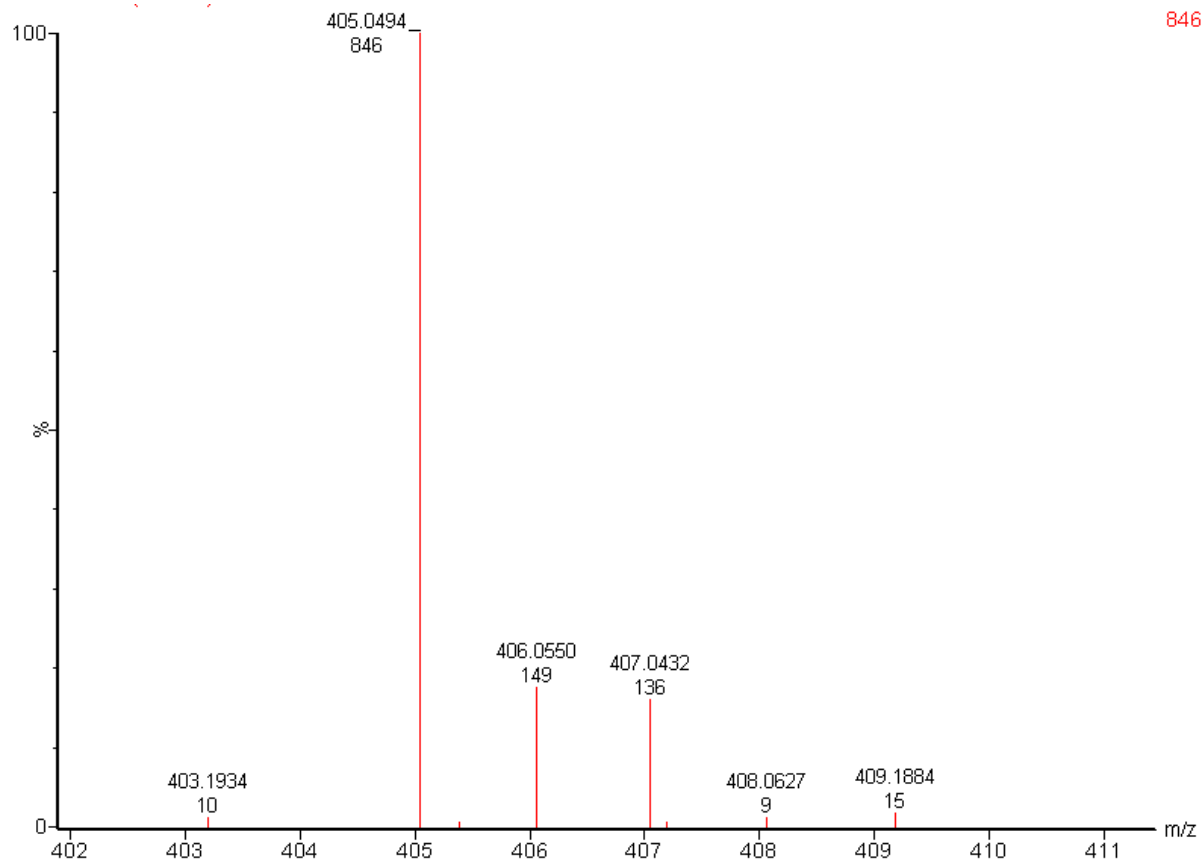


Figure 3.14 Mass spectrum at RT=19.754 minutes, showing the parent ion of metabolite 1150 at 405.0494 D and its isotopes at 406.0550 D and 407.0432 D. The pattern is further indication that this metabolite contains a sulphur atom. Sulphur isotopes naturally occur as ^{32}S: 95.02%, ^{33}S: 0.75%, ^{34}S:4.21% (http://en.wikipedia.org/wiki/Isotopes_of_sulfur) causing the 407.0432D isotope to elute higher than is typical for the distinctive "stairs" pattern.

different structures can hypothetically be formed with this collection of atoms, the exact structure of this metabolite is uncertain until further attempts at structure elucidation (e.g. LC-MS/MS or H NMR) have been performed.

The centrotype/metabolite 550 is a product of 12 clustered mass peaks, see table 3.4, all with high membership values but also very variable masses, which makes it more difficult to identify the parent ion. Heavier masses cannot be explained by atomic isotopes, but may instead be the result of adducts of two lower mass peaks. Also, MSClust groups mass peaks based on their behaviour (retention time and relative abundance) across sample runs, so these peaks may also represent distinct metabolites that have been grouped because they naturally exist in close correlation. The most abundant ion, with accurate mass of 355.0669D corresponds also has the highest membership, and is likely to be the parent ion. According to the library I used (Arita, www.metabolome.jp) the molecular formulae with masses closest to this parent ion are $C_9H_{17}N_4O_9P$ (m/z = 355.066038D, 2.4 ppm) and $C_{15}H_{16}O_{10}$ (m/z=355.06707D, 0.5 ppm). Both have a measure of similarity low enough (<5) to represent the mass peak, meaning there is no basis for distinguishing between these options and to do so would be purely speculative. Further analysis (e.g. LC-MS/MS) would be needed to elucidate the molecular and structural formula of this metabolite.

Table 3.4 Descriptives of metabolite 550.

metabolite	membership	RT (min)	detected mass (D)	max intensity
550	0.85736	11.52	147.03003	271
550	0.857429	11.52	191.01926	1884
550	0.927589	11.51	209.02937	2233
550	0.853449	11.52	210.03369	164
550	0.927671	11.51	355.0669	3904
550	0.927669	11.51	356.07135	675
550	0.92259	11.51	711.14142	365
550	0.925394	11.51	733.12439	431
550	0.919645	11.51	734.12787	147
550	0.924912	11.51	749.09827	554
550	0.920758	11.51	750.10443	187
550	0.92284	11.51	765.0639	183

Table 3.4 also shows that the mass peak I selected from the MetAlign differential analysis, m/z=749.1 D, also appears inside the 550 centrotype with high membership. Comparison to the two molecular formulae in the library, $C_{20}H_{30}O_{30}$ (m/z=749.0749D, 31.2ppm) and $C_{34}H_{38}O_{19}$ (m/z=749.1935D, 127.1ppm), shows that neither of these has a mass similar enough to this mass peak to draw any conclusions about its composition.

Finally, metabolite 2057 was selected from the OPLS-DA analysis, see table 3.5. In this centrotype, the highest membership corresponds with the lowest intensity and vice versa. This makes assumptions as to the identity of the parent ion (if the three mass peaks clustered within 2057 belong to the same metabolite in the first place) rather difficult. Looking at the difference in mass between the peaks, these are approximately 54D and 53D respectively (from top to bottom). These differences can be caused by fragments that have broken from the parent ion during ionization, and these could be

Table 3.5 Descriptives of metabolite 2057.

metabolite	membership	RT (min)	detected mass (D)	max intensity
2057	0.857433	26.30573	357.15555	449
2057	0.927671	26.28787	401.14667	385
2057	0.653794	26.28787	454.17319	633

fragments like C_3H_2O (m/z=54.01056D) and C_3H_3N (m/z=53.02655D). Supposing the peak with the highest intensity is the parent ion, it could have the molecular formula $C_{23}H_{29}NO_{10}$ (m/z=454.171896D, 2.9ppm) which can lose one or both of such fragments (these may be even be linked, the 53D group connected to the 54D group connected to the rest of the molecule). This is merely speculative though, and the low membership of this "parent ion" does not warrant the speculation. Another possible candidate for this metabolite may be $C_{26}H_{31}Cl_2N_3$ (454.1822266D, 19.9ppm). Although its molecular weight is not similar enough, it could yield the fragment ClH_4N (m/z=53.00323D).

DISCUSSION

Several metabolites and mass peaks were found in greater abundance in the accepted leaves than in rejected leaves, and thereby correlated with oviposition preference. However, this correlation was dominated by difference in leaf age.

The clusterings of figure 3.7 show that old leaves are consistently grouped together with Jaccard distance and also with Pearson correlation but with somewhat lower support values. Young leaves on the other hand; seem to be more variable in their chemical make-up. Leaves of a certain age are shown to be more similar to leaves of the same age from different plants, than to leaves of a different age from the same individual. The greater variation in young leaves was also clear in the SPSS t-tests and MetAlign differential analyses, because there were many more metabolites that were significantly occurring in greater abundance in the old and/or rejected samples over young/accepted samples than vice versa (these results are not listed above).

The few young rejected in the analyses, still clustered within the accepted leaves in the cluster analysis and PCA. The OPLS-DA found several metabolites that could cause a differential preference, however with this small sample size and low significance, this is far from certain.

It may well be that physical characteristics (like water content or cuticle thickness) rather than chemical characteristics causes the difference in oviposition preference between young and old leaves. This could mean that with regard to oviposition stimulant, there may be no difference between young and old leaves.

Nevertheless, the validation of the "proof of principle" results I have obtained in this untargeted metabolomics approach, lies in the fact that the methods I have used has, to my knowledge, never been used before in an attempt to discover plant chemicals that elicit oviposition in a specialist herbivore. It appears that my method of sampling, however unorthodox in the field of metabolomics, would still yield quite comparable metabolic profiles.

CHAPTER 4 TESTING OF CPG AS OVIPOSITION STIMULANT

In this chapter I aim to test the hypothesis that CPG is an oviposition stimulant to *Cymothoe egesta* using two distinct experimental approaches; A: chemical analysis (LC-MS) is used to compare the CPG content of accepted and rejected leaves; and B: an attempt to test whether CPG stimulates egg-laying by *Cymothoe egesta* females, i.e. a behavioural experiment. Both of these methods require 2-(3'-cyclopentenyl)glycine in order to be successful, therefore acquiring this chemical compound was one of the first priorities of this investigation.

4.1 OBTAINING 2-(3'-CYCLOPENTENYL)GLYCINE

In order to test Clausen and coworker's hypothesis of 2-(3'-cyclopentenyl)glycine – CPG in short – being a chemical by which Cymothoe butterflies recognize their Rinorea host, we require this chemical in a purified form. The behavioral experiment is set up using CPG to test for its oviposition stimulatory effect directly, whereas the chemical analysis tests for the presence of CPG in Rinorea by comparing plant material to a purified CPG standard. Obtaining 2-(3'-cyclopentenyl)glycine however, is not an easy task.

We had hoped to obtain a sample of this purified CPG from the Royal Danish School of Pharmacy, where Clausen and coworkers did their original research. After several weeks, the head of this department, Jerzy Jaroszevski, was kind enough to send us a sample of powdery white crystals in a small vial marked simply "cyclopentenylglycine". We were unsure which exact type of cyclopentenylglycine this was so I identified the substance with the aid of Bart Rijksen of the Laboratory of Organic Chemistry using H NMR analysis.

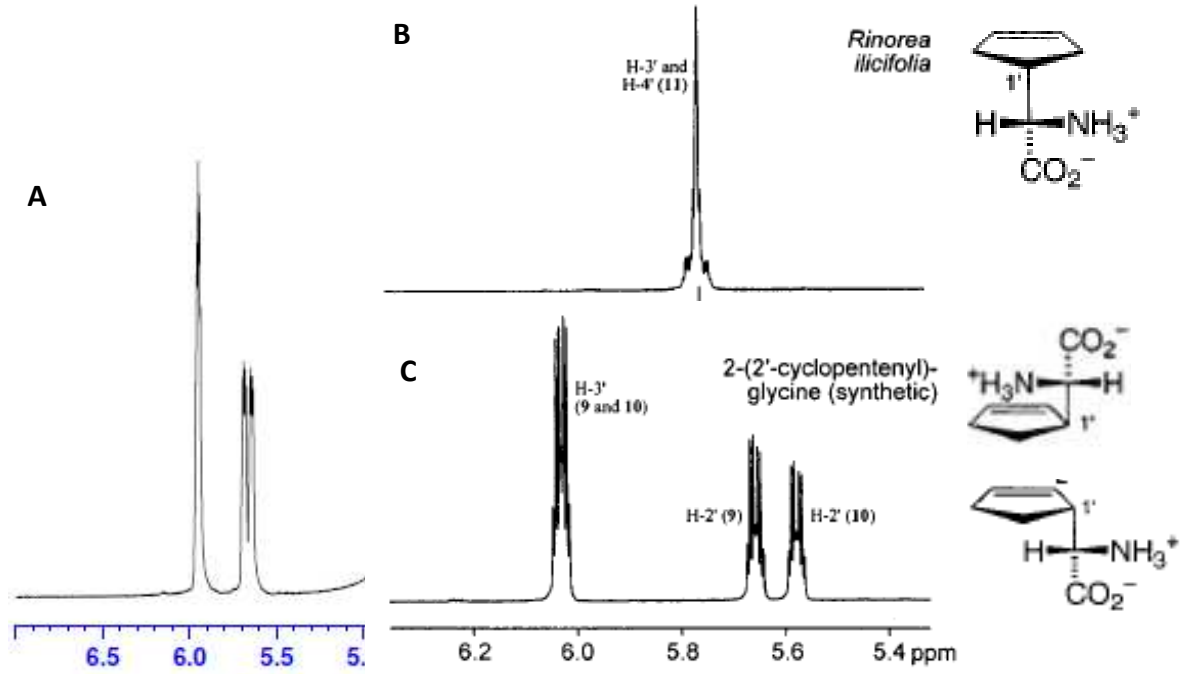


Figure 4.1 Comparison of H NMR spectra. **A** shows part of the spectrum made during this study. **B** and **C** are from Clausen et al (2002), with corresponding structural formulae. This shows spectra A and C are obtained from the same synthetic 2-(2'-cyclopentenyl)glycine. There is a slight shift in ppm values between A and C because we used CD3OD as solvent whereas Clausen and coworkers used D_2O for this spectrum. Values provided in Andersen, Nielsen & Jaroszevski (2000) for 2-(2'-cyclopentenyl)glycine in CD_3OD do match my values precisely.

For this, several crystals from the vial were dissolved in deuterized methanol (CD_3OD). The solubility was observed to be rather low in methanol, and the sample was heated to facilitate a good level of solution. The H NMR spectrum we obtained was unlike that described of 2-(3'-cyclopentenyl)glycine as described in Clausen et al (2002), see figure 4.1. It did however perfectly match the spectrum of synthetic 2-(2'-cyclopentenyl)glycine as provided in Andersen, Nielsen & Jaroszevski (2000). Jerzy Jaroszevski later confirmed that this was the chemical he had sent. No more 2-(3'-cyclopentenyl)glycine was available at the Royal Danish School of Pharmacy, which forced us to use 2-(2'-CP)G instead. Since this molecule is nearly identical to 2-(3'-CP)G in structure, I expect its hydrophobic interactions and mass – as opposed to NMR interactions due to a shift in the double bond – to be identical. Therefore, I expect its LC-MS spectrum to be practically identical. For the behavioral experiment, a difference in the reaction of butterflies is possible.

A manner of obtaining 2-3-CPG by means of chemical synthesis has been investigated. For this, I was eventually referred to Peter Botman of the chemical company Chiralix B.V. In his opinion, synthesis of CPG is likely to be possible by oxidizing 3-cyclopentene-1-carboxylic acid to an aldehyde and subsequently using Strecker amino acid synthesis to create a racemic mix of 2-(3'-cyclopentenyl)glycine. Because this synthesis has not been performed before with this specific chemical, a two week research period is required, costing around €3000 for the production of around 50-100mg of racemic chemical. Should it be necessary, converting the racemic mix of L- and D-CPG into a pure L-CPG – such as would be found in R. ilicifolia (Clausen et al, 2002) – would take another two weeks with Chiralix B.V. and include increased costs (Peter Botman, research director at Chiralix B.V., personal communication). Gipson, Skinner and Shive (1965) describe an alternative method of synthesizing 2-3-CPG, but according to Peter Botman, this method is apparently less efficient and more expensive than the method proposed by him.

4.2 TESTING FOR THE PRESENCE OF CPG USING LC-MS

This chapter attempts to test the CPG hypothesis using approach 2B. If *C. egesta* females use CPG as host recognition chemical, as Clausen and coworkers (2002) suggest, then it would seem likely that accepted *R. ilicifolia* leaves such as used in the previous chapter, contain higher levels of this chemical than the rejected ones. Testing for the presence of a specific chemical using this method is possible, using the metabolic profile, i.e. the retention time and molecular weight, of pure CPG therein, to identify this chemical in the metabolic profiles of *R. ilicifolia* leaf extracts. Subsequently, the levels, i.e. the intensities, of this chemical can in theory be compared between extracts of leaves which have displayed different bioactivity. Nymphalid butterflies have been found to be able to detect and discriminate between different levels of oviposition stimulant (Pereyra & Bowers, 1988). A difference in CPG level can therefore be responsible for host discrimination between old and young leaves.

METHOD

I had initially attempted to test for the presence CPG using LC-MS in negative mode, see chapter 3. Unfortunately, it turned out CPG does not readily take a negative charge, so I decided to focus more on the untargeted metabolomics approach, also because funding did not permit the analysis of all the samples in positive detection mode. Later, I prepared another CPG solution which was used in the orbi-trap device, which is a type of LC-MS that switches between positive and negative detection modes, which showed a better detection of CPG in positive ion detection mode, which – in consultation with Ric de Vos and Bert Schipper, operator of the LC-MS at PRI bioscience – led me to perform the experiment described here.

In positive detection mode, I analyzed six samples, four leaf extracts, one pure CPG solution and one mix of leaf extract and pure CPG solution. The latter was added in order to test whether CPG would not be competitively excluded by other ions in the extract. The concentration of CPG was equal in both the mix and the pure extract, as the latter was diluted to half the original concentration.

For CPG stock solution, 0.7 mg of CPG was weighed and dissolved into 1 mL of 75 MeOH, 25% H_2O and 0.1% FA and sonicated for 15 min, providing a stock solution of 0.7 mg/mL. Plant extracts included in this analysis were two accepted, young leaves 12 and 14 and two rejected, old leaves 24 and 17A. These extracts were made from dry material using the same method as in chapter 3. One run was added with a 1:1 mix of leaf 12 extract and CPG stock, in order to test for ion suppression of CPG by plant metabolites. The pure CPG sample therefore also consisted of two-fold diluted stock solution. The LC-MS method used was the same as described in chapter 3, with the exception of the detection mode.

As a measure of checking the data, the actual CPG content in *R. ilicifolia* leaves may be inferred by comparing the intensity measured in the pure CPG solution with known concentration, to that found in leaf extract, and subsequently calculating CPG concentration therein.

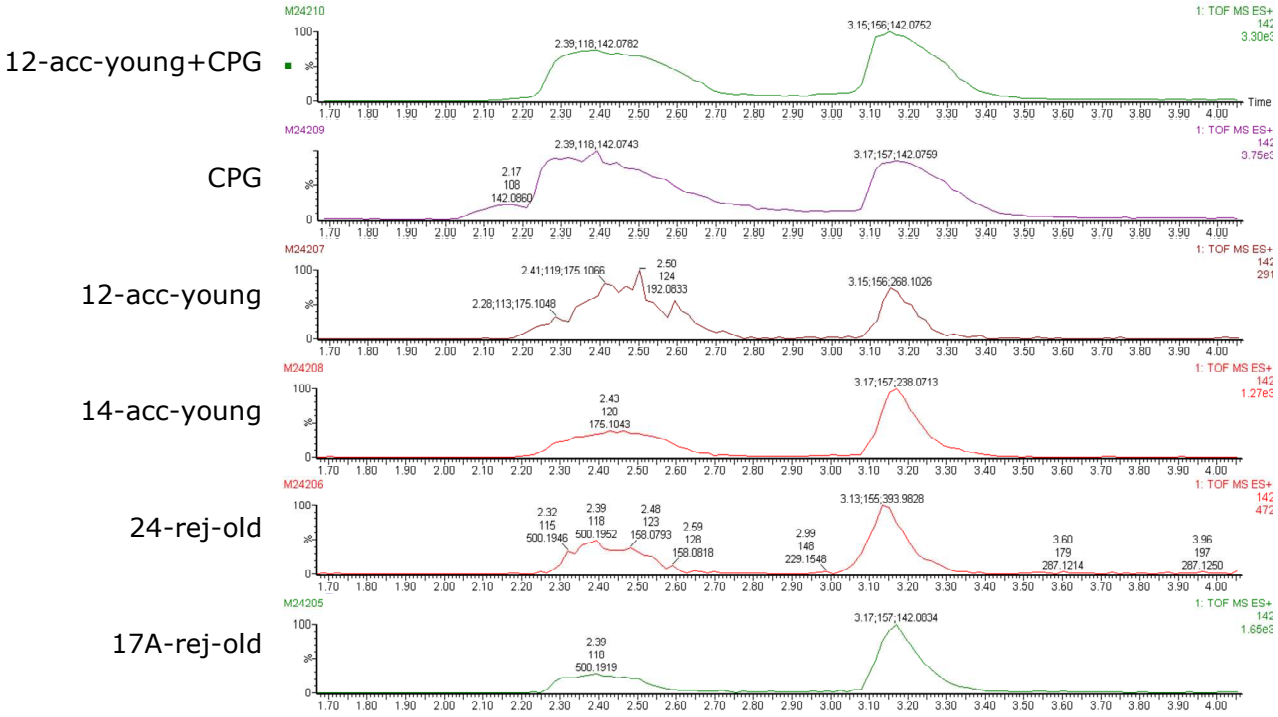


Figure 4.2 Chromatograms of the runs in positive mode, filtered for m/z=142D, i.e. the approximated molecular mass of CPG in positive ion detection mode. Retention time is on the horizontal, and intensity on the vertical axes.

RESULTS

In positive detection mode, CPG could be detected fairly well using LC-MS, see figure 4.2. **Recommendation:** for measuring the level of CPG using LC-MS, use positive ion detection mode. The second peak represents CPG eluting with the liquid phase. The first peak at RT=2.4min, which apparently also contains some CPG, can be explained by the polar nature of the molecule. It probably has some affinity for the MeOH solvent, the initial burst of which passes through the column very rapidly, taking some CPG with it. The CPG that does stay bound to the column, follows quite shortly, at around RT=3.17min. This peak apparently exists in all six runs, showing conclusively the presence of CPG in my samples.

The intensity of CPG measured in the LC-MS in the pure solution was 3177, which corresponds to an actual CPG concentration of 350µg/mL. Measured intensities of CPG in the leaf extracts ranged from 201 to 1647, which inferred from the pure solution would correspond to actual concentrations of approximately 23.5µg/mL and 192µg/mL respectively. The leaf extracts were prepared using a dry leaf matter to solvent ratio of 57mg/mL, meaning the CPG content range in dry leaf material calculated from these measurements is approximately 0.04%-0.3%.

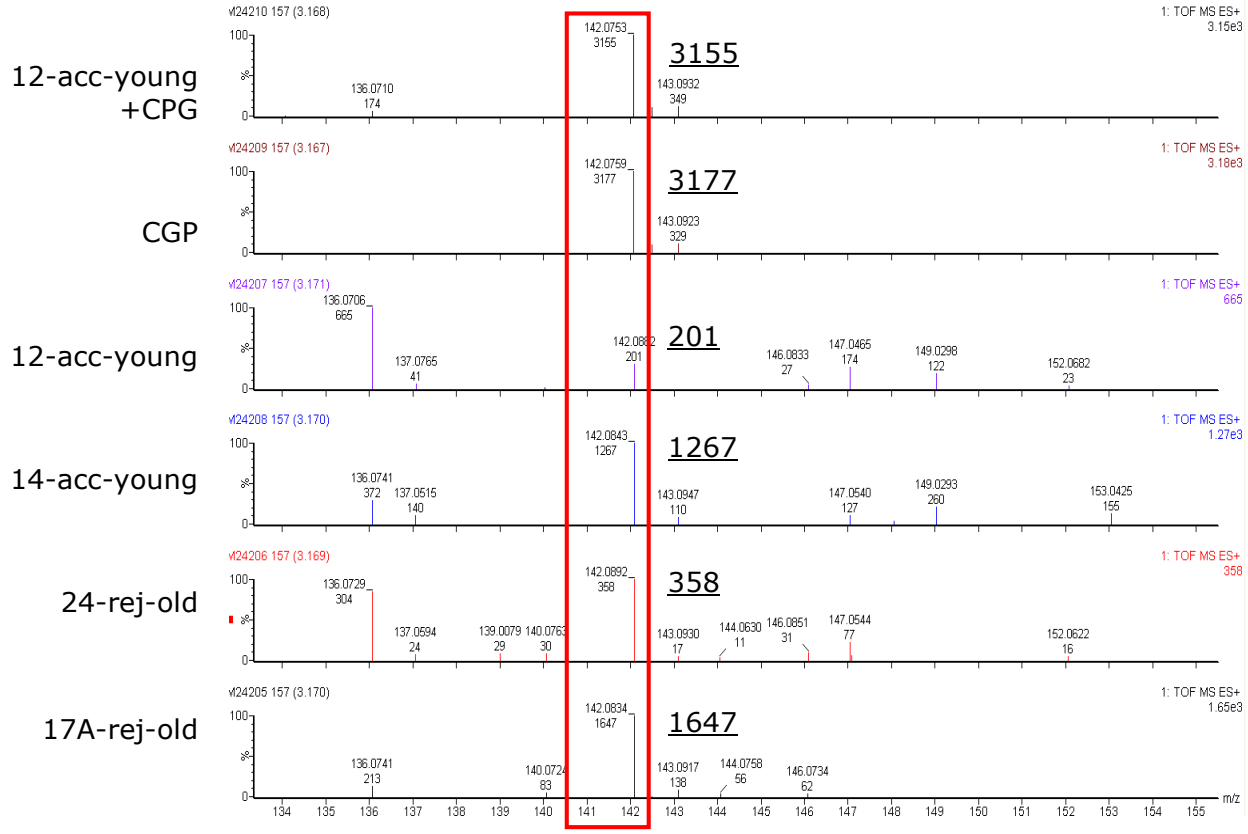


Figure 4.3 Comparison of mass spectra from the runs in positive mode at RT of CPG at RT≈3.17min. Inside the red box is the mass peak of CPG ($C_7H_{11}NO_2$) at m/z≈142.08D in positive ion detection mode. The underlined figures right of this box are the intensities (TIC) at which this peak elutes in each metabolic profile.

DISCUSSION

Interestingly, the highest level of CPG found using LC-MS (approach 2B) (TIC=1647) was in the extract of a rejected, old leaf (17A-rej-old), whereas the lowest CPG content (TIC=201) was measured in a young, accepted leaf (12-acc-young). Conversely, the other young, accepted leaf (14-acc-young) has a CPG content some six-fold higher (TIC=1267) whereas the other old, rejected leaf (24-rej-old) has a CPG content approximately 4-fold lower than 17A-rej-old (TIC=358). A conclusion that can be drawn from this analysis, however statistically insignificant with only four extracts tested, is that there appears to be no correlation between CPG content and leaf age or butterfly behaviour. More extracts should be analyzed for statistical testing that may further strengthen or refute this conclusion. Given the premise that butterflies select the oviposition site with the highest levels of oviposition stimulant, these results do not corroborate the hypothesis that CPG is such an oviposition stimulant to *C. egesta.*

Comparing these measurements to the CPG content of 0.07% found by Clausen and co-workers (2002) in dried *R. ilicifolia* leaf material, i.e. 144mg of CPG purified from 200g material, suggests that the figures obtained in my analysis can be fairly accurate. A factor that may be influencing this accuracy is the possibility of ion suppression due to the presence of other ions in the leaf extracts, a factor that may be assessed by comparison of the LC-MS run of the mix of 12-acc-young extract and CPG solution to the pure CPG solution and 12-acc-young extract. It is possible that some ion suppression exists as the mix shows a slightly lower CPG content than the pure solution, whereas one would expect the CPG intensity to be somewhat higher even as some CPG from 12-acc-young is added to the mix. This ion suppression may mean that the actual CPG content

is somewhat higher in the dry leaf matter. However, figures may fall within standard error, which cannot be tested in this small sample size.

Another factor that may cause the actual CPG content in *R. ilicifolia* leaves to be underestimated in this analysis is the possibility that not all CPG in the leaf material was fully dissolved into the solvent. Similarly, the CPG content of 0.07% measured by Clausen and co-workers may also have been an underestimation as their purification efficiency has logically been a factor limiting the amount of CPG they have measured in dry leaf material. Nevertheless, CPG contents as measured in this study using LC-MS appear to be consistent with the content inferred from the purification procedure using Thin Layer Chromatography and fractionation as used by Clausen and co-workers (2002).

4.3 BEHAVIORAL EXPERIMENT

Behavioral experimentation has been used extensively to find the oviposition stimulants for Lepidoptera (Schoonhoven, Van Loon & Dicke, 2005; Honda, 1995). In fact, should a method like untargeted metabolomics yield a positively identified metabolite that correlates – in this case – with butterfly ovipositing behavior, a behavioral bioassay would ultimately be necessary to prove this metabolite's biological significance. In this case I attempted to test the biological significance of 2-(3'-cyclopentenyl)glycine (CPG). In order to test this hypothesis, a method for performing an ovipositing assay had to be chosen that had a chance of succeeding under primitive field conditions. This part of the experimental approach was also a pilot of this type of experimentation on *Cymothoe,* which has never been done before. Even if *C. egesta* females laid no eggs on artificial substrate that offered pure CPG, the test would still be successful if the positive control received oviposition, as the experimental method could be repeated in the future with different chemicals.

I looked at several methods of performing oviposition experiments available in the literature, and preference was given to those involving Nymphalidae, such as those used to identify the oviposition stimulants listed in table 1.1. Most of these involve extracting plant metabolites from leaf material, and applying this to an artificial substrate. I however felt this would be too difficult to perform under field conditions, and my method was therefore modified from Pereyra & Bowers (1988) and Hovanitz & Chang (1964), who used plates of agar to simulate leaves. The advantage of this method was mainly that Pereyra and Bowers had used dried, ground *Plantago lanceolata* material directly added into the agar, to induce oviposition in the Buckeye butterfly *Junonia coenia*, which I felt confident to repeat for dried and ground *R. ilicifolia* leaf material.

The behavioral experiment originally proposed was a binary (two choices) choice-experiment using artificial leaves to test for the oviposition stimulatory effect of 2-(3'-cyclopentenyl)glycine (CPG). Such choice tests are a sensitive method of establishing the bioactivity of a chemical (Schoonhoven, Van Loon & Dicke 2005).

METHODS AND MATERIALS

Catching, keeping alive, and optimizing conditions: The experiments were performed using *C. egesta* female butterflies that were caught with a butterfly net in the rainforest of Kakum NP. Preferably, butterflies were used that were confirmed to be capable of egg-laying, i.e. which had been observed to oviposit on *R. ilicifolia.* This required tracking butterflies through the undergrowth until oviposition was observed, then simultaneously catching the female and gathering the leaves for the metabolomics approach described in chapter 3. Because of the difficulties involved herein I also subjected some female *C. egesta* to experimentation that had not shown to be capable of egg laying, on the advice of Robin van Velzen.

First, survival of butterflies in captivity was tested and the conditions, butterfly food and location of the cage were optimized accordingly. In an experiment testing tropical butterfly longevity, Molleman and coworkers (2007) were able to keep *Cymothoe* species alive for up to 38 days in captivity, in a cage placed close to the forest's edge. From field reports of Robin van Velzen, keeping *Cymothoe* butterflies alive in captivity is possible using a cage and some (rotten) fruit, which is the natural food of *Cymothoe* (Larsen, 2005). I kept butterflies alive in captivity in a pop-up cage. Experimentation was also performed inside these pop-up cages, of which I had six at my disposal so I could conduct several experiments at once.

Since I did not stay at the Kakum Lodge, I brought the cages to the garden of my guestroom where, and adapted the placement of these cages to the resulting lifespans and butterfly behaviour. The first time, I hung the cage on the washing line. To simulate tree cover and temperature found in the forest, the cage was later hung under a cocoa tree that grew in the garden. **Recommendation:** Eventually, cages were placed inside

the rainforest, which is probably the best place for experimenting with wild-caught butterflies.

Ovipositing choice experiment: Once butterflies could be kept alive and fed inside the pop-up cages, I initiated the choice experiment using artificial ovipositing substrate: agar discs in 5.5 cm Petri dishes (Pereyra & Bowers, 1988). 0.5 g of noble (i.e. without added nutrients) agar is cooked with 15 mL of water and experimental components are added. This can be dried, ground *R. ilicifolia* material for a positive control or a CPG solution. For oviposition testing of *Junonia coenia* (Nymphalidae), Pereyra and Bowers used 0.5 g of dried *Plantago lanceolata* leaf material and 0.01 g of iridoid glycoside (IG) in their experiments. IG content was then 2% of dry leaf weight, which apparently is not an uncommon IG content in *P. lanceolata* (Pereyra and Bowers, 1988). Clausen et al found only 0.072% of CPG in dry *R. ilicifolia* leaf material, which in this set-up would come down to 0.36 mg CPG together with 0.5 g agar. Positive controls were made using 0.5 g of agar and 0.5 g of dried *R. ilicifolia* leaf material whereas the negative controls had no additives. Both the positive control and the CPG test agar were colored using green food coloring. 10 mg of CPG crystal was dissolved in 25 mL filtered and demineralized water. 25 eppendorff tubes were filled with 1 mL of this 0.4 mg/mL stock solution, so that agar plates could be poured with pre-dissolved CPG under field conditions.

For experimentation, agar discs were initially placed in a holder made from wire that elevated the discs some 10 cm off the cage floor, see figure 4.4B&C, to simulate a branch with free hanging leaves. Later, I heightened this holder to the top of the cage, following results from the experiment with live *R. ilicifolia* branches described below, seeing as these branches reached higher into the cage.

Potted *R. ilifolia:* Because the artificial substrate described above yielded no ovipositions (see Results), a series of extra tests were performed to test whether butterflies would lay eggs at all in captivity. First, I exhumed two *R. ilicifolia* specimens, placed these in improvised plastic pots that I introduced into the cage and presented to female *C. egesta*, see figure 4.4D. Having observed young leaf preference from ovipositing *C. egesta* females in the forest, I made sure the second of these exhumed plants had young leaves.

Young Leaf Preference Experiment: Later, cages were taken into the forest and hung over live *R. ilicifolia* specimens with a branch inside the cage and *C. egesta* females were left undisturbed for two days like this, see figure 4.4E. The branch was inspected prior to the experiment to make sure that it had both young and old leaves, and that no older eggs were already present. Results gained in this manner led me to repeat this experiment several times, in an attempt to further corroborate suspicions of young leaf preference in *C. egesta* females. The number of leaves and the number of eggs per leaf were recorded, as well as the type of leaves, i.e. young or old.

Tabel 4.1 Overview of butterflies caught and the experimental treatments to which they were subjected. The circumstances in which these butterflies were caught are described in table 2.1

but-terfly code	Catch date MM/DD	Prior egg laying observed?	Experimental treatment	Experiment date (MM/DD)	Observation/result/remark
♂1	11/09	N/A	Survival test, cage outside the guesthouse on the washing line with orange	11/09	Feeding (see figure 4.4A)
				11/10	Dead
♀1	11/09	No	Survival test, cage outside the guesthouse on the washing line with orange	11/09	Dead
♀2	11/11	Yes	Agar disc choice experiment: blanc vs. positive control, cage placed outside guesthouse under cover of cocoa tree (see figure 4.4A & C)	11/12	No eggs on agar Feeding on coconut
			Introduced potted *R. ilicifolia* of 11/13 (see table 2.1 and figure 4.4D)	11/13	No eggs on potted *R. ilicifolia*
♀3	11/25	Yes	Introduced ♀3 in cage with potted *R. ilicifolia* of 11/13 and coconut, banana and cocoa. Observed ♀3 for 5 hours and scored behavior	11/26	No eggs or oviposition behavior, feeding on banana
♀4 ♀5	11/27 11/27	No No	Introduced potted *R. ilicifolia* of 11/27 (see table 2.XX), that had young leaves, to ♀3, ♀4 and ♀5	11/28-29	No eggs on the potted plant. ♀3 observed feeding on banana.
♀3 ♀4 ♀5	11/25 11/27 11/27	Yes No No	Test for oviposition in the forest environment on natural host in captivity, as well as a test of young leaf preference. ♀3, ♀4 and ♀5 placed in a cage hung over an *R. ilifolia* specimen with a branch inside the cage. The cage was left undisturbed inside the forest for two days.	11/30	Start test, no oviposition observed that day
				12/03	One butterfly had died. 35 eggs were found on leaves in the cage:

#eggs	#old leaves	#eggs	#young leaves	Total #eggs
0	52	4	2	35
1	7	6	1	
2	3	8	1	

♀6	12/03	No	Agar disc choice experiment: blanc vs. positive control, cage placed in the forest. Agar disc were placed on raised standard reaching into the top of the cage, following the result of from the previous experiment.	12/03	Start experiment
				12/05	No eggs on agar
			Agar disc choice experiment: blanc vs. CPG, cage as in previous experiment.	12/05	Start experiment
				12/06	No eggs on agar
			Test for oviposition on natural host, with cage hung over *R. ilicifolia* with a branch inside the cage.	12/06	Start experiment
				12/09	No eggs, unsure whether ♀6 was capable of oviposition
♀7	12/06	Yes	Test for oviposition in the forest environment on natural host in captivity, as well as a test of young leaf preference. ♀7 was placed in a cage hung over an *R. ilicifolia* specimen with a branch inside the cage (see figure 4.4E). I gathered the young accepted leaves as samples 13A and 13B, as well as two old rejected leaves as samples 13B and 13C, see table 3.1.	12/06	Start test
				12/10	9 eggs were found on leaves in the cage (see figure 4.4F):

#eggs	#old leaves	#eggs	#young leaves	Total #eggs
0	18	4	1	9
1	2	3	1	

			Agar disc choice experiment: blanc vs. positive control, as in the ♀6 12/03 experiment.	12/10	Start experiment
				12/13	No eggs on agar

Table 4.1 continued

♀8	12/07	No	Test for young leaf preference, as in the ♀7 12/06 experiment	12/07	Start experiment
				12/10	♀8 dead, no eggs so unsure if she was capable of ovipositing
♀9	12/10	No	Test for young leaf preference, as in the ♀7 12/06 experiment	12/10	Start experiment
				12/13	No eggs
			Test for the ovipositing capability of ♀9, using an *R. ilicifolia* that had been accepted (oviposited upon) earlier by a free flying *Cymothoe*	12/13	Start experiment
				12/16	No eggs, ♀9 nearly dead, unsure if this individual was capable of ovipositing
♀10	12/12	Yes	Test for young leaf preference, and whether a young leaf of a once rejected *R. ilicifolia* could be accepted. ♀10 had rejected an older leaf of *R. ilicifolia* (sample 17A_rej_old, see table 2.1). The cage was hung over a branch of this same *R. ilicifolia* individual that had some young leaves. I gathered the young accepted leaf as sample 17B, as well as two old rejected leaves as samples 17C and 17D, see table 3.1.	12/12	Start experiment
				12/15	1 egg was found on the leaves in the cage:

#eggs	#old leaves	#eggs	#young leaves	Total #eggs
0	14	0	1	1
		1	1	

RESULTS AND DISCUSSION

During my field work in Kakum NP, I caught a total of 10 female- and one male *C. egesta*, which were kept in a pop-up cage and experimented with. Table 4.1 provides an overview of these butterflies and the treatments they were subjected to. As stated in chapter 2, overall numbers of butterflies were low, presumably due to persisting rains. Secondly, the difficulty of tracking butterflies through the undergrowth until they laid an egg, whilst simultaneously gathering leaves for the chemical analyses, made it hard to capture fertilized females.

The experiments I performed were not as sequential as they appear below. Rather, it was an empirical process of ruling out certain circumstantial factors (such as surroundings or height of the agar disc standard) that could be influencing the outcome of the choice experiment, that eventually led me to use the potted *R. ilicifolia* specimens and the experiment in the rainforest.

Survival in captivity: For most of the time, captive butterflies survived well on fed bananas, but orange and coconut were also eaten. Butterflies could survive in the pop-up cage for several days under cover of cocoa trees in the garden of my guesthouse or in the rainforest, as long as some shade was provided. Tree cover proved to increase the longevity of captive butterflies, as the longest living captive female - ♀3 from table 4.1 - survived for approximately 8 days in a cage that was hung under a cocoa tree or was placed inside the forest.

Agar experiment: Unfortunately, neither the agar discs with the positive control nor those with CPG yielded any oviposition during the whole period of experimentation. There can be several causes for this.

Butterflies appeared not to respond to the agar as though these were leaves at all. This may be due to the stress of captivity, but also because wild-caught butterflies have already habituated to actual leaves.

With regard to the positive control, there was bacterial and fungal growth on the agar. In hindsight, I should have anticipated this, because although noble agar itself contains no nutrients apart from agarose (indeed, few micro-organisms would grow on the negative control or CPG test discs), adding dried plant material to the noble agar allowed for the development of bacteria and fungi. Whether this is what repelled the butterfly, or whether the butterfly simply did not respond to discs of agar as to actual leaves is unsure, although I did not observe butterflies alighting on the agar discs at all, indicating the latter.

Because of the result of the young leaf preference experiment, I decided to also perform the choice experiments with agar discs inside the rainforest. Because the branch reached higher up in the cage than the 10 cm of the petri dish holder, I used wire to make a taller holder. Unfortunately, this still did not yield any eggs and if anything, the humid rainforest environment only further promoted the bacterial growth on the agar.

I only used one agar disc with CPG in the experimentation, with ♀6. The reason for this was that I tried to be conservative with this chemical, and I first wanted to make sure the butterflies would oviposit in captivity at all. I had wasted much time with the experimentation at the guesthouse rather than inside the rainforest, and once I had established this to be a better location, I also wanted to test the young leaf preference.

Another cause the refusal of ♀6 to oviposit on the CPG agar, may be that this disc contained 2-(2'-cyclopentenyl)glycine rather than 2-(3'-cyclopentenyl)glycine. An investigation by Gipson, Skinner and Shive (1965) proves that an *in vivo* difference exists between 2-2'-CPG and 2-3'-CPG as functional antagonists for different amino acids in *E. coli*. Also, this may be the very basis of differential host specificity, as was mentioned in chapter 1.5.

However, since the female I had offered CPG to had not been confirmed to be capable of oviposition, I can draw no conclusion on the stimulant effect on oviposition by CPG, based on this experiment.

Potted *R. ilicifolia:* The two exhumed *R. ilicifolia* specimens that were placed inside the cage in improvised plastic pots also did not receive any eggs from captive females. During a 5 hour observation of the behaviour of ♀3, this female never alighted on the leaves of the potted *Rinorea*, although I did observe her feeding. Having noticed the young leaf preference in wild females, I made sure the second of these exhumed specimens had young leaves. This specimen turned out to be a vegetative offshoot from a larger individual, see chapter 2.3, and – being disconnected from its main water source - its young leaves quickly withered, which may also have explained the lack of oviposition on this young specimen.

Young leaf preference experiment: The first young leaf preference test for oviposition by captive *C. egesta* females (♀3, ♀4 and ♀5) in the forest yielded a total of 35 eggs on the branch inside the cage. This shows that in their natural habitat, wild-caught butterflies in captivity will lay eggs more readily than in the unnatural surroundings, probably due to climatic factors, the lack of disturbance from children or livestock as opposed to the village, or the fact that a live, healthy *Rinorea ilicifolia* was available for oviposition rather than potted plants or agar discs.

The distribution of the 35 eggs showed a striking preference for young apical leaves over older leaves. The branch that was placed inside the cage had 66 leaves in total, of which only 4 were young apical leaves. Yet, these four leaves had 22 eggs with 4-8 eggs per leaf. Clearly, these young leaves were accepted and oviposited upon more. The 13 eggs on the older leaves were spread out across the branch with only 1 or 2 eggs per leaf. Repeating of this experiment with ♀7 showed a similar result, with 9 eggs laid in total with 7 eggs laid on the two young leaves on the branch and only two laid on older leaves.

In order to increase the number of butterflies I could use, I started catching females of which I had not observed oviposition in freedom, because it was much easier to just catch them quickly without having to track them through the undergrowth until they oviposited. Also, now it was known that *C. egesta* would lay on live *Rinorea* in captivity, I still wanted to test agar discs. One such female, ♀6, did not lay any eggs in captivity, neither on agar nor subsequently on *R. ilicifolia*. Another, ♀8, was first placed over a *R. ilicifolia* specimen but was dead after two days in the cage without having laid any eggs.

For both ♀8 and ♀6, not laying eggs in captivity could either be because they were physically incapable of egg-laying (because they had not yet been fertilized or had already laid all their eggs) or because the *Rinorea* I placed them over were simply not recognized as suitable hosts and subsequently rejected. The latter scenario would provide me with the basis for the intraspecific comparison for which I could use these *Rinorea*. In order to test this, I placed a third such female, ♀9, which I had also not observed ovipositing, in a cage with a branch of a *R. ilicifolia* specimen on which I had found a *Cymothoe* egg and so was sure of its attractiveness. Unfortunately, no eggs were laid on this specimen by ♀9.

The last female, ♀10, provided some conclusive proof of young leaf preference. I had observed this specimen accept a young leaf of a different *R. ilicifolia* individual after it had rejected an older leaf of *a R. ilicifolia* specimen which also had a few young leaves that had not been tested by this female. I decided to hang the cage with this female over this specimen with the young leaves inside the cage. In this way, I could test whether this *Rinorea* was recognized as host at all. If not, this again provided a basis for intraspecific comparison. However, after two days ♀10 had laid one egg on a young leaf showing that, although this *R. ilicifolia* specimen had previously been rejected, it was still recognized as host, indicating once more that within *R. ilicifolia*, leaf age is more important for oviposition preference than a consistent difference (e.g. genotype) between individuals.

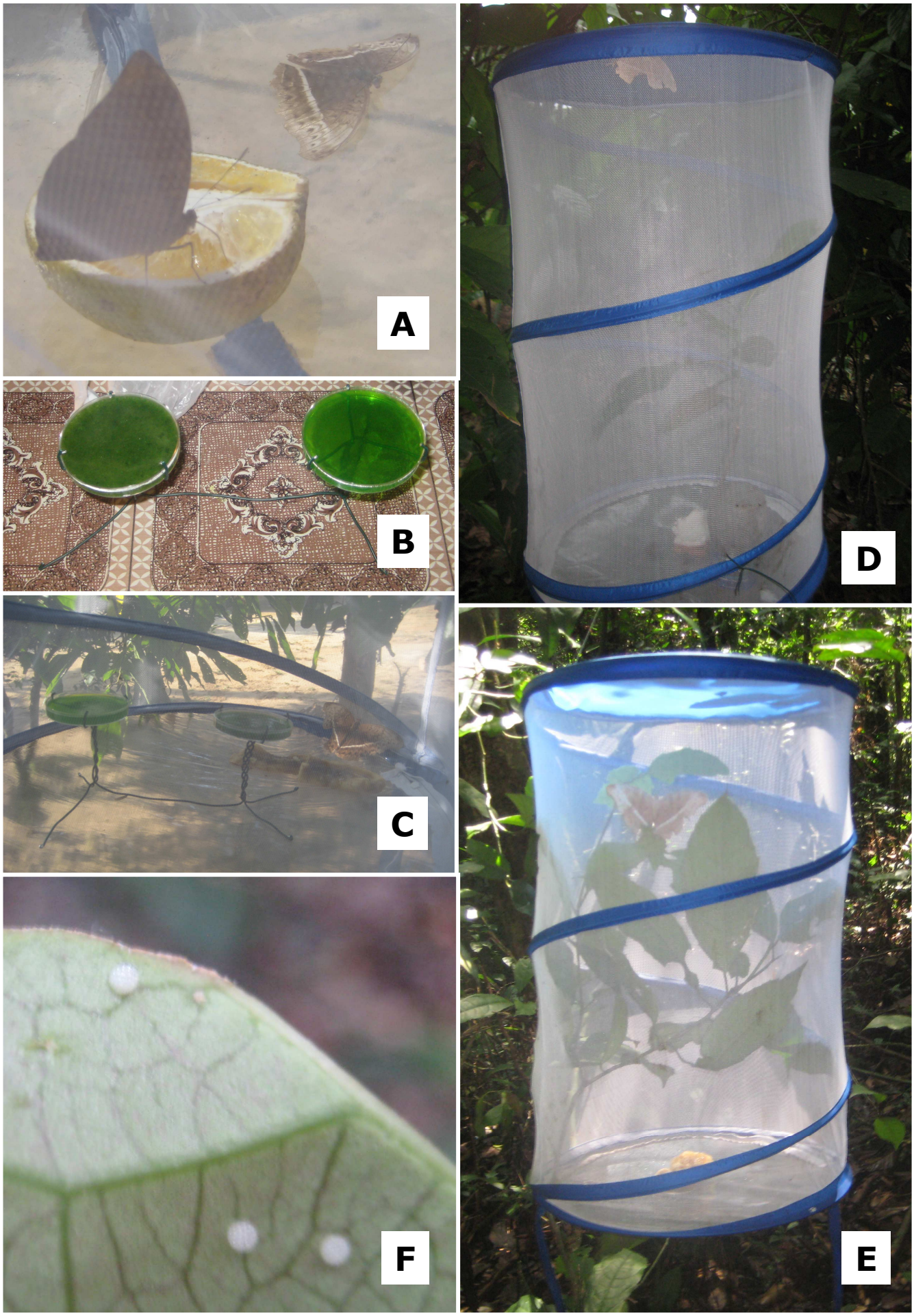


Figure 4.4 A: ♂1 feeding on orange in captivity. **B:** agar discs in holder for choice test. Left: positive control with ground *R. ilicifolia* leaf material; right: negative control with food coloring (CPG disc was indistinguishable). **C:** Choice test set-up for ♀2 (see table 4.1). **D:** potted *R. ilicifolia* introduced to ♀2. **E:** test inside Kakum National Park, with pop-up cage suspended over an *R. ilicifolia* specimen, with a live branch inside the cage presented to ♀7. **F:** 3 of a total 9 eggs laid by ♀7 as a result of the set-up in 4.4E, see table 4.1

CHAPTER 5 CONCLUSIONS

From the two main experimental approaches described in chapter 1.8, the following conclusions have been reached:

Untargeted metabolomics

The aim of the untargeted metabolomics approach was to compare metabolic profiles of *Rinorea ilicifolia* leaves, in order to identify plant metabolites that correlate with observed oviposition behaviour of *C. egesta* females. Using the methods described in this thesis, I selected metabolites that were present in greater abundance in accepted than in rejected leaves and I made attempts to deduce the molecular formulae of some of these metabolites. However, the premise that chemical differences between *Rinorea ilicifolia* individuals cause differential host selection by *Cymothoe egesta*, could not be assumed, following my observations and experimental corroboration of young leaf preference by ovipositing *C. egesta* females. This means that the correlating metabolites found, probably have little to do with host preference, but are simply caused by leaf ontogeny.

Testing of 2-(3'-cyclopentenyl)glycine (CPG) as oviposition stimulant

This approach aimed at testing the hypothesis that CPG is an oviposition stimulant to *Cymothoe egesta.* One approach to this was to test whether a correlation could be found between CPG levels in accepted and rejected leaves and observed butterfly behavior. No such correlation was found, therefore assuming the premise that *C. egesta* select leaves with the greatest amount of oviposition stimulant, CPG does not appear to be such an oviposition stimulant. An increased sample size may however change this conclusion.

The result of the behavioral experiment was inconclusive. No oviposition on the agar disc with CPG was recorded, but I only tried this once. Also, the positive control was also never accepted, so the experimental method of using agar discs as artificial leaves did not appear to be appropriate as a choice experiment for these wild butterflies.

CHAPTER 6 RECOMMENDATIONS

Bioassay-guided fractionation experiments: The "traditional" approach to identifying host recognition chemicals/oviposition stimulants (Prince & Pohnert 2010) may work well to "start from scratch". It is probably best to rear butterflies in captivity for this approach, as it is likely to be a lengthy one.

Choice Assay: Different methods of presenting candidate oviposition stimulants or extract fractionations to female *Cymothoe* may be attempted. Artificial substrates other than agar, such as green sponges (Haribal & Renwick, 1996; Haribal & Feeny, 1998), a green plastic plate (Honda et al., 1997; Honda et al., 2004), and paper (Honda, Nishii & Hayashi, 1997; Honda et al., 2001) can be tried. Chemically "uncomplicated" plants may also function as "artificial" substrate plant such as pea, which is generally thought to contain few deterrents (Van Loon, personal communication).

Metabolomics (levels of comparison): As described in chapter 1.7, pp. 18-21, different levels of comparison may be used to look for chemicals that correlate with butterfly behavior. Such investigations may be performed elsewhere Be aware that, if interspecific variation is investigated, leaf ontogeny should be taken into account as well, i.e. compare only the accepted leaves, which are possibly all young leaves. The dried Cameroonian samples collected by Robin van Velzen may still be used in these types of analyses.

CPG Detection: Test for CPG levels using different methods of detecting non-proteogenic amino acids, such as Thin Layer Chromatography (Clausen et al., 2002). Clausen and coworkers also never analyzed *Rinorea* species other than *R. ilicifolia* so it is unknown whether other species contain the same cyclopentenoid compound or a different one. This is a question that should be addressed, for if cyclopentanoids are indeed used for host recognition by *Cymothoe* as Clausen and coworkers hypothesize, different cyclopentanoids are likely to be present in different Rinorea species. If 2-3-CPG is found in Rinorea species other than *R. ilicifolia*, this would be an indicator that wither cyclopentanoids are not involved in specific host recognition, or that they may be a universal oviposition stimulant to *Cymothoe*, and host specificity may be caused instead by deterrents.

Sexual selection: Quental, Patten & Pierce (2007) hypothesized that host specificity may be caused by sexual selection of the female butterfly on male pheromones sequestered from host plant specific compounds. This may be of influence in *Cymothoe* as well. An approach towards testing this could be to use Gas Chromatography coupled Electro-antennographic detection to detect volatile components emitted by a male butterflies that elicit a response in the antennae of the opposite sex, and subsequently identify these using GC-MS (Nieberding et al., 2008). Whether this is a good method to uncover the mechanism of host recognition remains to be investigated however, possibly testing the sexual selection hypothesis is best performed after the proximate causes of *Cymothoe* host specificity have been studied in more detail.

Experimental recommendations for similar follow-up experiments throughout the report, including (chapter, page):

- Go to Ghana towards the end of November to escape the rainy season (2.2, pp. 27).
- Observe the ovipositing sequence closely, look for the accepted rather than the oviposited leaf (2.2, pp. 29).
- Consider storing dried metabolomics samples in paper bags on silica (3.1, pp. 35)
- Puncture holes in eppendorff lids, do not open them prior to lyophilization (3.1, pp. 35).
- Apply pressure to the syringe gently when filtering LC-MS extract (3.1, pp. 36).
- Use positive ion detection mode when measuring CPG using LC-MS (4.2, pp. 58)
- Experiment with wild-caught butterflies in the rainforest (4.3, 61).

ACKNOWLEDGEMENTS

My gratitude goes out to:

Freek Bakker and Robin van Velzen, for their supervision on my work during this Origin of Thesis and their patience during this process. They have offered me the opportunity to fulfill my dream of conducting fieldwork in a tropical rainforest, which has been a profound personal as well as scientific experience for me;

Charles Owusu and Andrews Kankam for the indispensable assistance they offered me during my fieldwork and for making me feel very welcome in a different culture;

Ric de Vos for his practical supervision on the metabolomics methods that I used, and Maarten Jongsma for his official supervision at PRI Bioscience;

Richard Kankam and family for providing me with a home away from home, deepening my experience of Africa and keeping me well fed;

Renske Onstein for her advice, contacts and friendship;

Gideon Kankam for saving Africa in the future and the rest of the Abrafo kids for getting on my nerves;

Wiebe Draijer, Willem-Albert & Manon Toose and Dennis & Anette van Buuren for providing the Dutch connection in Ghana;

The Island and staff of the Biosystematics group for their compatible humour and many agreeable lunches and coffee breaks;

Bert Schipper, Harry Jonker and Roland Mumm for their advice and assistance during my metabolomics work;

Szabolcs Sáfián for lending me the research permit of BCGhana;

Jerzy Jaroszewski for providing me with cyclopentenylglycine;

And for providing me with the financial support without which this study would not have been possible:

- Hugo de Vries Foundation
- Alberta Mennega Foundation
- Biosystematics Funds

REFERENCES

Achoundong G (1996) Les Rinorea comme indicateurs des grands types forestiers du Cameroun. In J. v. d. Maesen [ed.], *The Biodiversity of African Plants*, pp. 536-544. Kluwer Academic Publishers.

Agrawal AA (2011) EVOLUTIONARY ECOLOGY OF PLANT DEFENCES: Current trends in the evolutionary ecology of plant defence. *Functional Ecology*. Vol. 25, pp. 420–432

Agrawal AA, Salminen JP, Fishbein M (2009) PHYLOGENETIC TRENDS IN PHENOLIC METABOLISM OF MILKWEEDS (*ASCLEPIAS*): EVIDENCE FOR ESCALATION. *Evolution. Vol. 63, No. 3, pp. 663–673.*

Amiet JL, Achoundong G (1996) An example of trophic specialization within the Lepidoptera: The Cameroonian Cymothoe feeding on Rinorea (Violaceae) (Lepidoptera, Nymphalidae). *Bulletin de la Societe Entomologique de France*, Vol. 101, pp. 449-466.

Amiet JL (1997) Trophic specialization and early stages in the Cymothoe: Taxonomic involvement (Lepidoptera, Nymphalidae). *Bulletin de la Societe Entomologique de France.* Vol. 102, pp. 15-29.

Andersen L, Nielsen B, Jaroszewski JW (2000) Synthesis of Epimers of L-Cyclopentenylglycine Using Enzymatic Resolution. *Chirality.* Vol. 12, pp. 665–669

Aublet JBCF (1775) Histoire des plantes de la Guiane Francoise. Didot, Paris.

Bakker FT, van Gemerden BS & Achoundong G (2006). Molecular systematics of African *Rinorea* Aub. (Violaceae). Pp. 33—44 in Taxonomy and Ecology of African Plants, their Conservation and Sustainable Use, Proceedings of the 17th AETFAT Congress, Addis Ababa, Ethiopia. Kew Publishing.

Baur R, Haribal M, Renwick JAA (1998) Contact chemoreception related to host selection and oviposition behaviour in the monarch butterfly, *Danaus plexippus*. *Physiological Entomology*, Vol. 23, pp. 7-19.

Becerra JX (2003) Synchronous coadaptation in an ancient case of herbivory. *PNAS.* Vol. 100, No. 22, pp. 12804–12807

Becerra JX, Venable DL (1999) Macroevolution of insect–plant associations: The relevance of host biogeography to host affiliation. *Proc Natl Acad Sci USA.* Vol. 96, pp. 12626–12631.

Becerra J, Noge K, Venable D (2009) Macroevolutionary chemical escalation in an ancient plant–herbivore arms race. *Proc Natl Acad Sci USA*, 10.1073/PNAS.0904456106.

Berdegué M, Reitz SR, Trumble (1998) Host plant selection and development in Spodoptera exigua: Do mother and offspring know best? Entomologia Experimentalis et Applicata. Vol. 89, No. 1, pp. 57-64

Bernays EA (2001) Neural Limitations in Phytophagous Insects: Implications for Diet Breadth and Evolution of Host Affiliation. *Annu. Rev. Entomol.* Vol. 46, pp. 703-727.

Bernays EA and RF Chapman (1994) Host-plant selection by phytophagous insects. Chapman & Hall, New York.

Clausen V, Frydenvang K, Koopmann R, Jorgenson LB, Abbiw DK, Ekpe P, Jaroszewski JW (2002) Plant Analysis by Butterflies: Occurrence of Cyclopentenylglycines in Passiflaraceae, Flacourtiaceae, and Turneraceae and Discovery of the Novel Nonproteinogenic Amino Acid 2-(3'-Cyclopentenyl)glycine in *Rinorea*. *J. Nat. Prod*, Vol. 65, pp. 542-547.

Cornell HV, Hawkins BA (2003) Herbivore Responses to Plant Secondary Compounds: A Test of Phytochemical Coevolution Theory. *The American Naturalist.* Vol. 161, No. 4, pp. 507-522.

Cramer P (1775-1782) De uitlandsche Kapellen voorkomende in de drie Waereld-Deelen Asia, Africa en America - Papillons exotiques des trois parties du monde l'Asie, l'Afrique et l'Amerique.

Dethier VG (1954) Evolution of feeding preferences in phytophagous insects. *Evolution.* Vol. 8, pp. 32–54.

Dethier VG, Barton Browne L, Smith CN (1960) The Designation of Chemicals in Terms of the Responses They Elicit from Insects. *Journal of Economic Entomology*, Vol. 53, No. 1, pp. 134-136.

De Vos RCH, Moço S, Lommen A, Keurentjes JJB, Bino RJ, Hall RD (2007) Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols.* Vol. 2, No. 4, pp. 778-791.

Ehrlich PR, Raven PH (1964) Butterflies and Plants: A study in Coevolution. *Evolution*, Vol. 18, No. 4, pp. 586-608.

Engler HS, Spencer KC, Gilbert LE (2000) Preventing cyanide release from leaves. *Nature*, Vol. 406, pp. 144-145.

Farell BD, Dussourd DE, Mitter C (1991) Escalation of Plant Defense: Do Latex and Resin Canals Spur Plan Diversification? *American Naturalist.* Vol. 138, No. 4, pp. 881-900.

Feeny P, Städler E, Åhman I, Carter M (1989) Effects of Plant Odor on Oviposition by the Black Swallowtail Butterfly, *Papilio polyxenes* (Lepidotera: Papilionidae). *Journal of Insect Behavior*, Vol. 2, No. 6, pp. 803-827

Futuyma DJ, Agrawal AA (2009) Macroevolution and the biological diversity of plants and herbivores. *PNAS.* Vol. 106, No. 43, pp. 18054-18061.

Gipson RM, Skinner CG, Shive W (1965) Relationship of Structure to Biological Activity of Some Unsaturated Derivatives of Cyclopentenylglycine. *Archives of Biochemistry and Biophysics.* Vol. 111, pp. 264-269.

Haribal M, Renwick JAA (1996) Oviposition stimulants for the Monarch Butterfly: flavonol glycosides from *Asclepias curassavica*. *Phytochemistry*, Vol. 41, No. 1, pp. 139-144.

Haribal M, Feeny P (1998) Oviposition stimulant for the zebra swallowtail butterfly, *Eurytides marcellus*, from the foliage of pawpaw, *Asimina triloba*. *Chemoecology*. Vol. 9, pp. 99-110.

Hawthorne WD, Jongkind CCH (2006) Woody plants of western African forests. A guide to the forest trees, shrubs and lianes from Senegal to Ghana. Royal Botanical Gardens, Kew, Richmond.

Heinz CH, Feeny P (2005) Effects of contact chemistry and host plant experience in the oviposition behaviour of the eastern black swallowtail butterfly. *Animal Behaviour*. Vol. 69, pp. 107-115.

Hekking WHA (1988) Flora Neotropica Monograph 46, Violaceae Part I - *Rinorea* and *Rinoreocarpus*. New York: The New York Botanical Garden

Honda K (1995) Chemical Basis of Differential Oviposition by Lepidopterous Insects. *Archives of Insect Biochemistry and Physiology.* Vol. 30, pp. 1-23.

Honda K, Hayashi N, Abe F, Yamauchi T (1997) Pyrrolizidine Alkaloids mediate Host-Plant recognition by ovipositing females of an old world Danaid butterfly, *Idea leuconoe. Journal of Chemical Ecology.* Vol. 23, No. 7, pp. 1703-1713.

Honda K, Nishii W, Hayashi N (1997) Oviposition Stimulants of Sulfur Butterfly, *Colias erate poliographys*: Cyanoglucosides as synergists involved in host preference. *Journal of Chemical Ecolog.* Vol. 23, No. 2, pp. 323-331.

Honda K, Ômura H, Hayashi N, Abe F, Yamauchi T (2001) Oviposition-stimulatory activity of phenanthroindolizidine alkaloids of host-plant origin to a danaid butterfly, *Ideopsis similis. Physiological Entomology*. Vol. 26, pp. 6-10.

Honda K, Ômura H, Hayashi N, Abe F, Yamauchi T (2004) Conduritols as oviposition stimulants for the Danaid butterfly, *Parantica sita*, identified from host plant, *Marsdenia tomentosa*. *Journal of Chemical Ecology.* Vol. 30, No. 11, pp. 2285-2296

Hovanitz W, Chang VCS (1964) Adult oviposition responses in *Pieris rapae. Journal of Research on the Lepidoptera*. Vol. 3, No. 3, pp. 159-172.

Hübner J (1819) Verzeichniss bekannter Schmetterlinge. Jacob Hübner, Augsburg.

Janz N, Nyblom K, Nylin S (2001) Evolutionary Dynamics of Host-Plant Specialization: A case study of the tribe Nymphalini. *Evolution.* Vol. 55, No. 4, pp. 783–796

Janz N, Nylin S (1998) Butterflies and plants: A phylogenetic study. *Evolution.* Vol 52, No. 2, pp. 486–502.

Janz N, Nylin S (2008) The oscillation hypothesis of host–plant range. *Specialization, Speciation, and Radiation: The Evolutionary Biology of Herbivorous Insects*. Ed. Tilmon K (Univ California Press, Berkeley). Pp. 203–215.

Janz N, Nylin S, Wahlberg N (2006) Diversity begets diversity: host expansions and the diversification of plant-feeding insects. *BMC Evolutionary Biology.* Vol. 6, No. 4

Jermy T (1993) Evolution of insect-plant relationships – a devil's advocate approach. *Entomol. Exp. Appl*. Vol. 66, pp. 3-12.

Kerpel SM, Moreira GRP (2005) Absence of Learning and Local Specialization on Host Plant Selection by *Heliconius erato*. *Journal of Insect Behavior*. Vol. 18, No. 3, pp. 433-452.

Kuntze CEO (1891) *Revisio Generum Plantarum.* Vol. 1, pp. 42. Leipzig

Larsen TB (2005) Butterflies Of West Africa. Apollo Books, Stenstrup.

Lederhouse
RC, Ayres MP, Nitao JK, Scriber JM (1992) Differential use of lauraceous hosts by swallowtail butterflies, *Papilio troilus* and *P. palamedes* (Papilionidae). *Oikos*. Vol. 63, pp. 244-252.

Maier TS, Kuhn J, Müller C (2010) Proposal for field sampling of plants and processing in the lab for environmental metabolic fingerprinting. *Plant Methods.* Vol. 6, No. 6.

McBride CS, Van Velzen R, Larsen TB (2009) Allopatric origin of cryptic butterfly species that were discovered feeding on distinct host plants in sympatry. *Molecular Ecology.* Vol. 18, pp. 3639-3651.

Merz E (1959) Pflanzen und Raupen. Über einige Prinzipien der Futterwahl bei Gross-schmetterlingsraupen. *Biol. Zentr.* Vol. 78, pp. 152- 188.

Moço SIAV (2007) *Metabolomics Technologies applied to the Identification of Compounds in Plants*. PhD thesis, Wageningen University, The Netherlands

Molleman F, Zwaan BJ, Brakefield PM, Carey JR (2007) Extraordinary long life spans in fruit-feeding butterflies can provide window on evolution of life span and aging. *Experimental Gerontology.* Vol. 42, pp. 472–482

Murphy SM, Feeny P (2006) Chemical facilitation of a naturally occurring host shift by *Papilio machaon* butterflies (Papilionidae). *Ecological Monographs.* Vol. 76, No. 3, pp. 399-414.

Nahrstedt A and Davis RH (1981) The Occurrence of the Cyanoglucosides, Linamarin and Lotaustralin, in *Acraea* and *Heliconius* Butterflies. *Comp. Biochem. Physiol*. Vol. 68B, pp. 575 to 577.

Nakayama T, Honda K, Ômura H, Hayashi N (2003) Oviposition Stimulants for the Tropical Swallowtail Butterfly, *Papilio polytes*, Feeding on a Rutaceous Plant, *Toddalia asiatica*. *Journal of Chemical Ecology*. Vol. 29, No. 7, pp. 1621-1634.

Nieberding CM, de Vos H, Schneider MV, Lassance J-M, Estramil N, et al. (2008) The Male Sex pheromone of the Butterfly Bicyclus anynana: Towards an Evolutionary Analysis. PLoS ONE 3(7): e2751. doi:10.1371/journal.pone.0002751

Nishida R (2005) Chemosensory Basis of Host Recognition in Butterflies – Multi-component System of Oviposition Stimulants and Deterrents. *Chem. Senses.* Vol. 1, pp 1293-1294.

Novotny V, Basset Y, Miller SE, Weiblen GD, Bremerk B, Cizek L, Drozd P (2002) Low host specificity of herbivorous insects in a tropical forest. *Nature*. Vol. 416, pp. 841-844

Nyman T (2010) To speciate, or not to speciate? Resource heterogeneity, the subjectivity of similarity, and the macroevolutionary consequences of niche-width shifts in plant-feeding insects. *Biological Reviews.* Vol. 85, pp. 393-411

Nyman T, Paajanen R, Heiska S, Julkunen-Tiitto R (2011) Preference-performance relationship in the gall midge Rabdophaga rosaria: Insights from a common-garden experiment with nine willow clones.

Ecological *Entomology*,
Volume 36, Issue 2, April 2011, pp. 200-211

Ødegaard F (2000) How many species of arthropods? Erwin's estimate revised. *Biological Journal of the Linnean Society.* Vol. 71, pp. 583-597.

Onstein (2010) Evolution of Rinorea (*Violaceae*) as host plant for *Cymothoe* (Nymphalidae): new insights from extended sampling of Ghanaian populations, and development of a new nuclear phylogenetic marker. *WUR MSc thesis*

Parmesan C, Singer MC, Harris I (1995) Absence of adaptive learning from the oviposition foraging behaviour of a checkerspot butterfly. *Anim. Behav.* Vol. 50, pp. 161-175.

Pereyra PC and Bowers MD (1988) Iridoid Glycosides as Oviposition Stimulants for the Buckeye Butterfly, *Junonia coenia* (Nymophalidae). *J. Chem. Ecol.*, Vol. 14, No. 3, pp. 917-928.

Poelman EH, Van Dam NM, Van Loon JJA, Vet LEM, Dicke M (2009) Chemical diversity in Brassica oleracea affects biodiversity of insect herbivores. *Ecology*, Vol 90(7), pp. 1863-1877.

Potter DA, Kimmerer TW (1988) Do holly leaf spines really deter herbivory? *Oecologia.* Vol. 75, pp. 216-221

Prince EK, Pohnert P (2010) Searching for signals in the noise: metabolomics in chemical ecology. *Anal. Bioanal. Chem.*, Vol. 396, No. 1, pp. 193-197.

Quental TB, Patten MM, Pierce NE (2007) Host Plant Specialization Driven by Sexual Selection. *The American Naturalist*, Vol. 169, No. 6, pp. 830–836.

Renwick JAA, Chew FS (1994) Oviposition Behaviour in Lepidoptera. *Annu. Rev. Entomol*, Vol. 39, pp. 377-400.

Renwick JAA, Radke CD (1983) Chemical Recognition of Host Plants for Oviposition by the Cabbage Butterfly, *Pieris rapae* (Lepidoptera: Pieridae). *Environ. Enlomol.* Vol. 2, pp. 446-450

Reudler Talsma JH, Biere A, Harvey JA, Van Nouhuys S (2008) Oviposition Cues for a Specialist Butterfly - Plant Chemistry and Size. *J. Chem. Ecol*., Vol. 34, pp. 1202-1212.

Shikano I, Akhtar Y, Isman MB (2010) Relationship between adult and larval host plant selection and larval performance in the generalist moth, Trichoplusia ni. *Arthropod-Plant Interactions*. Vol. 4, Iss.3 pp. 197-205

Schoonhoven LM, Van Loon JJA, Dicke M (2005) Insect-Plant Biology. Oxford University Press, Oxford.

Stermitz FR, Belofsky GN, David NG, Singer MC (1989) Quinolizide Alkaloids obtained by *Pedicularis semibarbata* (Scrophulariaceae) from *Lupinus fulcratus* (Leguminosae) fail to influence the specialist herbivore *Euphydryas editha* (Lepidoptera). *Journal of ChemicalEeology.* Vol. 15, No. 11

Stevens PF (2001 onwards). Angiosperm Phylogeny Website. Version 9, June 2008 http://www.mobot.org/MOBOT/research/APweb/ last accessed August 2011

Strong DR, Lawton JH, Southwood R (1984) *Insects on Plants* (Harvard Univ Press, Cambridge, MA).

Van Velzen (2007) Evolution of Host-Plant us in Cymothoe (Nymphalidae) Feeding on Rinorea (Violaceae). *WUR MSc thesis*

Thompson, J. N. 1994. The coevolutionary process. University of Chicago Press, Chicago.

Traynier RMM (1984) Associative learning in the ovipositional behaviour of the cabbage butterfly. *Physiological Entomology.* Vol. 9, pp. 465-472.

Traynier RMM, Truscott RJW (1991) Potent natural egg-laying stimulant for cabbage butterfly *Pieris rapae*. *J. Chem. Ecol.*, Vol. 17, No. 7, pp. 1371-1380

Tree of Life Web Project (2002) Malpighiales. Version 01 January 2002 (temporary). http://tolweb.org/Malpighiales/21033/2002.01.01 *in* The Tree of Life Web Project, http://tolweb.org/

Van Velzen R (2006). Evolution of host-plant use in *Cymothoe* (Nymphalidae) feeding on *Rinorea*

(Violaceae). *Wageningen University,* thesis report.

Vorst O, De Vos RCH, Lommen A, Staps RV, Visser RGF, Bino RJ, Hall RD (2005) A non-directed approach to the differential analysis of multiple LC–MS-derived metabolic profiles. *Metabolomics*. Vol. 1, No. 2, pp. 169-180.

Wahlberg N, Brower AVZ (2009) Nymphalidae Rafinesque 1815. Brush-footed Butterflies. Version 15 September 2009 (under constriction). http://tolweb.org/Nymphalidae/12172/2009.09.15 *in* The Tree of Life Web Project, http://tolweb.org

Wheat CW, Vogel H, Wittstock U, Braby MF, Underwood D, Mitchell-Olds T (2007) The genetic basis of a plant–insect coevolutionary key innovation. *PNAS.* Vol. 104, No. 51, pp. 20427-20431

Wiklund C (1975) The Evolutionary Relationship between Adult Oviposition Preferences and Larval Host Plant Range in *Papilio machaon* L. *Oecologia* (Berh). Vol. 18, pp. 185-197.

Wiklund S, Johansson E, Sjöström L, Mellerowicz EJ, Edlund U, Shockcor JP, Gottfries J, Moritz T, Trygg J (2008) Visualization of GC/TOF-MS-Based Metabolomics Data for Identification of Biochemically Interesting Compounds Using OPLS Class Models. *Analytical Chemistry.* Vol. 80, No. 1, pp. 115-122.

APPENDIX PRELIMINARY TESTING OF *R. YAUNDENSIS*

Prior to the field work in Ghana, I performed a test to optimize the sampling protocol using a specimen of *Rinorea yaundensis* from the Burgers' Bush greenhouse of Burgers' Zoo in Arnhem, originally collected by Carel Jongkind in 1996. I tested 4 methods for quenching metabolism: immersion in EtOH and drying with an oven or hairdryer. Dry mass of *R. yaundensis* was also determined (table A1). Freezing in liquid nitrogen was used for a reference treatment. Metabolic profiles showed considerable difference between treatments (figure A1), but the same treatment grouped together. The methods were found to be consistent enough to allow for comparison between extracts of the same quenching method, although the detection of artifactual signal cannot be fully prevented using either method. The ethanol protocol was not used adducts may form to plant metabolites (Ric de Vos, personal communication). The extracts of oven dried leaves in the *R. Yaundensis* samples showed many breakdown products in the later retention times, therefore the plant press and hair dryer method was used to make air dried samples in the field.

Table A1 List of *Rinorea yaundensis* samples treatments

Nr	Treatment	Nr. Discs	Fresh Mass (mg)	Dry mass (mg) (%)
1	Liquid N2	24	1166	n.a
2	EtOh	25	1210	n.a.
3	Oven	20	1025	295 29%
4	Hairdryer	25	1119	357 32%
5	Liquid N2	26	1095	n.a.
6	Hairdryer	27	1251	411 33%
7	EtOH	25	1208	n.a.
8	Oven	50	2027	640 32%

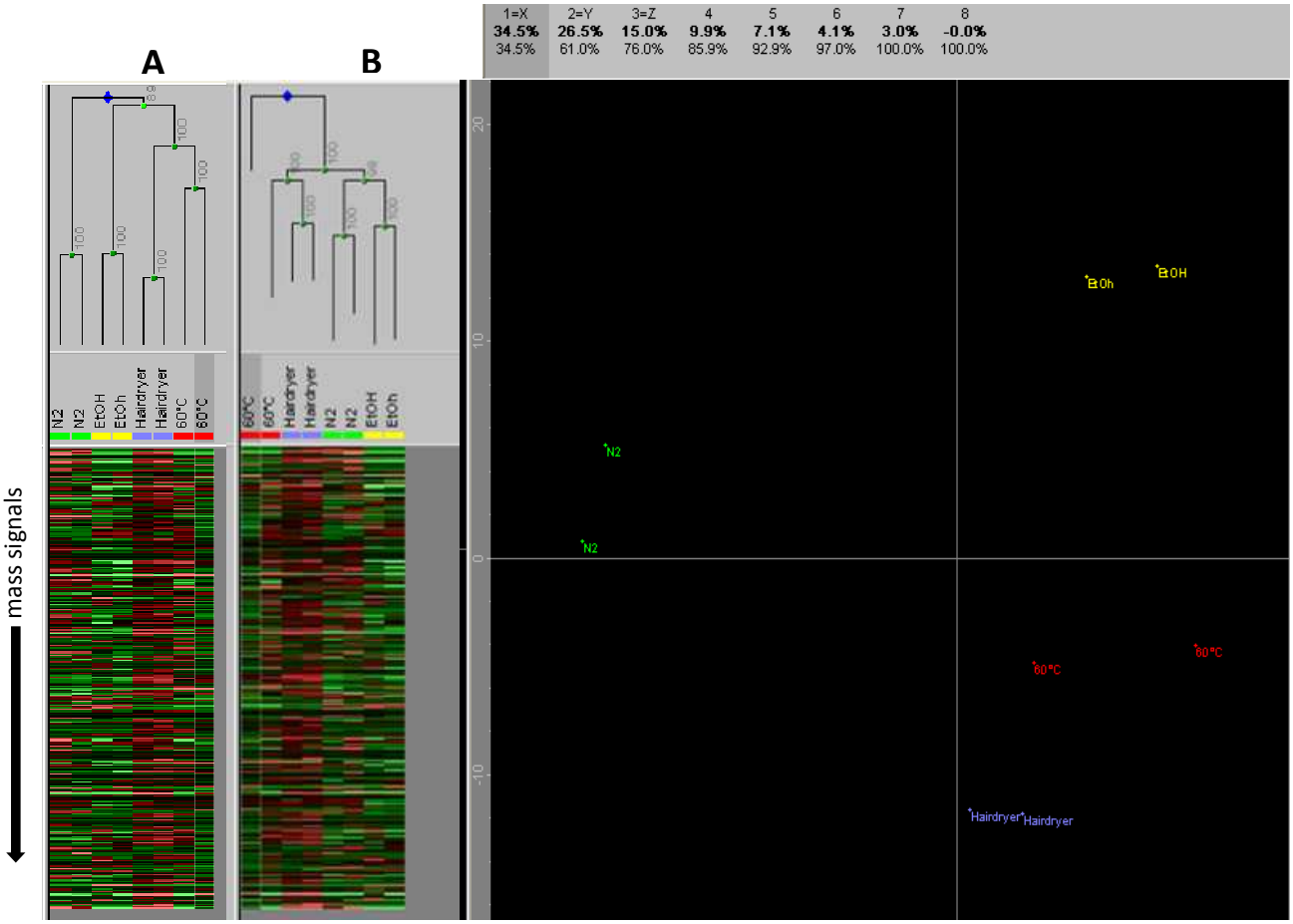


Figure A1 Genemaths output of the processed dataset of the *R. yaundensis* LC-MS runs. **A** and **B** show a clustering of the samples with Pearson Correlation clustering/UPGMA tree and Jaccard distance/Neighbor Joining tree respectively, Values are Bootstrap values. Mass signals are arranged by Retention Time, colors represent relative abundance (red high, green low). On the right is a Principal Component analysis based on the same dataset, with the components at the top, X and Y together covering 61% of variation between samples. Colors of the samples in the PCA correspond to the colors below sample names in A and B.