Phylogenetic Studies of *Catocala* by Amplification and Sequence Analysis of the Cytochrome Oxidase Subunit I Gene

by

John Alan Troutman

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in the

Department of Biology

Program

YOUNGSTOWN STATE UNIVERSITY

November, 1997

Phylogenetic Studies of *Catocala* by Amplification and Sequence Analysis of the Cytochrome Oxidase Subunit I Gene

John Alan Troutman

I hereby release this thesis to the public. I understand this thesis will be housed at the Circulation Desk of the University library and will be available for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

John A. Troutman 11-14-97
Student Date

Approvals:

David Asch 11/14/97
Thesis Advisor Date

David B. MacLean Nov. 14, 1997
Committee Member Date

Gary R. Walker Nov 14, 1997
Committee Member Date

Peter J. Kaczynski 11/17/97
Dean of Graduate Studies Date

ABSTRACT

The insect genus *Catocala* Schrank pose many interesting evolutionary questions, particularly between chromatic and achromatic species. A comprehensive phylogenetic assessment of the *Catocala*, may help elucidate these relationships and offer hypotheses which explain the enormous amount of diversity that characteristically describes its more than 200 members.

Using a pair of polymerase chain reaction (PCR) primers, we have amplified a 500 base pair (bp) region within the mitochondrial cytochrome oxidase subunit I (CO-I) gene from five *Catocala* species and *Euparthenos nubilis*. Amplified products were not produced, however, using DNA isolated from *C. mira*. These genera belong to the same subfamily Catocalinae. *Euparthenos* may be an appropriate outgroup for cladistic analysis.

DNA sequence data, obtained from the amplified CO-I gene regions of *C. piatrix*, *C. grynea*, and *C. sordida*, were submitted for comparisons over the world wide web using a BLAST search. The results of the search showed varying amounts of homology to other CO-I gene regions from a wide variety of organisms. DNA sequence comparisons of *C. piatrix* and *C. grynea* revealed 6 out of 125 nucleotide sites as being phylogenetically informative. The A+T content was also shown to be comparable for these species: *C. grynea* (77%) and *C. piatrix* (75%). Amino acid sequence comparisons from the same region showed replacement at one codon from alanine (*C. piatrix*) to arginine (*C. grynea*). These results, although preliminary, suggest that this region may be rich in phylogenetic signal. DNA sequence data from each species in its entirety, along with other species of *Catocala*, may facilitate an accurate phylogenetic estimation of its members.

ACKNOWLEDGEMENTS

I would like to extend my appreciation to Dr. David K. Asch, who has not only given me the opportunity to work in his lab, but who has also provided an enormous amount of insight and support for this project and throughout my graduate career. His expertise in the field of molecular biology has truly been inspirational.

I would also like to thank Dr. David MacLean. He has been invaluable in providing the materials necessary to obtain moth specimens from the field and in making available literature to ensure proper characterization of the insects. Thank you for giving me the opportunity to extend my appreciation of science beyond the laboratory.

To Dr. Gary Walker, I am especially appreciative of the time he has spent reviewing my thesis. He has also played a major part in the images that are found within this document. I have always valued his input, suggestions, and encouragement.

A word of thanks to John Peacock, who has literally gone out of his way to provide moth specimens. His efforts have ensured this project to never fall short of *Catocala*!

I would like to dedicate this work to my family, friends, and Stacy Shiley, who have helped make my education a reality. Without their support, love, and encouragement, this work would be meaningless.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

INTRODUCTION

I. *Insect Evolution*

More than 300 million years ago, in the mid-Paleozoic era, fossil records indicate that insects were a major terrestrial life form (Evans, 1984; New, 1996). By the upper Carboniferous period, insects were already abundant, diverse, and fully winged (Evans, 1984). By the end of that era, approximately 230 million years ago, 8 of the 21 insect orders which had appeared were already extinct (Wootton, 1981). Today, insects total more than a million named species (Evans, 1984; New, 1996) and comprise the largest class of organisms (Ross, 1982). Worldwide estimates of 5-10 million insect species have been proposed by several entomologists, and some recent estimates, based on collections in tropical rainforests, range as high as 80 million species (New, 1996). In the United States, insect biomass has been estimated to average about 45 kilograms per hectare, in comparison to the weight of humanity, estimated at about 1.6 kilograms per hectare (Evans, 1984).

II. *Characterization of the Arthropods*

Arthropoda, the largest animal phylum, is characterized by a segmented body, jointed appendages, and an exoskeleton. It includes spiders, mites, centipedes, crabs, woodlice, and insects, its largest class (New, 1996). The Lepidoptera (butterflies and moths), with approximately 112,000 species, constitutes the second largest order of insects (Arnett, 1993). The body, wings, and other appendages are covered with scales that are often brilliant in color and arranged in elaborate patterns (Ross, 1982), making them the most popular among amateur collectors (Arnett, 1993). The Noctuoidea is estimated to be the second largest moth superfamily of Lepidoptera, with over 50,000 described species (Common, 1990) and an estimated total of 80,000 species (Miller, 1991). The Noctuoidea is comprised of many well-known families, including the Noctuidae, its largest family

(Covell, 1984; Ross, 1982).

With approximately 20,000 species worldwide and about 2,900 North American species (Arnett, 1993; Covell, 1984), the family Noctuidae (Owlet or Noctuid moths) includes many species that are serious pests of forest trees and crops, resulting in millions of dollars in crop losses and pesticide expenses annually (Ross, 1982). A few species, however, may be considered beneficial because they feed on scale insects or other caterpillars (Covell, 1984). The adults of nocturnal species are mostly grey to brown. Many diurnal species, however, are brightly colored with complex patterns of lines and spots (Covell, 1984). The hindwing patterns of most noctuid species are often simpler than those of the forewing, but in species of *Catocala* (commonly referred to as underwing moths), the hindwings often strongly contrast (Covell, 1984).

III. *Characterization of Catocala Schrank*

The noctuid genus *Catocala* Schrank contains more than 200 species worldwide, most of which are North Temperate (Holarctic) (Sargent, 1976) and equally divided between the Old and New Worlds (Mitter and Silverfine, 1988). All Lepidoptera, including *Catocala* species, undergo complete metamorphosis (i.e. holometabolism) (Ross, 1982; Sargent, 1976). Holometabolic insects pass through four distinct life-cycle stages: egg, larva, pupa, and adult.

The eggs of *Catocala* species are deposited on tree trunks by adult female moths during the summer months; and due to cryptic coloration, are able to evade detection by potential predators (Sargent, 1976). Immune to freezing temperatures and dry conditions of winter, the eggs lie dormant until the buds on their hostplants open in the spring (Sargent, 1976).

Largely monophagous, the larvae (caterpillars) feed on representatives of the following plant families: Juglandaceae (Hickories and Walnuts), Fagaceae (Oaks), Salicaceae (Willows and Poplars), Rosaceae (Apples, Cherries, etc.), or Ericaceae

(Blueberries) (Sargent, 1976). In common with other arthropods whose growth is restricted by an exoskeleton, *Catocala* larvae undergo a series of molts, accompanied by an increase in size and varying degrees of change in appearance and behavior (Sargent, 1976; New, 1996). Typically, *Catocala* larvae undergo five larval instars (the period between each molt), and characteristically feed on different locations of the hostplant during this time, thought to enhance crypsis (Sargent, 1976).

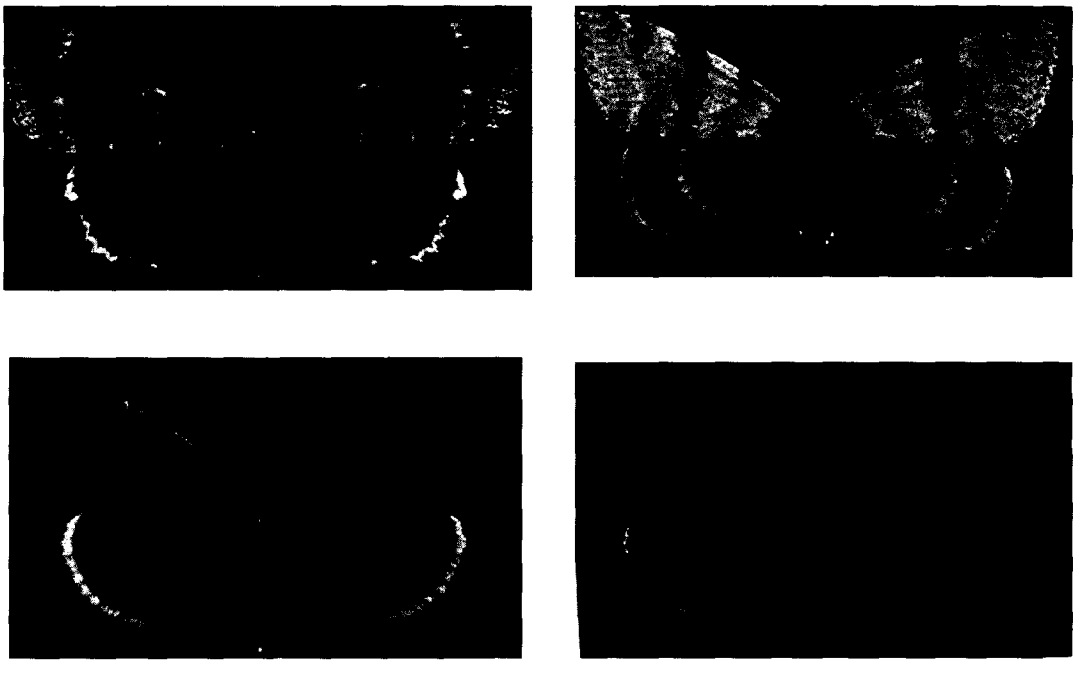
Similar to other leaf-feeding larvae (Ross, 1982), the first instar larva is typically translucent in many *Catocala* species, showing the green color of the leaf on which it feeds (Sargent, 1976). As the larva continues to grow, it becomes darker in coloration and is often found resting on the petiole or on small twigs with which it resembles (Sargent, 1976). The mature larva pupates within a silken cocoon covered by leaves of the hostplant on the ground among leaf litter. The pupal stage lasts from three to four weeks, during which time the wings and other adult structures develop internally (New, 1996; Sargent, 1976).

IV. *Forewing and Hindwing Variability*

Adult *Catocala* moths are characterized by bark-like cryptic forewings and hindwings which are often contrastingly colored (figure 1). Chromatic species are characterized by banded hindwing patterns with various shades of yellow, orange, pink, red, or white (Sargent, 1981). In contrast, the achromatic or melanic species display no trace of color or banding, having only black hindwings, or in some species, fringed with white (figure 1; bottom right) (Sargent, 1981). Sargent (1976) provides an excellent discussion regarding the functional significance of both hindwing and forewing patterns in underwing moths, which are considered to be palatable and preyed upon heavily, primarily by birds.

Crypsis is the resemblance of the color of an animal with its background; and is hypothesized to be a defense mechanism against visually oriented predators (Edmunds,

Figure 1) Adult *Catocala* moths showing bark-like cryptic forewings and hindwing variability. Top left: *C. marmorata*; Top right: *C. relicta*; Bottom left: *C. concumbens*; Bottom right: *C. maestosa*



1990). Indeed, crypsis has evolved in many insect taxa as an important means of survival (Ross, 1982). The behavior of underwing moths to select appropriate resting places helps to maximize their crypsis (Sargent, 1976). Experiments have suggested that these moths posses an innate preference for particular background reflectances (Sargent, 1968), as well as an ability to orient themselves on a substrate so as to align their markings with surface artifacts (Sargent, 1969). *C. relicta*, a species with greyish-white forewings, is commonly found resting on light colored trees, such as white birch (figure 2; top). The resting behavior of *C. retecta*, and its preference for darker backgrounds such as hickory, is shown in figure 2 (bottom). In spite of their cryptic appearance, the forewings of many *Catocala* species are polymorphic.

Polymorphism is the occurrence of two or more distinctly different forms of a species occurring sympatrically (Sargent, 1976). Polymorphism in many *Catocala* forewings (figure 3), may enable cryptic moths to escape detection from predators which hunt by means of a search-image (Edmunds, 1990). Each predator would therefore have to learn a specific search image for each distinct morph in the population (Edmunds, 1990). Croze (1970), using wild crows as predators, found that trimorphic prey suffered less predation than monomorphic prey at the same density.

Melanism, the darkening of the forewings caused by an increase in melanin production and deposition, has also been investigated by Sargent (1976) in regard to underwing moths. The population of melanic forms of some *Catocala* species appear to have increased in frequency, particularly in industrialized areas where pollutants darken tree trunks and other resting substrates such as stone buildings and bridges (Sargent, 1976). Darker moths are undoubtedly selected for in such environments. Figure 4 illustrates *C. relicta* "phrynia" taken from two different localities: Leverett, Mass. (top) and Pittsburgh, PA (bottom) (Sargent, 1976). Melanic forms of *C. relicta* have been taken in Pittsburgh for decades (Sargent, 1976), once a leading steel manufacturing city. Evidence such as this

Figure 2) Resting behavior. Top: *C. relicta* at rest on white birch in characteristic 'head up' position. Bottom: *C. retecta* resting on hickory in characteristic 'head down' position. The underwings tend to align their markings with surface artifacts of their backgrounds.

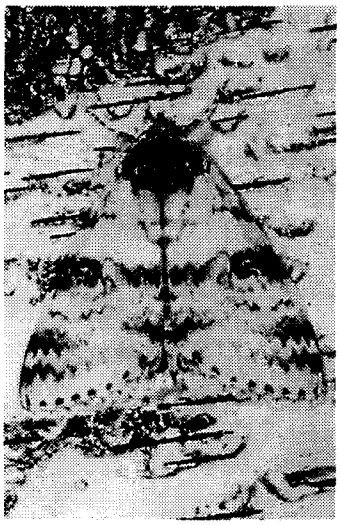


Figure 3) Forewing polymorphism of *C. micronympha*. Top left: typical; Top right: unnamed; Bottom left: "hero"; Bottom right: "grisela"

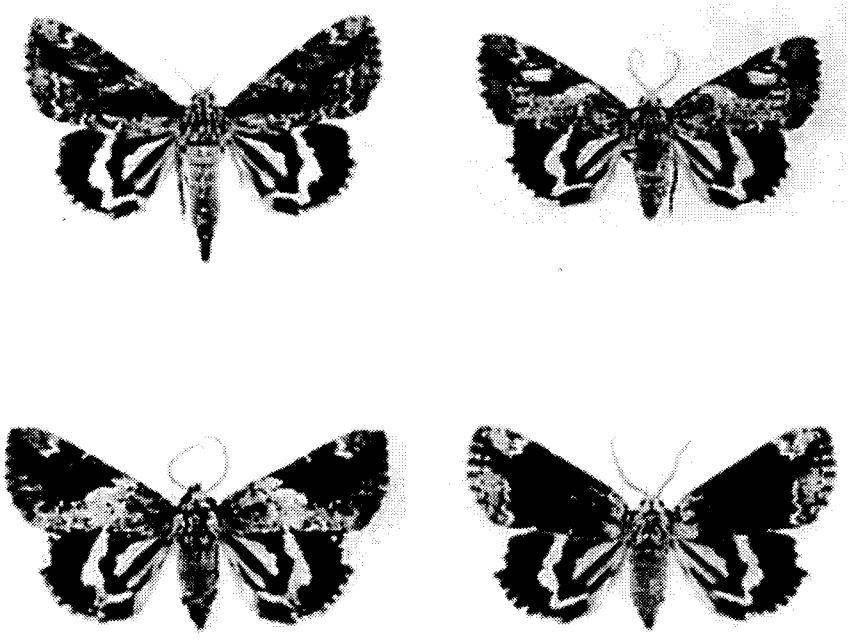
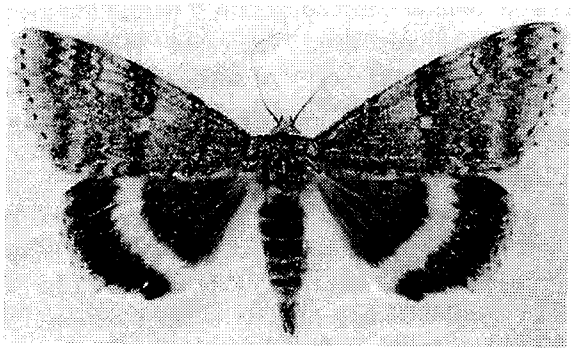
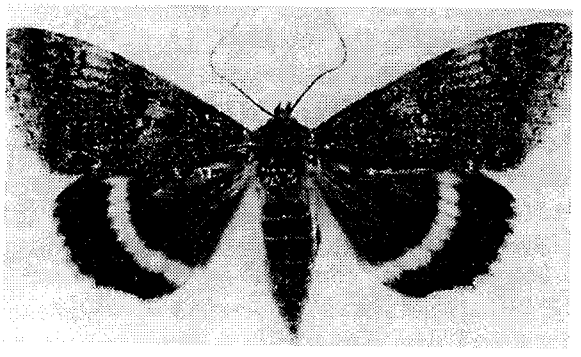


Figure 4) Melanic forms of *C. relicta* "phrynia." Top: Taken from Leverett, Massachusetts; Bottom: Taken from Pittsburgh, Pennsylvania.





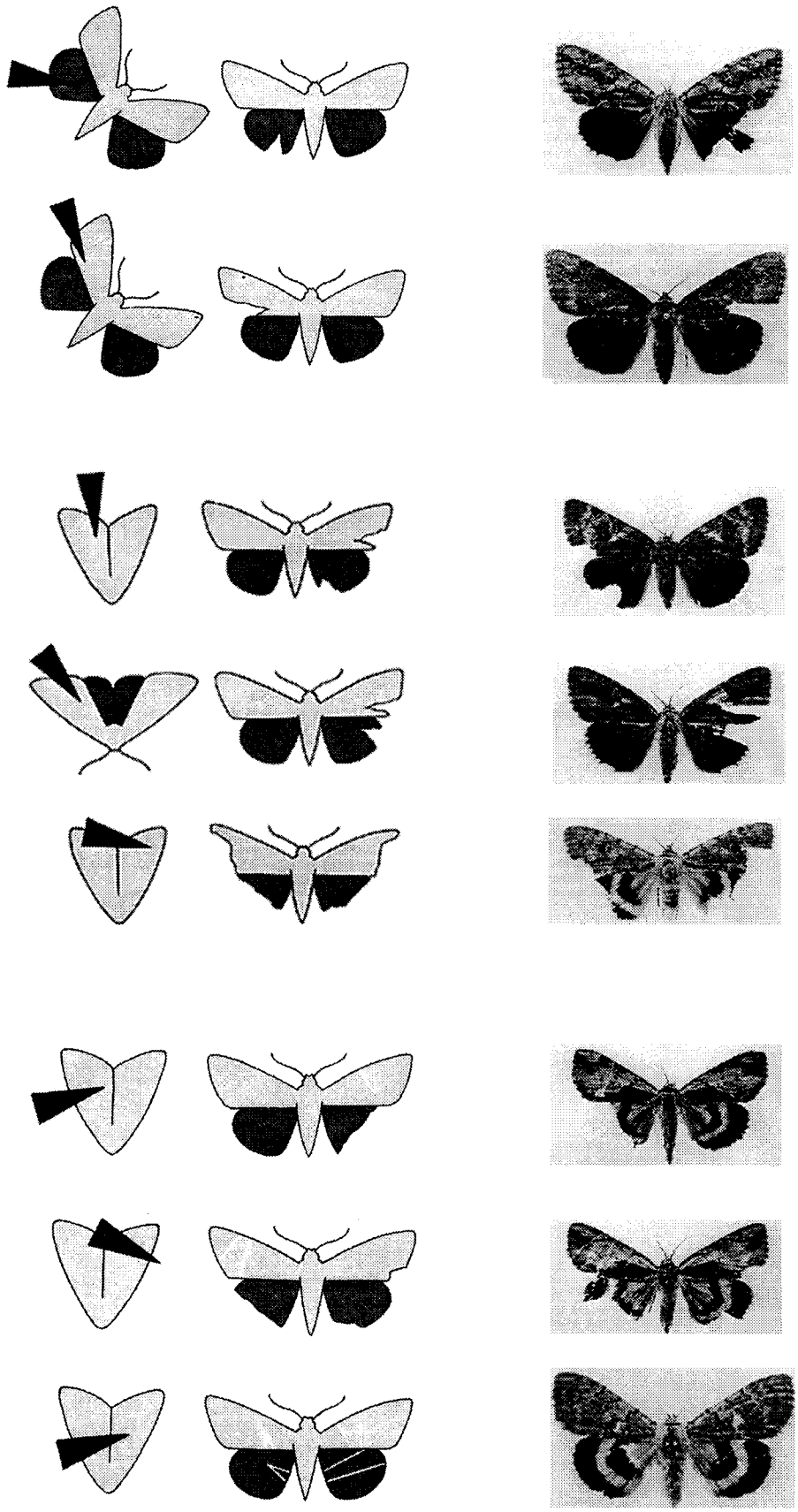
lends support to the notion that intraspecific variability as well as adaptation to changing environments, may be critical for the survival of *C. relicta* and similar species. These findings, according to Sargent (1976), corroborate much of the pioneering work of Kettlewell (1955, 1956) and his associates who studied melanism in the Salt and Pepper moth, *Biston betularia.*

An alternative explanation to the occurrence of melanic forms in areas free of high pollutant levels, is the idea of plant successionism. Sargent has proposed that changes in plant communities from "light-barked" trees (e.g. white birch) to "dark-bark" trees (e.g. hickories) would favor melanic species of *Catocala.* Therefore, an increase in frequency of melanism may be due to successive change in plant communities favoring "dark-bark" species (pers comm).

The relative success of these mechanisms for evading detection and subsequent predation includes a second line of defense for these moths, the elicitation of a startle-response in avian predators. A startle-response is a widely recognized, yet poorly understood phenomenon (Sargent, 1990), initiated only after a primary defense mechanism (crypsis) has failed (Edmunds, 1974). At rest, the cryptic forewings lie folded over the hindwings, concealing their bold patterns (Sargent, 1990). When disturbed, underwing moths reveal their hindwings which are thought to elicit a startle-response in their avian predators and thereby reducing the efficacy of the attack, or directing it to a less vulnerable area (MacLean, 1984; Sargent, 1976).

Evidence to support a startle-response due to *Catocala* hindwings was provided by Sargent (1973), who studied beak-damage patterns found on specimens caught in the field and specimens which successfully escaped attack after release in an aviary (Sargent, 1976). These observations led to the classification of three major types of beak-damage patterns (figure 5) (Sargent, 1976) as follows:

Figure 5) Beak-damage patterns showing site of attack and the resulting damage as described in text.



Type I: Characteristic damage: unilateral; tear from one wing only

a. hindwing tear
b. forewing tear

Type II: Characteristic damage: corresponding tears from ipsilateral forewing and hindwing

a. forewing and hindwing tears overlapping when wings are fully closed; unilateral
b. forewing and hindwing tears overlapping when wings are partially closed; unilateral
c. forewing and hindwing tears overlapping when both wings are fully closed; bilateral

Type III: Characteristic damage: crisp beak imprint on one forewing; imprint or tear from ipsilateral hindwing

a. apex of beak-imprint directed toward (but not across) inner margin of forewing; unilateral
b. apex of beak-imprint directed toward (and usually across) costal margin of forewing; unilateral or bilateral
c. apex of beak-imprint directed toward (and across) inner margin of forewing; bilateral

In contrast to polymorphic forewings exhibited by many *Catocala* species, the hindwing patterns remain largely invariant within species, although they are not necessarily indicative of taxonomic affinities (MacLean, 1984). Hindwing diversity among sympatric species, particularly between chromatic and achromatic forms, appears to also serve as anti-predator devices by introducing the element of anomaly (the unexpected) into the predator-prey system (Sargent, 1981).

Habituation, considered to be the simplest form of learning, involves the loss of sensitivity to recurrent environmental stimuli (Ross, 1982). On the basis of past experiences of the avian predator, expectations of hindwing patterns to be encountered in the future are built up (Sargent, 1976). After several successful attacks of one hindwing type, an anomalous stimulus (a different hindwing type) interferes with the habituation process by causing a startle-response allowing the moth to escape (Sargent, 1976). Evidence of this apparent dis-habituation is seen between chromatic and achromatic

hindwing types with type III damage patterns (figure 6), namely beak-imprint markings on the forewings and an ipsilateral tear on the hindwing. Achromatic species comprise approximately 20% of *Catocala* populations in most areas. When encountered, they represent an anomalous stimulus to avian predators (Sargent, 1976).

Ingalls (1993) conducted experiments designed to elucidate whether banding patterns on novel stimuli enhanced the startle reaction (experiment I) and whether the rate of habituation is affected by the number of different stimuli types encountered (experiment II). In both experiments, blue jays were required to manipulate a 24-well feeding apparatus in order to obtain a food reward (mealworm). Each well consisted of a flap which had to be pushed open by the jay to reveal a second flap. The second flap, painted in various colors (banded or unbanded), had to be removed to obtain the reward.

Results from experiment I showed that banded stimuli accounted for the increased latency in obtaining the reward, especially for red-yellow colored stimuli which is representative of three *Catocala* hindwing colors (refer to figure 1) (Ingalls, 1993). Furthermore, experiment II demonstrated a positive correlation between stimulus diversity and the number of trials needed to habituate the jays to a set of startling stimuli (Ingalls, 1993). Admittedly however, insufficient data in this experiment precluded any conclusions regarding hindwing diversity among sympatric *Catocala* species from being made (Ingalls, 1993), as suggested by other evidence discussed below.

Variation among the underwings is confounded further when one considers the amount of similarity observed in forewing patterns between some chromatic and achromatic species. Sargent (1981) discusses this issue and includes several plates (which have been reproduced here) to illustrate these similarities.

Figure 7 are examples of chromatic and achromatic patterns seen in closely related species which share the same hostplants (Hickory), exhibit similar seasonal occurrences and behavior. Others, on the other hand, feed on different foodplants and do not share other common attributes and are thought to be distantly related (figure 8). These examples

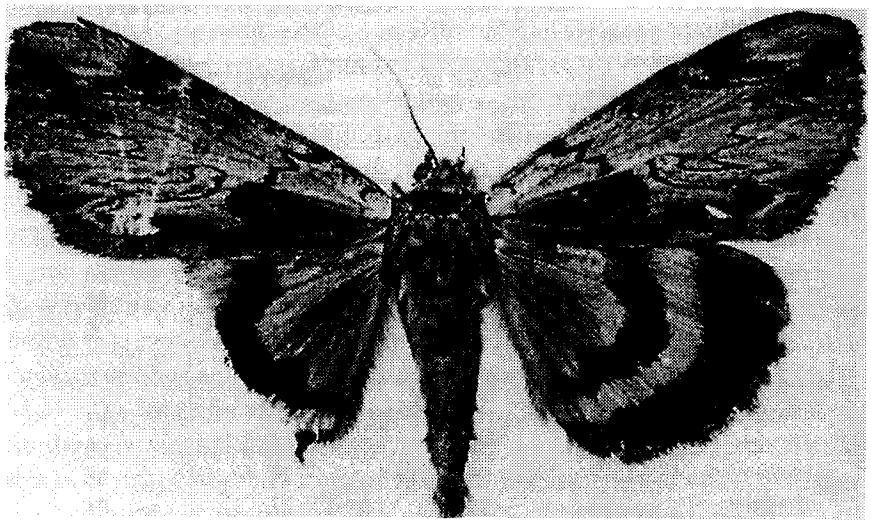


Figure 6) Type III damage pattern; evidence of dis-habituation.

Figure 7) Parallel forewing morphs among closely related species.

Top: *C. palaeogama* "denussa"; Bottom: *C. lacrymosa* "zelica"

Figure 8) Parallel forewing morphs among distantly related species.

Top: *C. marmorata*; Bottom: *C. maestosa*

further support the idea that achromatic hindwings introduce anomaly into the predator-prey system which act as important startle devices, especially those in sympatric localities with chromatic forms (Sargent, 1981).

The amount of diversity that is displayed by *Catocala* species represents a long evolutionary history in which changing environments and many accounts with visually acute predators produce selective pressures that beg for diversity among its members. To a large extent, however, the evolutionary relationships between the members of this large genus remain enigmatic. Many interesting questions regarding the mode of speciation and the functional significance of the achromatic forms have been raised (Sargent, 1981). Mechanisms which maintain the genetic integrity of these assemblages, particularly those that are found sympatrically, as well as their ability to compete for resources have yet to be elucidated. The *Catocala* are an interesting group for such evolutionary considerations.

V. *Evolution and Phylogenetic Inference*

The field of systematics encompasses a broad range of disciplines whose primary objective is to study the amount and type of diversity that exists between organisms. This endeavor began centuries ago as naturalists began to detect, describe, and explain diversity in the biological world (Moritz and Hillis, 1996). Initially, the reconstruction of evolutionary histories (phylogenies) of organisms were based on few, if any, objective criteria (Moritz and Hillis, 1996). Many researchers have argued that more objective criteria should be used, as in other areas of science, for addressing taxonomic decisions (Quicke, 1993).

The emergence of new ideas in the 1960's led to the development of largely separate areas of systematic study (Moritz and Hillis, 1996; Quicke, 1993). Much debate over the relative merits of each respective approach for classification and phylogenetic inference followed in the 1970's and the 1980's (Smith, 1994). Hull (1970) recognized three distinct schools involved in this debate, namely, phenetics, cladistics, and

evolutionary systematics.

Phenetics, also known as numerical taxonomy from its inception, was largely developed and popularized by Sokal and Sneath (1973). This approach to classification involves the clustering of organisms into groups on the basis of similarity by applying various mathematical procedures to numerically encoded character state data. The taxa are then arranged into a hierarchy based on the products of the mathematical operations (Quicke, 1993). This form of classification, said to be 'unbiased' indicators of similarity between taxa (Quicke, 1993), is unable to distinguish between similarity caused by common ancestry or similarity due to convergent evolution (Scott-Ram, 1990). Thus, phenetic methods have been criticized for ignoring phylogenetic relationships among organisms (Gingerich, 1979), a notion that Hennig (1966) considered vital.

Different works were being published at about the same time as the development of numerical phenetic methods by the German systematic entomologist Willi Hennig (Quicke, 1993). Although not fully realized until translated into English, they were ultimately to lead to a revolution in taxonomic thinking (Hull, 1979). Hennig's methods and their subsequent derivations are referred to as cladistics or phylogenetic systematics in which the the pathways of evolution are defined (Li and Graur, 1991). It is based on the knowledge of clades, which are part of a phylogeny (evolutionary history). Therefore, the phylogeny must be known, or presumed before one can attempt cladistic analysis (Gingerich, 1979). The basic units of cladistic analysis are the species which are characterized by at least one unique derived character (Hennig, 1966). Unlike phenetic methods which does not discriminate between the types of characters used in the analysis, cladists include only characters which infer common ancestry, i.e. homologous characters (Williams, 1993).

Cladistics requires that for as many homologous characters as possible, a decision is made as to which of its states are the primitive condition (plesiomorphous) and which are the advanced or derived condition (apomorphous) (Quicke, 1993). Plesiomorphous and apomorphous characters may be shared by two or more groups however. Homologous

plesiomorphous character states shared by two or more taxa are referred to as a symplesiomorphy. If two or more taxa share the same apomorphous state for a character, it is said to be a synapomorphy for the taxa possessing it. Synapomorphies are shared-derived charactersitics, meaning they are common to both taxa, and the character state has changed compared to the most recent common ancestor. Therefore, characters which are shared-derived are unique to the descendants of an ancestor and can serve to delineate taxonomic relationships.

The third method of classification is called evolutionary systematics. This approach attempts to maximize evolutionary theory in order to prepare classifications which are most useful (Ashlock, 1979). It is both phenetic in that it addresses overall similarity and cladistic because it emphasizes the importance of phylogeny in classifying organisms (Ashlock, 1979; McNeill, 1979). All evolutionary systematists employ a principle rule regarding its methodology which is finding the correlation between a classification and a phylogeny (Scott-Ram, 1990).

VI. *Characters Used for Phylogenetic Estimates*

The raw data used in all phylogenetic analyses, regardless of the method, are taxa versus character state matrices (Smith, 1994; Quicke, 1993). Essentially, characters are any particular attribute of a specimen that can take two or more forms, called character states (Platnick, 1979). Some of the most common types of data used as characters for inferring phylogenies are morphological, physiological, biochemical, and macromolecular sequences (Quicke, 1993).

Data can be collected in two fundamentally different forms: 1) as character data (discrete characters), or 2) as distance data (Crozier, 1993). Discrete data provides data about an individual species, measured as a series of discrete characters (states) (Moritz and Hillis, 1996). Character data are usually transformed into distance values which are quantitative comparisons of taxa, and therefore, not a property of any one taxon (i.e. each

measure describing a pairwise relationship) (Swofford et al., 1996). Unlike discrete characters which can be converted into distance values, distances can not be converted into character data (Crozier, 1993). Thus, character data allow new taxa to be added directly to the data set and analyzed (Moritz and Hillis, 1996); whereas, for distance data, the addition of new taxa is best done by determining the distances from it to each previously-studied taxon (Crozier, 1993).

Clearly the most direct evidence of past life on earth is the fossil record (Forey, 1992), referred to by Scott-Ram (1990) as the history of life. Haeckel (1866) first coined phylogeny as the history of the palaeontological development of species through time. Due to the incomplete nature of the fossil records of any organism, questions regarding the use of this evidence as characters in phylogenetic analyses have not been resolved (Hillis, 1987). According to Quicke (1993), phylogenetic evidence that can be used from fossils is two-fold.

First, fossils can provide an estimate regarding the minimum age of the origin of particular taxa and their character states. Phylogenetic trees have been described by Smith (1994) as cladograms calibrated against the fossil record. According to Forey (1992), cladograms are important for reconstructing phylogenetic trees because they are likely to be the only ancestors of later fossils and of modern species. Several well-preserved fossils of a species of tarantula spider (*Rosamygale grauvolgeli*) have been discovered which date the ancestry of this group, the Mygalomorpha, at 240 million years old (Selden and Gall, 1992).

Secondly, fossil evidence can be used for additional taxa and their character states can be included to estimate phylogenies. The problem arises when fossil taxa are included with incomplete data sets, particularly those with soft tissues, causing the resolution of the consensus cladogram to be reduced significantly (Quicke, 1993). Donoghue et. al. (1989) examined this problem by comparing cladograms of amniote vertebrates and seed plants, both with and without fossil evidence. They concluded that the fossil taxa should not be

excluded because this evidence was shown to have a profound affect on the topology (branching pattern) of the cladogram. Others, such as Patterson (1981), oppose the use of this evidence for phylogenetic anaylses, particularly when a more resolved cladogram is preferred over a less than accurate estimation of phylogeny. Novacek (1992) pointed out that there is not a significant amount of carefully analyzed data regarding fossil evidence and more work in this field is urgently needed.

Palaeontological evidence has been treated primarily by morphologists (Hillis, 1987). A few techniques for obtaining molecular information from well-preserved fossil specimens have met with some success however. DNA has been cloned from Egyptian mummies (Pääbo, 1985) and from human brain tissue found buried underwater for 8000 years (Doran et. al., 1986), offering an exiciting new direction for phylogenetic studies as methods improve.

VII. *Molecular Systematics*

In recent decades, the use of molecular sequences for systematics has gained considerable attention, oweing much of its popularity to advancements made in molecular biology, particularly, the advent of the polymerase chain reaction (PCR) (Brower and DeSalle, 1994). PCR reduces the amount of time necessary to sequence a gene, facilitates sequencing of DNA that has been recovered in minute amounts, and by using universal PCR primers, virtually any gene can be sequenced (Simon et. al., 1994). Although most systematists generally recognize that no one type of systematic data is better than another for inferring phylogenies, molecular data offer many distinct advantages (Hillis, 1987).

The extent of data that can be obtained from molecular sequences is enormous (Hillis, 1987), since each base in a sequence can be assumed to evolve independently (Quicke, 1993). The number of independent characters for any organism is limited by the number of nucleotide base pairs in its DNA, ranging from about 5×10^3 for the smallest virus, to nearly 4×10^{11} for some eukaryotes (Hillis, 1987). The amount of effort to find

an equivalent number of informative morphological characters for a relatively closely related set of taxa would require an exhaustive search, likely to be impossible (Quicke, 1993).

Secondly, there are very few homologous morphological characters that can be directly compared across all extant organisms (Hillis and Huelsenbeck, 1992). In contrast, many genes with similar biochemical functions can be found in all living organisms which can be sequenced, aligned, and analysed for phylogenetic studies (Hillis and Dixon, 1991). Caution must be used however in such comparisons because some molecular data sets consist of little phylogenetic signal and considerable random noise (Fitch, 1984).

Random noise refers to high rates of change between nodes of a phylogenetic tree which, when significantly high (a probability of change of $\geq 75\%$ between nodes for DNA sequences), cause the character states to be significantly randomized with respect to the phylogenetic history (Hillis and Huelsenbeck, 1992). Analyses which use such sequences contribute to the notion that morphological data are superior in comparison to molecular data (Hillis, 1987). Hillis and Huelsenbeck (1992) discuss this issue and offer a method which distinguishes phylogenetic signal from random noise within a data set.

The establishment of gene libraries offers yet another advantage that is unique to molecular data. They can contain an inexhaustible store of genetic information that can be shared among all workers, unlike museum collections which are rarely loaned (Hillis, 1987). In a similar theme, nucleic acid sequence data are compiled continuously in several data bases (Hillis et. al., 1996). Perhaps the best known and most widely used is GenBank. Release 102.0 (August, 1997) lists 1,053,474,516 bases from 1,610,484 loci of available DNA sequence information. This wealth of information makes available sequence data which can be 'downloaded' and included in a phylogenetic analysis.

Despite all of this, not all aspects concerning molecular data are advantageous. Cost can clearly be a major obstacle, particularly the initial lab set-up, costing tens of thousands to hundreds of thousands of dollars to establish and maintain (Hillis, 1987).

Secondly, most systematists are inept to molecular methods and techniques for isolating and sequencing genes or gene regions (Crozier, 1993). As sequence data among taxa become increasingly available in data bases, however, this may be less of a problem as long as the appropriate sequences are available for each taxon in the analysis. Thirdly, molecular methods often require the destructive sampling of the organism (Baverstock and Moritz, 1996). Sufficient precautionary measures must be taken in order to maximize the amount of data that is obtained and used in the analysis, particularly in the planning stages. And finally, unlike other characters, it is impossible to distinguish homology from analogy for the individual positions in a molecular sequence (Quicke, 1993).

In a general sense, both homology and analogy are terms which describe similarity between taxa. The most common meaning of homology is common ancestry (Williams, 1993). In the context of nucleic acid sequences, an exact base for base match of the same gene between two taxa is considered 100% homologous (Williams, 1993). Analogy, on the other hand, is similarity that is not indicative of common ancestry but are similarities caused by reversals, convergences, or parallelisms. As an example, the wings of a bird and bat are analogous characters which have resulted from convergent evolution (Scotland, 1993). Evolutionary changes such as these which lead to misleading conclusions about phylogeny are termed homoplasies (Quicke, 1993). Because there are only four nucleotide character states: G, A, T, and C, base substitutions at a single site may occur more than once (mutilple hits) (Simon et. al., 1994). Therefore, character states which are shared between two taxa may actually be the result of chance convergence rather than common ancestry. The strength of the phylogenetic signal must then be sufficiently large as to reduce the effect of homoplastic noise (Wake, 1996).

Shared ancestral polymorphisms, as discussed by Simon et. al. (1994), represents yet another potential pitfall for sequence data. The problem arises when variants fixed in different species originated from polymorphic variation in the common ancestor. The changes are therefore older than the species that carry them and have become fixed or lost

randomly among species. In such cases, the phylogenetic relationships of the genes (gene tree) are unlikely to reflect the phylogenetic relationships of the species (species tree). Especially when the time between the speciation event is small (Li and Graur, 1991).

Often the first and most critical step in molecular systematic studies is selecting the appropriate sequence(s) (Baverstock and Moritz, 1996; Hillis and Dixon, 1991). This decision is commonly based on the rate of evolution and mode of inheritance of the sequence(s) (Dowling et. al., 1996). A second consideration is the time frame of divergence within the study group (Baverstock and Moritz, 1996). Slowly evolving sequences may provide useful characters for studying distantly related taxa, whereas those with rapid evolutionary rates, yet moderate to low intra-species polymorphism, are most appropriate for studying recently diverged (closely-related) taxa (Dowling et. al., 1996). Therefore, genes are favored whose level of variability will minimize the effect of multiple hits while maximizing the number of nonhomoplasious shared character states (Simon et. al., 1994). Hillis and Dixon (1991) suggested selecting DNA regions that are greater than 70% but less than 100% similar among taxa for phylogenetic studies.

VIII. *Sequence Evolution*

The nuclear ribosomal RNA (rRNA) gene cluster is among the slowest evolving sequences found among living taxa (Brower and DeSalle, 1994). Due to its functional importance, it has been highly conserved across all genomes and used for phylogenetic analyses of the deepest branches of *the tree of life* (Simon et. al., 1994). All eukaryotic nuclear ribosomes consist of 18S, 5.8S, and 28S subunits, separated by two transcribed spacer regions (ITS1 and ITS2) (Brower and DeSalle, 1994). The rate of evolution along the length of the molecules vary and can be used at differet levels of divergence for phylogenetic estimations (Simon et. al., 1994).

The small rRNA (18S) subunit gene tends to be most conserved across taxonomic boundaries (Polumbi et. al., 1996) and is useful for reconstructing phylogenetic events

from the Precambrian (Hillis and Dixon, 1991). The large rRNA (28S) subunit gene has been used primarily to examine events through the Paleozoic and Mesozoic (Hillis and Dixon, 1991). The 5.8S rRNA gene is similar to the small subunit (18S) in its phylogenetic usefulness (Hillis and Dixon, 1991). However, the shortness of the sequence greatly reduces its effectiveness for inferring phylogenies (Halanych, 1991). Amplification of the two internal transcribed spacer regions (ITS1 and ITS2) is facilitated by conserved flanking regions of the 18S, 5.8S, and the 28S genes (Hillis and Dixon, 1991). These regions can be used among closely related taxa that have diverged within the last 50 million years (Hillis and Dixon, 1991).

Single-copy nuclear protein coding regions have been less studied for phylogenetic studies (Brower and DeSalle, 1994). Friedlander et. al. (1992) described 14 genes which are thought to be phylogenetically useful. However, single-copy amplifications are more difficult because there are fewer primer binding targets and eukaryotic genes are often interrupted by one or more introns (Polumbi et. al., 1996). Brower and DeSalle (1994) noted that a single-copy gene in a particular species may occur as multiple-copies in other species such as wingless.

In higher animals the circular, double-stranded mitochondrial DNA (mtDNA) molecule consists of approximately 16,000 base pairs (bp) (Quicke, 1993) containing only unique (non-repetitive) sequences (Li and Graur, 1991). The entire nucleotide sequence is known for man, mouse, cow, and Drosophila (Quicke, 1993). All of the molecules contain two ribosomal RNA (rRNA) genes, 22 transfer RNA (tRNA) genes, 13 protein coding genes, and a control region for replication (Wolstenholme and Clary, 1985). With the exception of tRNA genes, gene order appears to be highly conserved within vertebrates and within insects but not among animal phyla (Wolstenholme, 1992). MtDNA has become popular for studies of animal populations because, unlike nuclear DNA, its rate of base substitution is relatively high, its inheritance is uniparental and non-recombining, and it is relatively easy to isolate and analyze (Dowling et. al., 1996; Simon et. al., 1994).

Animal mtDNA has been estimated to evolve 5-10 times that of single-copy nuclear DNA sequences (Quicke, 1993). Generally, substitution rates have been attributed to the idea that greater constraints are placed upon sequences that are under strong selective pressures (Gillespie, 1991). Simon et. al. (1994) offer a detailed discussion on rates of substitution among mtDNA.

Mitochondrial rRNA genes are the most conserved. Although, in comparison to their nuclear counterparts, they have evolved much more rapidly (Simon et. al., 1994). At the other end of the spectrum are the protein-coding genes. These genes appear to have evolved the most which may be attributed to the degenerate nature of the genetic code. Many base substitutions within the codon may not result in an amino acid change (silent mutation). Therefore, the rate of substitution may be high in some sequences without altering the structure of the protein. Brown and Simpson (1982) investigated the cytochrome oxidase subunit II (CO-II) gene of two rat species (*Rattus norvegicus* and *R. rattus*) and found that out of 54 total base changes (8%), only 6% resulted in amino acid changes (96% were silent).

Wolstenholme and Clary (1985) compared the genes for six mtDNA protein sequences of *D. yakuba* and *D. melanogaster*. Nucleotide and amino acid comparisons between corresponding protein genes showed an overall sequence divergence for all of the sequences to be 7.2% and 2.7%, respectively. Nucleotide divergence ranged from 6.3% (URFA6L) to 7.6% (URF2). These regions, however, represented the highest amino acid divergence values, 3.8% (URFA6L) and 6.3% (URF2). Sequence divergence for the CO-I gene was 7.1%, with an amino acid sequence divergence of only 1%, the lowest for all amino acid sequences compared. Sequence comparisons for CO-II and CO-III showed nucleotide divergences to be 6.4% and 7.4%, respectively. Amino acid divergence for these genes were 1.8% (CO-II) and 1.9% (CO-III).

Simon et. al. (1994) also offer percentages of amino acid similarity at various divergence levels for various mitochondrial protein coding genes. The amino acid

sequences from the cytochrome oxidase I, II, and III genes appear to be highly conserved for closely related taxa such as Drosophila, estimated to have diverged between 3-19 Mya (million years ago). Similarity levels for these protein sequences become progressively low, however, as the time between the speciation event increases. For human-cow comparisons, estimated to have diverged more than 90 Mya, the CO-I sequence is 91% similar. Comparisons of mouse-nematode, estimated to have diverged more than 700 Mya, shows a 59% similarity for the CO-I sequence. The amino acid sequence from the NADH dehydrogenase subunit 6 (ND6) gene is 91% similar for Drosophila species and represents the lowest amount of similarity among the genes examined (13% similar for mouse-nematode comparison).

Liu and Beckenbach (1992) investigated the evolution of the CO-II gene among 10 insect orders. They found a total of 468 of the 688 nucleotide sites were variable (68%). Most of the variation occured at or near the 3' end of the gene. Highly conserved regions were shown to be involved in the electron transport system and some included the copper binding sites of the protein. Other individual residues known to be conserved in CO-II sequences from vertebrates, yeast, and maize were also conserved in all of the insects analyzed in the study. Although the CO-II gene is a highly conserved protein coding gene, they concluded that it is not useful for inferring the phylogenetic relationships among insect orders, which are thought to have diverged more than 300 million years ago (Wootton, 1981).

The purpose of this study is to estimate the phylogeny of the members of the genus *Catocala* Schrank. The underwings pose many interesting questions regarding their evolutionary history, particularly between chromatic and achromatic populations. Here, a portion of the CO-I gene has been amplified and sequenced from DNA isolated from various insect species. The CO-I gene was chosen based on its degree of amino acid similarity among closely related species. We hypothesize that this region is rich in phylogenetic signal and will facilitate the phylogeny of the underwings to be estimated.

MATERIALS AND METHODS

Materials

I. Ampicillin, sodium chloride [NaCl], potassium acetate [KOAc], 3-N-morpholino-propanesulfonic acid [MOPS], (tris[Hydroxymethyl]aminomethane) [Trizma Base], RNase A, Proteinase K, urea, isoamyl alcohol [$C_5H_{12}O$], polyoxyethylene-sorbitan monolaurate [Tween 20], N,N,N',N'-tetramethylehtylene-diamine [TEMED], N-Lauroyl sarcosine, polyehtylene glycol 8000 [PEG8000], sigmacote, developer & replenisher, fixer &replenisher, and lithium chloride [LiCl] were purchased from Sigma Chemical Co., St. Louis, MO. T4 DNA ligase (1U/µl), restriction endonucleases [*Sac*I (10 U/µl), *Pst*I (10 U/µl),*Sma*I (10 U/µl)], DIG *Taq* DNA Sequencing Kit for Standard Cycle Sequencing, blocking reagent, disodium 3-(4-methoxyspiro {1,2-dioxetane3,2'(5'chloro)tricyclo[3.3.1.1^{3,7}] decan}-4yl) phenyl phosphate [CSPD], anti-digoxigenin-AP fab fragments, bis-acrylamide, acrylamide, and positively charged nylon membranes were purchased from Boehringer-Mannhiem, GmbH, Germany. Ethylenediamineteraacetic acid • disodium salt [EDTA], isopropyl-ß-D-galactopyranoside [Xgal], sodium dodecyl sulfate [SDS], and ammonium persulfate were purchased from International Biotechnologies, Inc., New Haven, CT. *PERFECT prep* Plasmid DNA Kit and *PCR* Select-II Purification Columns were purchased from 5 Prime —> 3 Prime, Inc., Boulder, CO. Phosphoric acid [H_3PO_4] (85%), and ethidium bromide [EtBr], chloroform were purchased from FisherScientific, Fair Lawn, NJ. Bacto-trypton, bacto-agar, and yeast extract were purchased from Difco Laboratories, Detroit, MI. Hydrochloric acid [HCl], magnesium chloride [$MgCl_2$], chloroform [$CHCl_3$], and calcium chloride [$CaCl_2$] were made by Mallinckodt, Inc., Paris, KY. Agarose was made by EM Science, Gibbstown, NJ. Maleic acid and polaroid film were purchased from Eastman Kodak Co., Rochester, NY. Dextrose, phenol and acetic acid [HAOc] were made by J.T. Baker Chemical Co., Phillipsburg, NJ. Sodium hydroxide [NaOH] was purchased from VWR,

Media, PA. Isopropanol was made by Baxter Healthcare Corp., McGraw Park, IL. Ethanol [EtOH] was purchased from Aaper Alsohol and Chemical Co., Shelbyville, KY. Qiagen columns were purchased from Qiagen Inc., Chatsworth, CA. PCR primers 78X-1 and 78X-2 were purchased from DNA International Inc., Lake Oswego, OR.

Methods

II. *Strains and Media*

Recombinant plasmids were transformed into *Escherichia coli* strain JM101. *E. coli* JM101 was cultured in Luria broth [LB] (1% bact-tryptone; 0.5% yeast extract; 1% NaCl). Bluescript transformants were selected on Luria agar + ampicillan [LA_{100}] (LB; 1.5% bacto-agar; ampicillan (100 µg/ml), also containing 100 µl IPTG (200 mM) and 50 µl Xgal (2%). Transformants were picked from LB containing ampicillan (100 µg/ml) [LA_{100} broth].

Single-stranded M13 phages were infected into *E. coli* strain JM101. *E. coli* JM101 was cultured in YT broth (0.8% bacto-tryptone; 0.5% yeast extract; 0.5% NaCl). M13 transformants were cultured in 2 XYT broth (1.6% bacto-tryptone; 1% yeast extract; 0.5% NaCl) and selected on YT plates (YT broth; 1.5% bacto-agar) containing 100 µl IPTG (200 mM) and 50 µl Xgal (2%).

III. *Collection and Identification of Insect Species*

Catocala piatrix, C. sordida, C. grynea, C. subnata, and *Euparthenos nubilis* specimens were collected from three sites of Hvizdos Acres in Mahoning County, Ohio. *C. piatrix* was collected from site 1 by D. MacLean on August 2, 1995. *C. sordida* (site 2), *C. grynea* (site 3), and *C. subnata* (site 3) were collected by J. Haylett on August 5, 7, and 19, 1995, respectively. *Euparthenos nubilis* was collected from site 1 on July 26, 1995. *C. cara* and *C. neogama* were collected by J. W. Peacock on September 13 and 16, 1996, respectively, in Miami County, Ohio. After identification, the specimens were

stored at -70°C.

IV. *Isolation of Insect DNA*

The thorax and abdomen were ground with a mortar and pestle containing liquid nitrogen and sterile sand. The powdered tissue was resuspended in 3 ml of digestion buffer [100 mM NaCl; 10 mM Trizma Base, pH 8; 25 mM EDTA, pH 8; 0.5% SDS; 0.1 mg/ml proteinase K] and incubated at 50° C for 12-18 hours with shaking. The sample was centrifuged at 5,000 xg for 5 minutes. The supernatant was transferred to a sterile falcon tube to which an equal volume of phenol/chloroform/isoamyl alcohol was added and briefly vortexed. After centrifugation at 5,000 xg for 15 minutes, the aqueous (top) layer was transferred into sterile eppendorf tubes to which 0.5 volumes of 7.5 M ammonium acetate and 2 (original) volumes of 100% EtOH were added and centrifuged (12,000-16,000 xg) for 5 minutes. The precipitated DNA pellets were washed in 70% EtOH. After centrifugation (12,000-16,000 xg; 5 minutes), the pellets were allowed to air dry for 20 minutes and resuspended in 20 µl of 1X TE buffer and stored at -20°C.

V. *pBluescript II KS (+/-) Phagemid*

The 2,961 basepair (bp) phagemid, derived from pUC19, was purchased from Stratagene, La Jolla, CA. A portion of the *lac*Z gene, encoding ß-galactosidase, located within the vector, confers blue/white color selection of recombinants in the presence of IPTG and Xgal. The multiple cloning site [MCS] is oriented such that cloning into the region results in disruption of *lac*Z translation. Ampicillan selection utilized the ampicillan gene located within the phagemid vector.

VI. *M13mp18 and M13mp19 Plasmid*

The 7,249 bp M13 vectors, as described by Messing and Vieira (1982), were purchased from Boehinger-Mannheim, GmbH, Germany.

VII. *Agarose Gel Electrophoresis*

The condition of the DNA in all of the experiments was analyzed by agarose gel electrophoresis. DNA was loaded into wells of an agarose gel (1%) in 1X Tris-Phosphate [TPE] buffer (0.08 M Trizma base; 0.008 M EDTA; 85% H_3PO_4 [1.679 g/ml]) and electrophoresed at 70 volts. The gel was stained with EtBr (50 mg/ml) and visualized on a transilluminator.

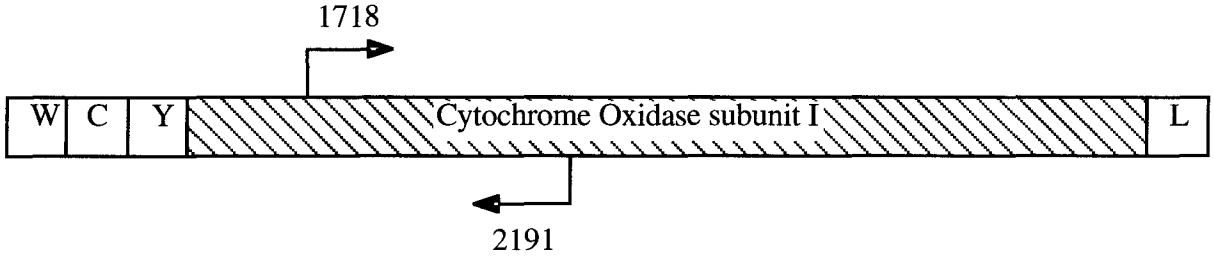
VIII. *Polymerase Chain Reaction Primers*

As described by Simon et al. (1994), the 26mer C1-N-2191, designed by R. Harrison, and the 26mer C1-J-1718, designed by B. Farrell, have been shown to exhibit homology with the minority and majority strand, respectively, of the cytochrome oxidase subunit I gene of many insect species, including members of the order Lepidoptera (Figure 9). A *Bam*HI recognition site (underlined) has been added to the 5' end of primer C1-N-2191and is referred to as 78X-1 [$^{5'}$ATGGATCCCGGTAAAATTAAAATATAAACTTC$^{3'}$]. A *Pst*I recognition site (underlined) has been added to the 5' end of primer C1-J-1718 and is referred to as 78X-2 [$^{5'}$ATCTGCAGGAGGATTTGGAAATTGATTAGTTCC$^{3'}$].

IX. *Amplification of Insect DNA*

Insect DNA was amplified by adding 10 µl of template (insect) DNA, 2 µl of primer 78X-1, 2 µl of primer 78X-2, 3 µl of *Taq* 10X buffer, 3 µl of 10X dNTP's, and 9 µl of sterile water to a sterile eppendorf tube. The mixture was denatured by heating at +95°C on a heating block for 5 minutes. Following the addition of 1 µl of *Taq* DNA polymerase, the reaction mixture was briefly centrifuged (12,000-16,000 xg) and overlayed with 10 µl of sterile mineral oil and placed inside a thermocycler. The program PCREX45 was utilized for all insect DNA amplifications as follows. The reaction mixture

Figure 9) Primer binding sites showing the location of the 3' base in reference to the cytochrome oxidase subunit I gene (shaded). Associated tRNA genes are also shown where W = tryptophan, C = cysteine, Y = tyrosine, L = leucine.


1718
W
C
Y
Cytochrome Oxidase subunit I
L
2191

was denatured at 92.5°C for 5 minutes. Initial primer annealing and elongation steps were maintained for 2 minutes each at 45°C and 72°C, respectively. One cycle included 92.5°C for 30 seconds, 45°C for 30 seconds, and 72°C for 2 minutes. The cycle was repeated 42 times. After amplification, the samples were stored at 4°C. Five microliters of sample was analyzed by agarose gel electrophoresis to determine if the expected 500 bp fragment was present.

X. *Purification of PCR Products*

The PCR reaction products were purified using *PCR* Select-II spin columns. The column buffer was removed by placing the column inside a collection tube and centrifuged at 1,000 xg for 1 minute, followed by a second centrifugation at 1,000 xg for 5 minutes. The column was then placed inside a second collection tube to which 50 µl of the PCR sample was added. After centrifugation at 1,000 xg for 3 minutes, the purified PCR product was collected in approximately 50 µl of 1X TE buffer and stored at -20°C. Twenty micoliters of the sample was analyzed to determine if the PCR product was still present.

XI. *Restriction Digestion of Vector*

Five micrograms of either vector was digested in the presence of water, a restriction endonuclease, and 10X buffer, as described by the manufacturer, overnight at 37°C. A small sample was ran on a 1% agarose gel. If the reaction was complete, 400 µl of neutralized phenol was added to the tube and centrifuged (12,000-16,000 xg) at room temperature for 15 minutes. The aqueous (top) layer was then transferred to a second sterile eppendorf tube to which 400 µl of chloroform was added. After centrifugation (12,000-16,000 xg) for 10 minutes at room temperature, the aqueous layer was again transferred to a sterile eppendorf tube followed by the addition of 400 µl of isopropanol. The sample was centrifuged (12,000-16,000 xg) at room temperature for 5 minutes. The liquid was then decanted, and the DNA pellet was allowed to air dry for 20 minutes. The

pellet was resuspended in 20 µl of 1X TE buffer (0.01 M Trizma base, pH 8.0; 0.001 M EDTA, pH 8.0).

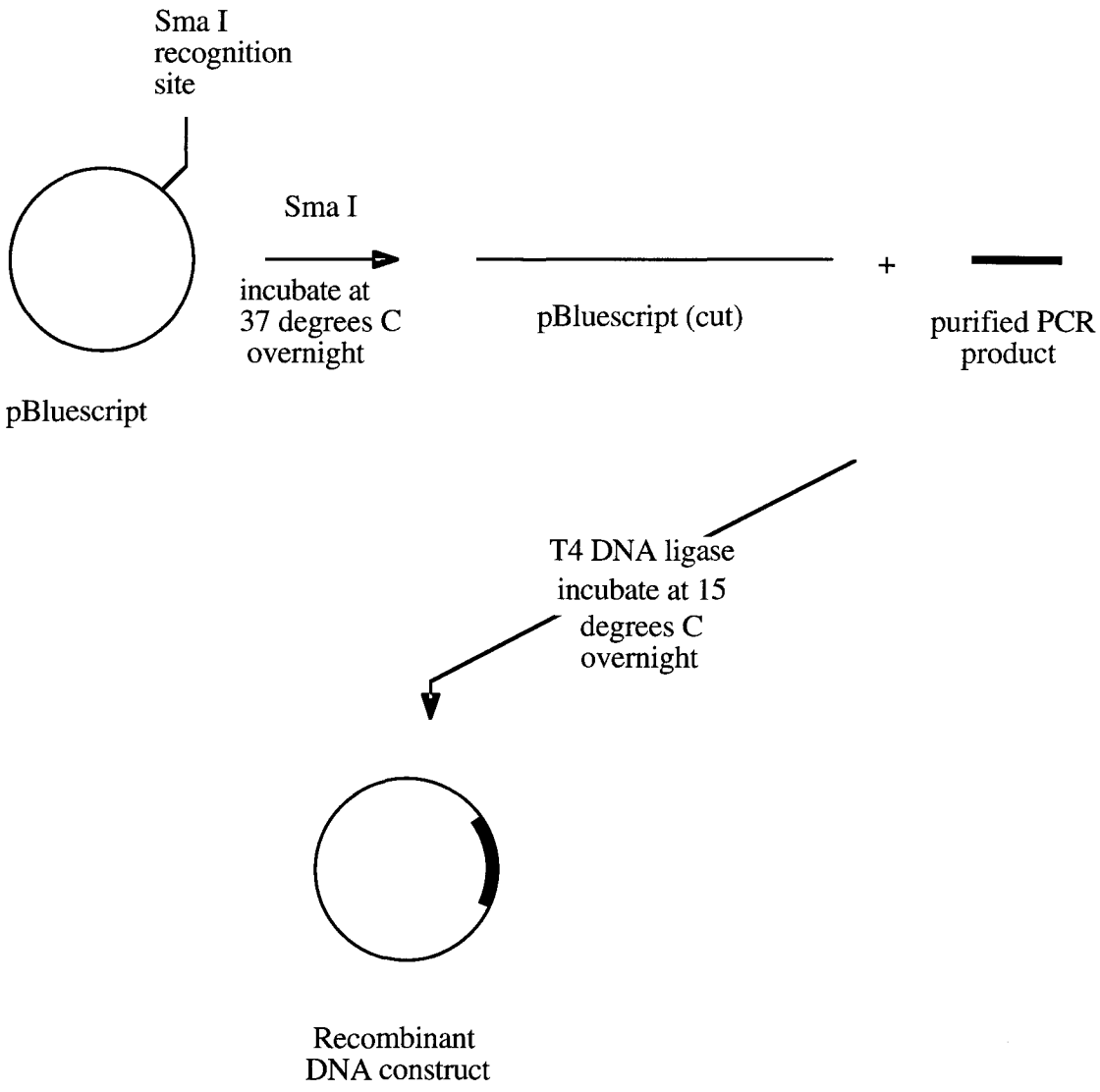
XII. *Construction of Recombinant pBluescript Plasmid*

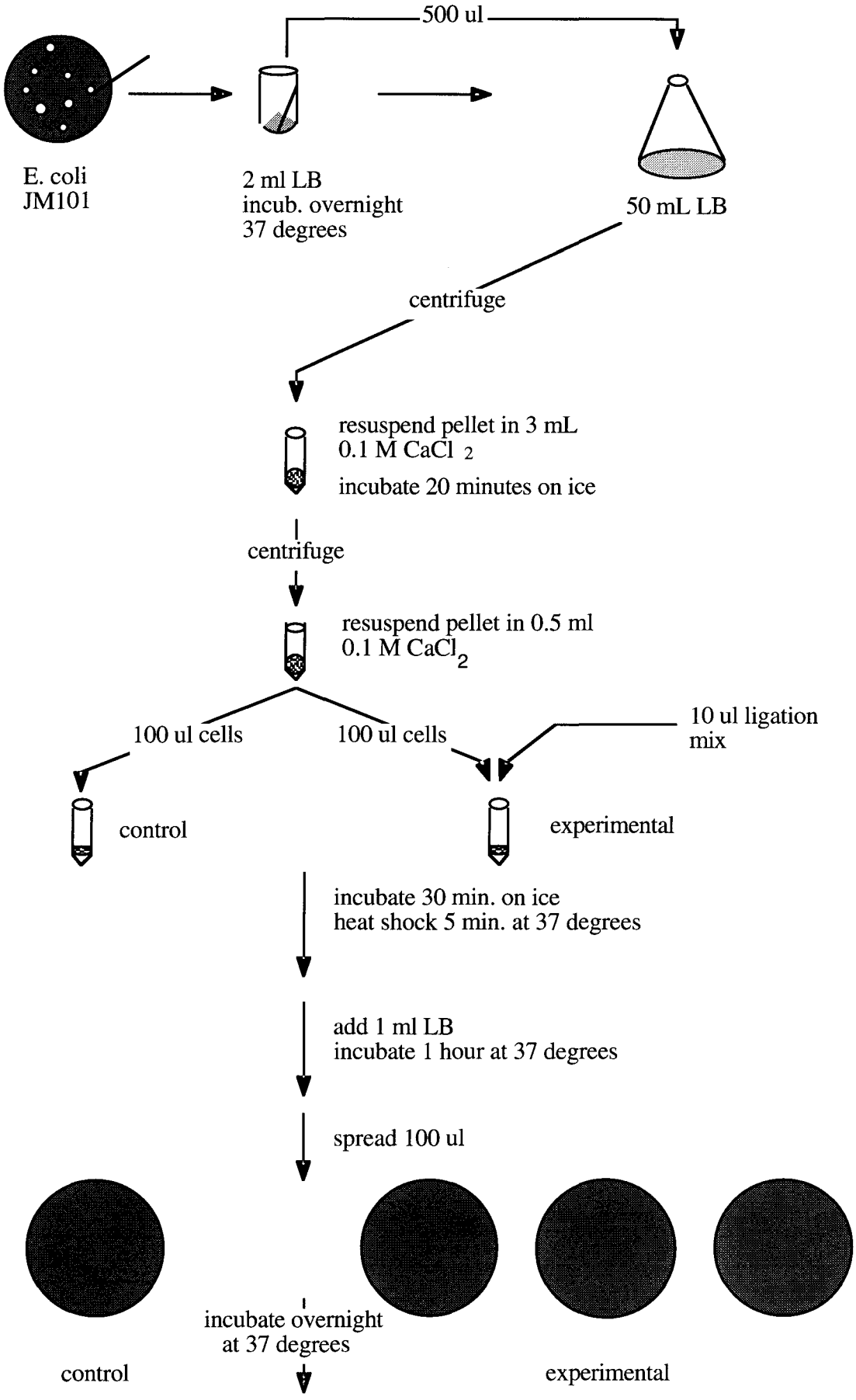
Recombinant DNA was constructed by ligating purified PCR products into pBluescript vector digested with restriction endonuclease *Sma* I (figure 10). Briefly, 15 µl of the insert, 5 µl of the vector, 3 µl of 10X ligase buffer, 6 µl of sterile water, and 1 µl of T4 DNA ligase were added to a sterile eppendorf tube and incubated at 15°C overnight.

XIII. *Transformation of Escherichia coli JM101 with pBluescript Phagemid*

E. coli JM101 was transformed with recombinant phagemid DNA using the $CaCl_2$ method (Figure 11). Two milliliters of LB was inoculated with JM101 and incubated overnight at 37°C. The following day, 50 ml of LB was inoculated with 0.5 ml of overnight cells and incubated at 37°C for 3 hours. The cells were chilled on ice for 5 minutes and centrifuged (10,000 xg) for 10 minutes. The pellet was resuspended in 15 ml of chilled 0.1 M $CaCl_2$ and placed on ice for 30 minutes. After incubation the cells were centrifuged (10,000 xg) for 10 minutes and the pellet was resuspended in 0.5 ml of chilled 0.1 M $CaCl_2$. One hundred microliters of competent cells were transferred into two sterile eppendorf tubes and labeled "control" and "experimental", respectively. Ten microliters of the pBluescript ligation mix was added to the experimental tube only, and both tubes were placed on ice for 30 minutes. The tubes were then incubated on a heating block at 37°C for 5 minutes. One milliliter of LB was added to both tubes which was then placed on a heating block at 37°C for 1 hour. After the incubation period, 100 µl of the transformation mix was spread on selective media (LA_{100}; IPTG; Xgal) and incubated overnightat 37°C .

Figure 10) Construction of recombinant pBluescript plasmid.


Sma I
recognition
site
pBluescript
Sma I
incubate at
37 degrees C
overnight
pBluescript (cut)
+
purified PCR
product
T4 DNA ligase
incubate at 15
degrees C
overnight
Recombinant
DNA construct


500 ul
E. coli
JM101
2 ml LB
incub. overnight
37 degrees
50 mL LB
centrifuge
resuspend pellet in 3 mL
0.1 M CaCl 2
incubate 20 minutes on ice
centrifuge
resuspend pellet in 0.5 ml
0.1 M CaCl2
100 ul cells
100 ul cells
10 ul ligation
mix
control
experimental
incubate 30 min. on ice
heat shock 5 min. at 37 degrees
add 1 ml LB
incubate 1 hour at 37 degrees
spread 100 ul
incubate overnight
at 37 degrees
control
experimental

XIV. *Isolation of Recombinant pBluescript DNA (Alkaline Plasmid Screen)*

Ampicillan resistant, white transformants were picked to 2 ml of LA_{100} broth and incubated at 37°C overnight. One and one-half milliliters of the overnight culture was transferred to a sterile eppendorf tube and centrifuged (12,000-16,000 xg) for 5 minutes. The pellet was resuspended in 0.2 ml of G buffer (50 mM dextrose; 25 mM Trizma base, pH 8.0; and 10 mM EDTA, pH 8.0). Four hundred microliters of denaturing solution (0.2 N NaOH; 1% SDS) was added to the tube and incubated on ice for 5 minutes. Three hundred microliters of chilled neutralizing solution (3 M KOAc; 2 M HOAc) was added to the tube and incubated on ice for 15 minutes. After the incubation period, the tube was centifuged (12,000-16,000 xg) for 5 minutes. The supernatant was then transferred to a sterile eppendorf tube to which 0.54 ml of isopropanol was added and mixed by inversion. Following a 5 minute centrifugation (12,000-16,000 xg) period, the isopropanol was decanted and the pellet was washed twice with 80% EtOH. The pellet was allowed to dry for 20 minutes and resuspended in 50 µl of 1X TE buffer and stored at -20°C.

XV. *Restriction Digest of Recombinant Plasmids*

Recombinant pBluescript DNA was double digested with restriction endonucleases *Pst*I and *Sac*I. Two-hundred microliters of recombinant DNA was incubated at 37°C for 6 hours with 30 µl of *Sac*I 10X buffer, 64 µl of sterile water, and 3 µl each of restriction endonucleases *Pst*I and *Sac*I.

XVI. *Large Scale Isolation of Recombinant DNA (Qiagen Preparation)*

Transformants were picked to 2 ml of LA_{100} broth and incubated at 37°C overnight. Fifty milliliters of LA_{100} broth was inoculated with 0.5 ml of the overnight culture and incubated at 37°C overnight. The cells were transferred into a sterile centrifuge tube and placed inside an ultracentrifuge (10,000 xg) for 10 minutes. The pellet was resuspended in 7.5 ml of buffer P1 (100 µl/ml RNase A; 50mM Trixma base; 10 mM

EDTA, pH 8.0). Seven and one-half milliliters of buffer P2 (0.2 M NaOH; 1% SDS) was added to the tube and gently mixed. The sample was placed on ice for 20 minutes and centrifuged (17,000 xg) for 30 minutes. During the centrifugation period, a Qiagen tip was primed with 4 ml of buffer QBT (0.75 M NaCl; 0.05 M MOPS; 15% EtOH, pH 7.0; 0.15% Trition X-100) and allowed to empty by gravity flow. After centrifugation, the supernatant was applied to the tip and allowed to enter the resin by gravity flow. The tip was washed twice with 10 ml of buffer QC (1 M NaCl; 0.05 M MOPS; 15% EtOH, pH 7.0). The DNA was eluted with 5 ml of buffer QF (1.25 M NaCl; 0.05 M Trizma base; 15% EtOH, pH 8.5) and precipitated with 0.7 volumes of isopropanol. The sample was centrifuged (17,000 xg) for 30 minutes and the pellet was washed with 70% EtOH. After centrifugation (17,000 xg; 10 minutes), the pellet was allowed to dry for 20 minutes and resuspended in 3 ml of 1X TE buffer. The sample was then stored at -20°C.

XVII. *Isolation of Fragments*

Fragments were prepared after restriction digestion of 5 µg of recombinant DNA, as described above (section XV), and isolated by agarose gel (1%) electrophoresis in 1X TPE buffer. The portion of the gel which contained the fragment of interest was removed and placed in dialysis tubing and filled with 0.5X Tris-Acetate [TAE] buffer (0.04 M Trizma base; 0.02 M NaOAc; 0.002 M EDTA, adjusted to pH 7.9 with 121 ml glacial acid). The tubing was clamped at both ends and placed into a gel tank which underwent electrophoresis for 1 hour. During that time, an Elutip column was primed with 3 ml of high salt buffer, pH 7.4 (1 M NaCl; 0.02 M Trizma base; 0.001 M EDTA) and 3 ml of low salt buffer, pH 7.4 (0.2 M NaCl; 0.02 M Trizma base; 0.001 M EDTA). After electrophoresis, the liquid in the tubing was transferred to the primed Elutip column. The DNA was eluted with 400 µl of high salt buffer, extracted and precipitated as described above (section XVI). The pellet was air dried for 20 minutes and resuspended in 20 µl of 1X TE buffer.

XVIII. *Construction of M13 Plasmid DNA*

Recombinant M13 DNA was constructed by ligating isolated fragments digested with *Pst*I and *Sac*I into M13 vector digested with *Pst*I and *Sac*I. Briefly, 15 µl of the isolated fragment, 5 µl of the vector, 3 µl of 10X ligase buffer, 6 µl of sterile water, and 1 µl of T4 DNA ligase were added to a sterile eppendorf tube and incubated at 15°C overnight.

XIX. *Transformation of Escherichia coli with M13 DNA*

E. coli JM101 was transformed with recombinant M13 DNA using the $CaCl_2$ method. Two milliliters of LB was inoculated with JM101 and incubated overnight at 37°C. The following day, 50 ml of LB was inoculated with 500 µl of the overnight culture and incubated at 37°C for 3 hours. The cells were chilled on ice for 5 minutes and centrifuged (10,000 xg) for 10 minutes. The pellet was resuspended in 15 ml of chilled 0.1 M $CaCl_2$ and placed on ice for 30 minutes. After incubation, the cells were centrifuged (10,000 xg) for 10 minutes and the pellet was resuspended in 0.5 ml of chilled 0.1 M $CaCl_2$. One hundred microliters of competent cells were transferred into two sterile eppendorf tubes and labeled "control" and "experimental", respectively. Fifteen microliters of M13 ligation mix was added to the "experimental" tube only, and both were placed on a on ice for 30 minutes. The tubes were then placed on a heating block at 37°C for 5 minutes. During the incubation period, 50 µl of IPTG (200mM), 50 µl of Xgal (2%), 200 µl of JM101 lawn cells were added to 2 test tubes and labeled "experimental" and "control", respectively. After the incubation period, 3 ml of soft YT agar (YT broth; 0.75% agar) was added to each test tube followed by the addition of 35 µl of transformation mix to each respective tube. The tubes were then poured onto hard YT plates (YT broth; 1.5% agar). After the media had solidified at room temperature, the plates were incubated at 37°C overnight.

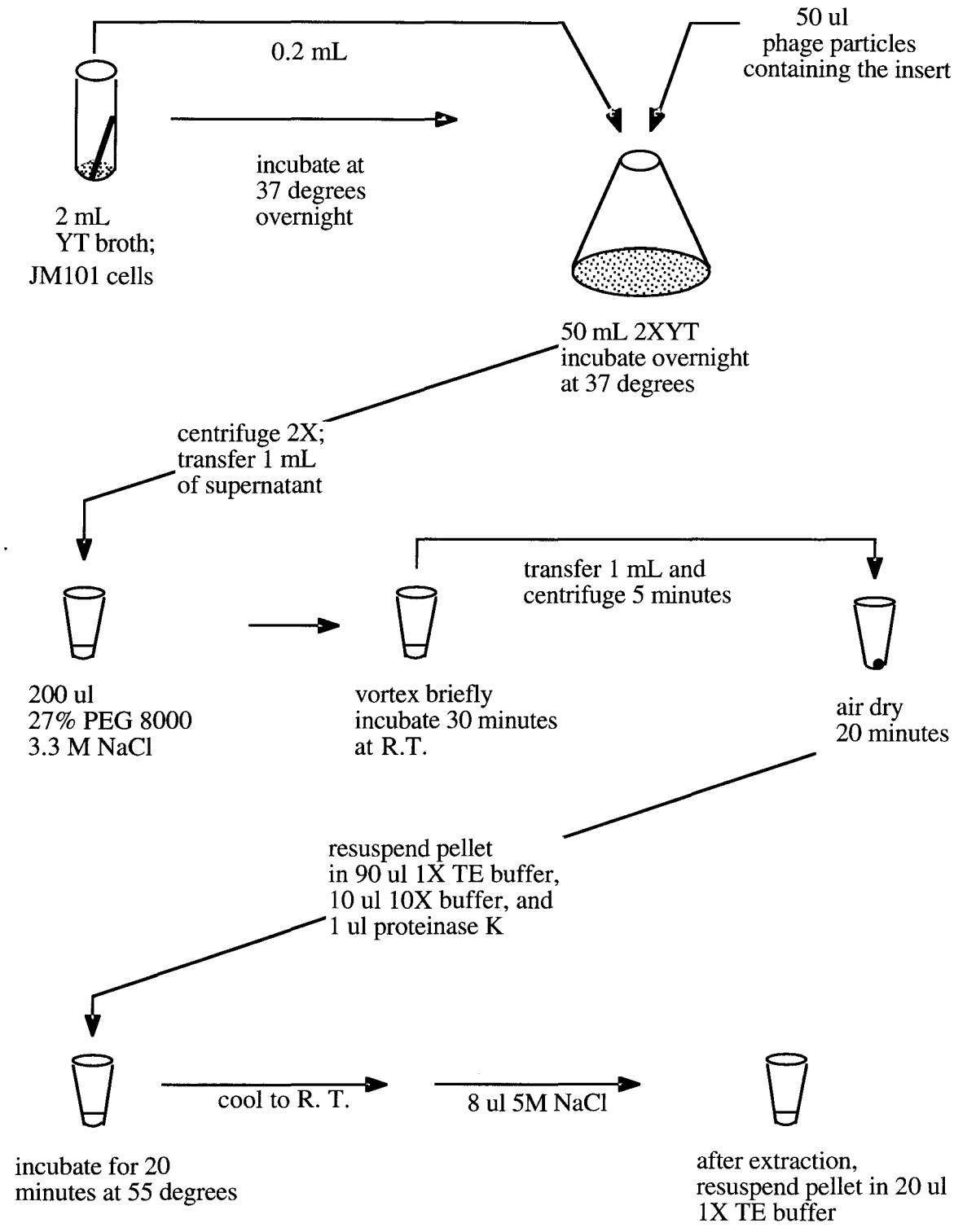
XX. *Isolation of Recombinant M13 DNA*

M13 transformants were picked to 2 ml of LB and incubated at 37°C overnight. One and one-half milliliters of the overnight culture was added to a sterile eppendorf tube and centrifuged (12,000-16,000 xg) for 5 minutes. Fifty microliters of the supernatant was transferred to a sterile eppendorf tube containing 5 µl of 2% SDS and 5 µl of dye (bromo-phenol blue) and loaded onto an agarose (1%) gel. The remaining amount of supernatant was transferred to a sterile tube and stored at 4°C as a stock solution.

XXI. *Isolation of Single-Stranded M13 DNA*

Two milliliters of YT broth was inoculated with JM101 and incubated at 37°C overnight (figure 12). The following day, 0.2 ml of the overnight culture and 50 µl of phage particles, containing the insert, were transferred to 50 ml of 2XYT broth and incubated at 37°C overnight. The broth was transferred to a sterile centrifuge tube and placed in an ultracentrifuge at 10,000 xg for 10 minutes. The supernatant was then transferred to a sterile tube. One and one-half milliliters of the supernatant was transferred to a sterile eppendorf tube and centrifuged (12,000-16,000 xg) for 5 minutes. Avoiding any bacterial debris, 1 ml of the supernatant was transferred to a second sterile eppendorf tube containing 200 µl of 27% PEG 8000 and 3.3 M NaCl and vortexed briefly. The sample was incubated for 30 minutes at room temperature. One milliliter of the PEG precipitated supernatant was transferred to a sterile eppendorf tube and centrifuged (12,000-16,000 xg) for 5 minutes. After the supernatant was decanted, the sides of the tube were carefully wiped twice with a sterile swab and allowed to air dry for 20 minutes. The pelleted DNA was resuspended in 90 µl of 1X TE buffer, 10 µl of 10X buffer [0.1 M Trizma base, pH 7.8; 0.01 M EDTA; 2% Sarkosyl], and 1 µl of Proteinase K (5mg/ml). The sample was incubated at 55°C for 20 minutes. After the incubation period, the sample was cooled to room temperature and 8 µl of 5M NaCl was added. The sample was extracted once with phenol m-cresol and twice with chloroform. Single-stranded phage

Figure 12) Isolation of single-stranded M13 DNA.


0.2 mL
50 ul
phage particles
containing the insert
incubate at
37 degrees
overnight
2 mL
YT broth;
JM101 cells
50 mL 2XYT
incubate overnight
at 37 degrees
centrifuge 2X;
transfer 1 mL
of supernatant
transfer 1 mL and
centrifuge 5 minutes
200 ul
27% PEG 8000
3.3 M NaCl
vortex briefly
incubate 30 minutes
at R.T.
air dry
20 minutes
resuspend pellet
in 90 ul 1X TE buffer,
10 ul 10X buffer, and
1 ul proteinase K
cool to R. T.
8 ul 5M NaCl
incubate for 20
minutes at 55 degrees
after extraction,
resuspend pellet in 20 ul
1X TE buffer

DNA was precipitated with 2 volumes of isopropanol and centrifuged (12,000-16,000) for 15 minutes. The DNA pellet was washed with 2 volumes of 80% EtOH and allowed to dry for 20 minutes. The pellet was resuspended in 20 µl of 1X TE buffer and stored at -20°C.

XXII. *Sequencing Reactions*

The preparation of the primer annealing mixture was in accordance to the manufacture's instructions. Briefly, 5 µl of template (single-stranded M13; see XX above) DNA was placed in a sterile eppendorf tube along with 2 µl of 10X reaction buffer, 2 µl of DIG-labeled M13/pUC19 forward or reverse sequencing primer, and 10 µl of *Taq* DNA polymerase (3U/µl). Four vials, labeled G, A, T, and C, respectively, were filled with 2 µl of the appropriate extension/termination mixture. Each vial was then filled with 4 µl of the primer annealing mixture and briefly centrifuged (16,000 xg). After centrifugation, 10 µl of sterile mineral oil was placed into each vial forming an overlay and placed in a thermocycler. Both mixtures were initialy denatured at 95°C for 5 minutes. For the forward sequencing primer, one cycle included 95°C for 5 minutes, 60°C for 30 seconds, and 70°C for 1 minute. For the reverse sequencing primer, one cycle included 95°C for 1 minute, 56°C for 1 minute, and 70°C for 1 minute. Both cycles were repeated 29 times. The samples were stored at 4°C after amplification. The PCR products were then separated from the mineral oil to which 2 µl of formamide buffer was added to stop the reaction. The samples were stored at -20°C.

XXIII. *Sequencing Gel Electrophoresis*

An 8% polyacrylamide gel [1X TBE (0.135 M Trizma base; 0.045 M boric acid; 0.0025 M EDTA), 8 M urea] was cast into a mold. The primer annealing mixtures were denatured at +95°C for 5 minutes. The vials were placed on ice and briefly centrifuged (16,000 xg). Three microliters of each vial (G,A,T,C) was loaded into corresponding wells of the sequencing gel. The gel was ran at 2000 volts for approximately 8 hours in

1X TBE buffer. After the 8 hour electrophoresis period, the same primer annealing mixtures were loaded into four different wells of the sequencing gel, and ran approximately 4 hours at 2000 volts.

XXIV. *Contact Blotting*

After electrophoresis, one of the glass plates was removed. A small amount of 1X TBE buffer was applied to a positively charged nylon membrane and placed on the sequencing gel. The membrane was covered with a sheet of 3 mm Whatman chromatography paper and the glass plate was then placed back on top. The sandwich was charged with a weight, approximately 2 kilograms, for 20 minutes. The membrane was then removed from the sandwich and placed inside a UV crosslinker. After crosslinking, the membrane was placed inside a plastic bag, sealed, and stored at 4°C until detection.

XXV. *Detection*

All detection steps were performed in pyrex© dishes at room temperature. The membrane was washed in washing buffer [Buffer 1 (0.1 M maleic acid; 0.15 NaCl, pH 7.5), 0.3% Tween 20] for 5 minutes. After the washing buffer was decanted, the membrane was incubated for 30 minutes in buffer 2 [10% blocking stock solution diluted 1:10 in buffer 1]. The blocking buffer was removed and the membrane was incubated for 30 minutes in antibody solution [anti-DIG-AP conjugate diluted 1:10,000 in buffer 2]. The membrane was then transferred to a second clean dish and washed twice, 15 minutes per wash, in washing buffer. The washing buffer was discarded and the membrane was equilibrated in detection buffer [0.1 M Trizma base; 0.1 M NaCl; 0.05 M $MgCl_2$, pH 9.5] for 5 minutes. After placing the damp membrane on sahran wrap, 1-2 ml of diluted CSPD [diluted 1:100 in buffer 3] was added. After 5 minutes, the excess CSPD was allowed to drain off of the membrane, and sealed in a second piece of sahran wrap. After a 20 minute incubation period at 37°C, the membrane was exposed to x-ray film for 8-12 hours.

RESULTS

I. *Isolation of Genomic DNA from Insect Specimens*

In order to estimate the phylogeny of the following insect species, genomic DNA was extracted and analyzed from *Catocala piatrix, C. grynea, C. sordida, C. cara, C. neogama, C. mira, C. subnata*, and *Euparthenos nubilis*. Genomic DNA that is analyzed by agarose gel electrophoresis typically appears as a relatively high molecular weight smear rather than as distinct fragments. Chromosomal DNA and other large molecules are sterically hindered from migrating through the gel matrix easily and are observed in the top portion of agarose gels.

Figure 13 shows samples of DNA from *C. cara* which were extracted and analyzed by agarose gel electrophoresis. As expected, each sample appears as a smear of DNA without the presence of any distinct fragments or bands. At the bottom of this gel there appears to be a cloud of RNA, which migrates much faster than DNA on agarose gels. These results indicated that total genomic DNA had been successfully isolated from the specimen facilitating further isolations of the other insect species.

All of the specimens produced isolated DNA similar to figure 13 with the exception of *C. subnata,* which resembled isolated RNA (figure 14, A). To determine whether the samples were extracts of RNA, a sample was treated with the RNase A enzyme prior to electrophoresis and analyzed. If RNA had been present in the sample, the enzyme would cleave the molecules into its constituent nucleotides. If, on the other hand, the sample did not contain isolated RNA, the sample would appear the same as in figure 14 (A). Figure 14 (B) does not show any trace of RNA (or DNA) as observed in figure 14 (A) indicating that RNA had been isolated from *C. subnata.* It has not been determined why RNA had been isolated from this specimen, since the same protocol was used for all genomic DNA isolations.

Figure 13) Genomic DNA isolated from *C. cara*. Lane 1: size standard; Lanes 2-4 show samples of isolated DNA from *C. cara* and represent typical genomic DNA isolates as analyzed by agarose gel electrphoresis.

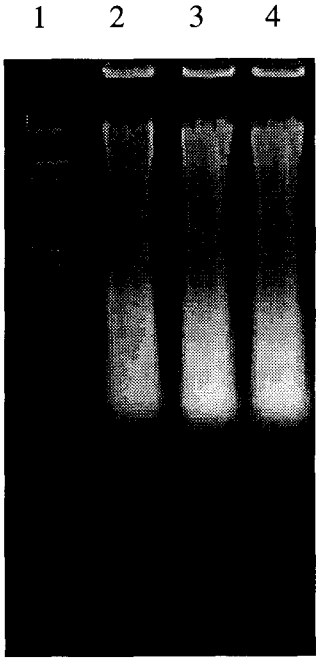
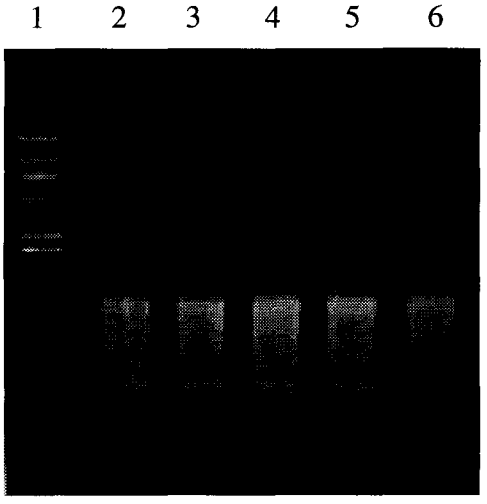
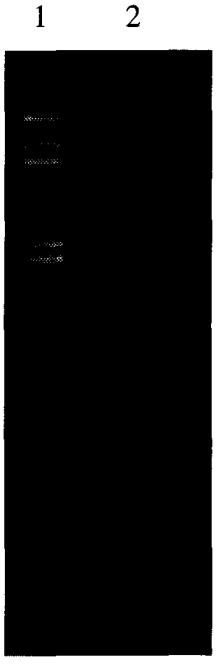

1
2
3
4

Figure 14) Genomic DNA isolation of *C. subnata*. A. The result of the genomic DNA isolation from *C. subnata*. Lane 1: size standard; Lanes 2-6 show the samples of the extracted material which resembles isolated RNA. B. RNase digestion of isolated material from *C. subnata*. Lane 1: size standard; Lane 2: shows no trace of isolated material which indicates that RNA was isolated from *C. subnata*.

A.



B.



II. *Amplification of Insect DNA*

Samples were chosen for amplification based on the amount and integrity of the DNA isolates following agarose gel electrophoresis. Polymerase chain reaction (PCR) primers 78X-1 and 78X-2 were used in an attempt to amplify a 500 base pair (bp) region within the cytochrome oxidase subunit I (CO-I) gene. The primers anneal to complementary sequences on opposite strands of the template and provide a 3' hydroxyl group which can be extended by the *Taq* polymerase. The primers had sequences that were conserved in mammals, insects, reptiles, and crustaceans. Successful PCR amplifications produce several copies of the target region which increase with the number of cycles in the PCR reaction. After amplification, a sample is analyzed by agarose gel electrophoresis and observed as a single bright band of DNA, indicating that multiple copies of the target sequence are present. The size of the PCR products are estimated by comparing it to a size standard (e.g. lambda DNA cleaved with restriction endonuclease *Hind* III). We expected primers 78X-1 and 78X-2 to produce a band of about 500 bp in length based on data from other species.

As expected, a 500 bp fragment from *C. grynea, C. cara,* and *C. neogama* was produced (figure 15). DNA extracts from *C. piatrix* and *C. sordida* also produced 500 bp PCR products (figure 16). *C. mira* failed to produce a band following PCR however (figure 17, A; lane 3). This result may suggest that the primer binding sites were not complementary to one (or both) of the primers which prevented binding and subsequent amplification of the target sequence. The reaction was repeated using a positive control (*C. grynea*) to verify the accuracy of the reaction. The results showed that no amplified PCR products were produced from *C. mira* (figure 17, B), confirming the results in the previous experiment.

Surprisingly, however, the same PCR primers amplified a 500 bp region from genomic DNA isolated from *E. nubilis*, commonly known as the Locust Underwing (figure 17, A; lane 2). *Euparthenos* and *Catocala* genera belong to the same subfamily Catocalinae

Figure 15) PCR products from *C. grynea*, *C. cara*, and *C. neogama*. Lane 1: size standard (arrow indicates 500 bp marker); Lane 2: *C. grynea*; Lane 3: *C. cara*; Lane 4: *C. neogama*.

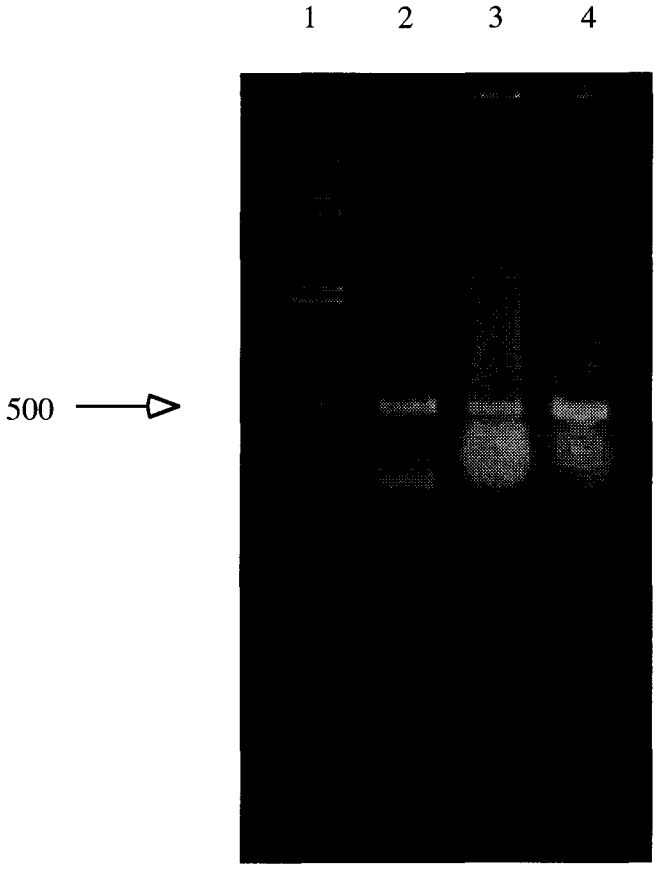

1
2
3
4
500

Figure 16) PCR products from *C. piatrix* and *C. sordida*. Lane 1: size standard (arrow indicates 500 bp marker); Lane 2: positive control (*C. grynea*); Lane 3: *C. piatrix*; Lane 4: *C. sordida*.

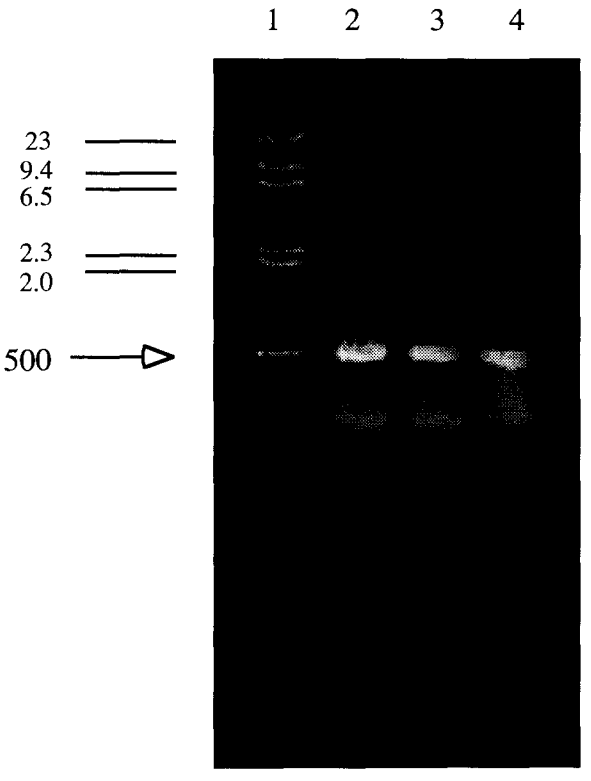
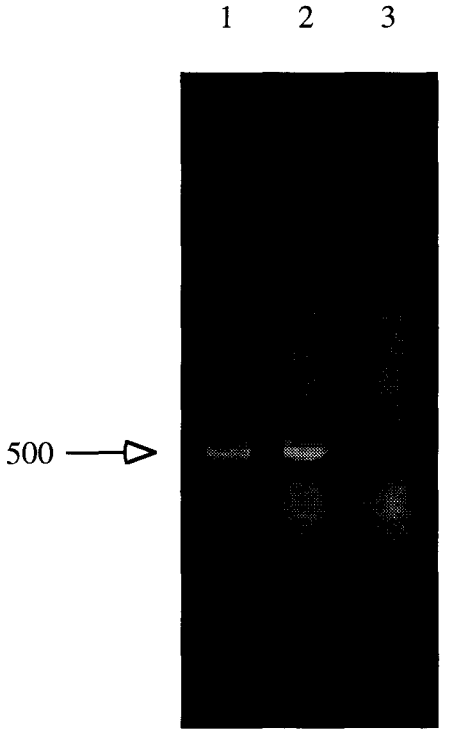
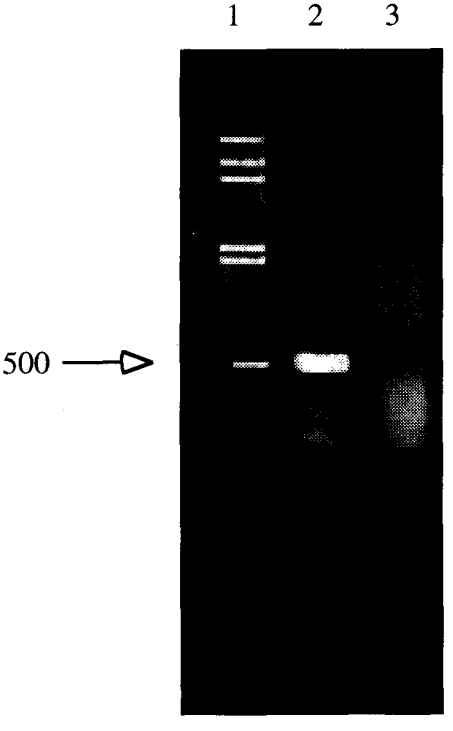

1
2
3
4
23
9.4
6.5
2.3
2.0
500

Figure 17) PCR products from *C. mira* and *E. nubilis*. A. Lane 1: positive control (*C. grynea*); Lane 2: *E. nubilis*; Lane 3: *C. mira* showing no PCR product. B. Lane 1: size standard; Lane 2: positive control (*C. grynea*); Lane 3: *C. mira* showing no PCR product.

A.



B.



and may share a common evolutionary history. Although this result may suggest that *E. nubilis* is related more closely than *C. mira* is to the other *Catocala* species examined, the results are preliminary and require further analysis.

The results of the PCR reactions show that total cellular DNA preparations are sufficient in amplifying mitochondrial targets. It was therefore, not necessary to isolate and purify mitochondrial DNA (mtDNA) prior to the amplification process.

III. *Construction of Recombinant DNA using Amplified CO-I Products*

Since PCR products are naturally blunt-ended, they can be ligated into blunt-ended vectors. Samples of the amplified PCR products from *C. piatrix*, *C. grynea,* and *C. sordida* were purified and blunt-end ligated into pBluescript cleaved with restriction endonuclease *Sma* I. The ligations were then transformed into *E. coli*. Clones containing the amplified PCR products were chosen by blue and white selection and confirmed by agarose gel electrophoresis. Recombinant plasmids containing the 500 bp CO-I gene region were expected to be 3.4 kilobase pairs (kb) in length.

Figure 18 shows samples from an alkaline plasmid screen of *C. sordida* clones. Lanes 2 and 4 appear to contain recombinant constructs in comparison to the size standard (lane 1). Lanes 3, 5, 6, and 7 appear to contain plasmids approximately 2.9 kb in length, indicating that these clones were transformed with naive plasmids.

In order to confirm whether the clones possessed recombinant plasmids, isolated samples were digested with restriction endonucleases *Pst* I and *Sac* I. The recognition sites for these enzymes flank the ligation site (*Sma* I) within the vector. After digestion, the 500 bp PCR product, if present, would be cleaved and separated from the vector by agarose gel electrophoresis. Figure 19 shows the result of a double-digest of clones containing the PCR product produced from *C. sordida*. A 500 bp fragment was produced following digestion which indicated that the corresponding clone contained the 500 bp CO-I gene

Figure 18) Alkaline Plasmid Screen of *C. sordida* clones. Lane 1: Size standard showing the sizes of the fragments; Lanes 2 and 4: Show samples containing plasmids approximately 3.4 kb in length. They represent recombinant DNA constructs; Lanes 3, 5-7: Show samples containing plasmids approximately 2.9 kb in length and represent naive plasmids.

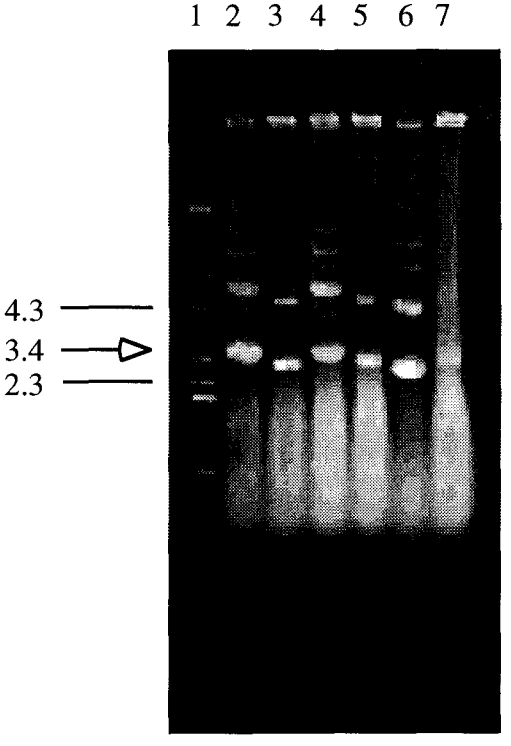
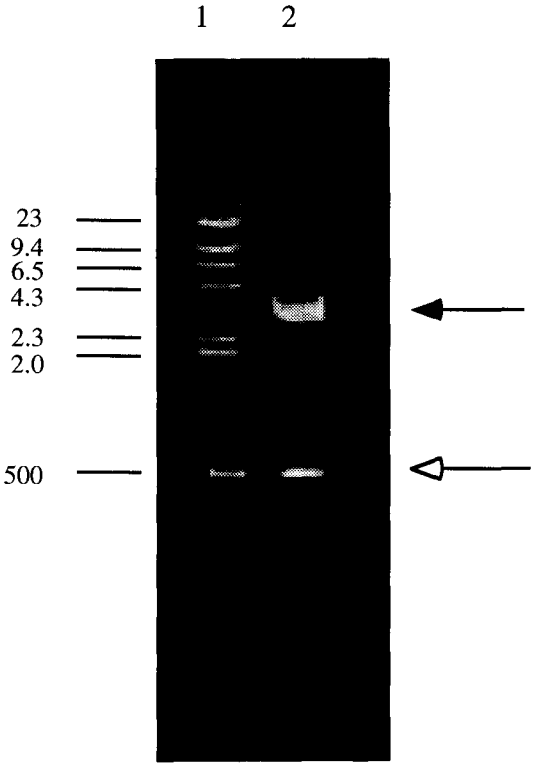

1 2 3 4 5 6 7
4.3
3.4
2.3

Figure 19) Double-digest of recombinant DNA construct of *C. sordida*. Lane 1: size standard with sizes of the fragments; Lane 2: Double-digest with restriction endonucleases *Pst* I and *Sac* I. Closed arrow indicates cut plasmid (approximately 2.9 kb in length). Open arrow indicates PCR product of *C. sordida* (approximately 500 bp in length).


1
2
23
9.4
6.5
4.3
2.3
2.0
500

region. Clones containing the PCR products of *C. piatrix* and *C. grynea* were also analyzed using this method. The results of these experiments identified clones which contained 500 bp CO-I gene regions from each respective moth species.

IV. *Construction of Single-stranded Template DNA*

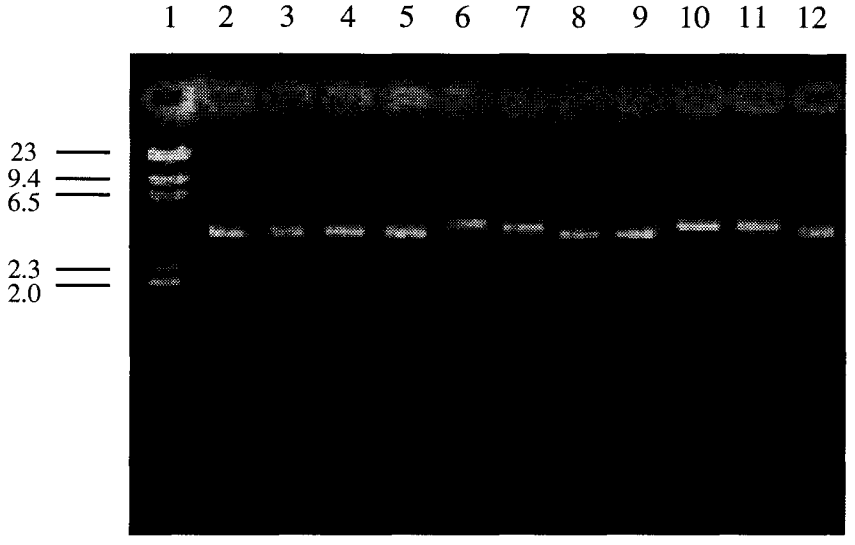
To facilitate DNA sequencing of the CO-I gene regions of *C. piatrix*, *C. grynea*, and *C. sordida*, the PCR products were ligated into the M13mp18 cloning vector cut with restriction endonucleases *Pst* I and *Sac* I. The ligations were transformed into *E. coli* and recombinant M13 clones were identified on the basis of blue and white plaque selection. After selection, single-stranded template DNA was extracted and analyzed by agarose gel electrophoresis. The samples were compared to a negative control which allowed the recombinant constructs to be identified on the basis of shift differences between the M13 samples.

Figure 20 shows single-stranded M13 DNA isolated from *E. coli*. Lanes 6, 10, and 11 have migrated at a slower rate than the negative control (lane 2). These results indicated that the corresponding clones were transformed with recombinant DNA constructs containing the PCR products. Clones containing amplified CO-I gene regions from *C. piatrix*, *C. grynea*, and *C. sordida* were all isolated and identified using this technique.

V. *DNA Sequencing of C. piatrix*, *C. grynea*, and *C. sordida*

DNA sequence data was obtained from *C. piatrix*, *C. grynea*, and *C. sordida* using the M13/pUC forward sequencing primer. The primer recognizes and binds to a flanking region of the multiple cloning site and allows DNA sequencing to proceed towards the *Pst* I recognition site of the polylinker region. DNA Strider, version 1.0 was used to analyze the sequence data and to generate restriction maps for each sequence.

Figure 20) Isolated single-stranded M13 DNA. Lane 1: size standard showing sizes of the fragments; Lane 1: negative control (naive ss M13 DNA); Lanes 6, 10, and 11: Show DNA which has migrated at a slower rate than the negative control, indicating that these samples contained the 500 bp PCR product; Lanes 3-5,7-9, and 12: Show DNA which has migrated at a rate similar to that of the negative control, indicating that these samples did not contain the 500 bp PCR product.


1 2 3 4 5 6 7 8 9 10 11 12
23
9.4
6.5
2.3
2.0

The amount of sequence data that was obtained from the reactions of *C. piatrix* generated 271 bp (figure 21). The sequencing reactions from *C. grynea* (figure 22) and *C. sordida* (figure 23) produced 125 and 132 bp of sequence data, respectively. Because the original clones were constructed from blunt-end ligations into the pBluescript cloning vector, the PCR products were capable of being ligated in two pausible orientations. The sequence data that was obtained from the reactions revealed that *C. sordida* was ligated in the opposite orientation to that of *C. piatrix* and *C. grynea.* Therefore, the sequence data from *C. sordida* was obtained from the opposite (complementary) strand to that of *C. piatrix* and *C. grynea.* To facilitate direct comparisons of the sequence data, the PCR product from *C. sordida* was isolated and ligated into M13mp18 and M13mp19 cloning vectors and were transformed into *E. coli.* The clones will allow sequence data from both strands to be obtained in future experiments.

VI. *Sequence Analysis using GenBank Search Capabilities*

In order to determine whether the sequence data of the *Catocala* species were from the CO-I gene region, the sequences were entered into the GenBank database on the World Wide Web. The uniform resource locator (URL) used for the search was BLAST (Basic Alignment Search Tool) (URL://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast). This site allows a query sequence to be entered and compared to other nucleic acid sequences within the database.

Searches optimized to find nearly identical non-redundant sequences using *C. piatrix* sequence data was used. All of the corresponding sequences that were obtained from this search were from CO-I gene regions. Figure 24 shows matches obtained from three insect species. The amount of sequence homology appears to decrease as taxonomic divergences increase. *Spodoptera frugiperda* exhibits the highest degree of homology (88%) and belongs to the same family as the underwings (Noctuidae). *Manduca sexta* (Lepidoptera: Sphingidae) and *Anopheles gambiae* (Insecta: Diptera) show sequence

Figure 21) Sequence analysis and restriction map of the CO-I gene region from *C. piatrix* generated by DNA Strider, version 1.0, using the M13/pUC forward primer.

Figure 21 -> Restriction Map

DNA sequence 271 b.p. tttttgacttct ... attacagctttc linear

Figure 22) Sequence analysis and restriction map of the CO-I gene region from *C. grynea* generated by DNA Strider, version 1.0, using the M13/pUC forward primer.

Figure 22 -> Restriction Map

DNA sequence 125 b.p. cttctaatattc ... tacgattaaata linear

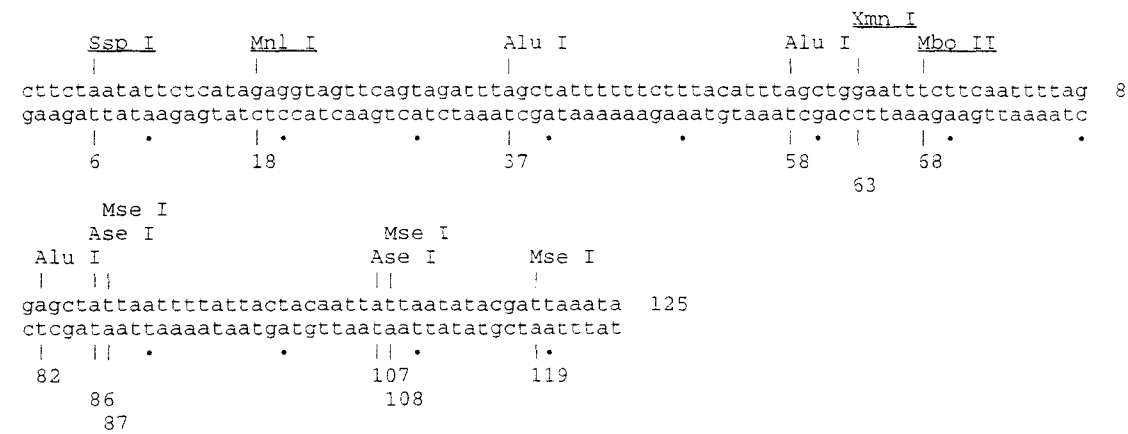


Figure 23) Sequence analysis and restriction map of the CO-I gene region from *C. sordida* generated by DNA Strider, version 1.0, using the M13/pUC forward primer.

Figure 23 -> Restriction Map

DNA sequence 132 b.p. ttttaaatttcg ... aaatattaaatt linear

Figure 24) Results of non-redundant sequence comparisons of *C. piatrix*, using the M13/pUC forward primer, performed by BLAST. Query sequence represents *C. piatrix*, Sbjct represents the known sequence listed above each comparison. The lines between each base represent homologous sites.

gb|U72977|SFU72977 Spodoptera frugiperda strain rice haplotype 1 cytochrome oxidase subunit I gene, mitochondrial gene encoding mitochondrial protein, partial cds
Length = 683

Plus Strand HSPs:

Score = 916 (253.1 bits), Expect = 1.3e-79, Sum P(2) = 1.3e-79
Identities = 204/230 (88%), Positives = 204/230 (88%), Strand = Plus / Plus

```
Query:    41 ATTGTAGAAAATGGAGCAGGAACTGGATGAACAGTATCACCCCCTCTTTCTTCCAATATT 100
             ||||||||||||||||||||||||||||||||||||||| | ||||| || || || ||||||
Sbjct:   340 ATTGTAGAAAATGGAGCAGGAACTGGATGAACAGTTTACCCCCCCCTCTCCTCTAATATT 399

Query:   101 CGTCATAGAGGTAGTTCAGTAGATTTAGCTATTTTTTCATTACATTTAGCTGGAATTTCT 160
               |||| | ||||||||||||||||||||||||||| ||| | ||||||||||||||||
Sbjct:   400 GCTCATGGTGGTAGTTCAGTAGATTTAGCTATTTTCTCACTTCATTTAGCTGGAATTTCA 459

Query:   161 TCAATTTTAGGAGCTATTAATTTTATTACCACAATTATTAATATAGCATTAAATAACTTA 220
             || |||||||||||||||| |||||||||| |||||||||| |||||||| |||
Sbjct:   460 TCTATTTTAGGAGCTATTAACTTTATTACCACTATTATTAATATACGATTAAATAATTTA 519

Query:   221 ATATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGAATTACAGCTTT 270
               |||||||||||||||||||||||||||||||||||| ||||| || ||
Sbjct:   520 TCATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGTATTACCGCATT 569
```

gb|U09843|MSU09843 Manduca sexta cytochrome c oxidase subunit 1 mRNA, mitochondrial gene encoding mitochondrial protein, partial cds
Length = 1616

Plus Strand HSPs:

Score = 808 (223.3 bits), Expect = 1.7e-59, P = 1.7e-59
Identities = 192/230 (83%), Positives = 192/230 (83%), Strand = Plus / Plus

```
Query:    41 ATTGTAGAAAATGGAGCAGGAACTGGATGAACAGTATCACCCCCTCTTTCTTCCAATATT 100
             |||||||||||||||| || || || ||||||||||  || ||| | || || ||||||
Sbjct:   425 ATTGTAGAAAATGGAGCTGGTACAGGTTGAACAGTATACCCACCTTTATCATCTAATATT 484

Query:   101 CGTCATAGAGGTAGTTCAGTAGATTTAGCTATTTTTTCATTACATTTAGCTGGAATTTCT 160
               |||||| || || || || |||||||| |||||||| |||||||||| || |||||
Sbjct:   485 GCTCATAGTGGAAGATCTGTTGATTTAGCAATTTTTTCTTTACATTTAGCAGGTATTTCA 544

Query:   161 TCAATTTTAGGAGCTATTAATTTTATTACCACAATTATTAATATAGCATTAAATAACTTA 220
             || |||||||||| |||||||||||||| ||||||||||||  | | |||||  ||
Sbjct:   545 TCTATTTTAGGAGCAATTAATTTTATTACTACAATTATTAATATGCGAATTAATAATATA 604

Query:   221 ATATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGAATTACAGCTTT 270
               |||||||||||||| |||||| ||||||||||||||||||||| ||
Sbjct:   605 TCATTTGATCAAATACCATTATTTGTTTGAGCTGTAGGAATTACAGCATT 654
```

gb|L20934|MSQMTCG Anopheles gambiae mitochondrial genome sequence.
Length = 15,363

Plus Strand HSPs:

Score = 622 (171.9 bits), Expect = 2.2e-46, P = 2.2e-46
Identities = 170/227 (74%), Positives = 170/227 (74%), Strand = Plus / Plus

```
Query:    41 ATTGTAGAAAATGGAGCAGGAACTGGATGAACAGTATCACCCCCTCTTTCTTCCAATATT 100
             || |||||||| || || ||||| |||||||| || |  || ||||| |||||     |||
Sbjct:  1766 ATAGTAGAAAACGGGGCTGGAACAGGATGAACTGTTTATCCTCCTCTATCTTCTGGAATT 1825

Query:   101 CGTCATAGAGGTAGTTCAGTAGATTTAGCTATTTTTTCATTACATTTAGCTGGAATTTCT 160
               ||||   ||   |||||||||||||| |||||||  | ||||||| |||||||||
Sbjct:  1826 GCTCATGCTGGAGCTTCAGTAGATTTAGCAATTTTTTCTCTTCATTTAGCAGGAATTTCT 1885

Query:   161 TCAATTTTAGGAGCTATTAATTTTATTACCACAATTATTAATATAGCATTAAATAACTTA 220
             || ||||||||||  | ||||||||||| ||| | |||||||||   |        |
Sbjct:  1886 TCTATTTTAGGAGCAGTAAATTTTATTACAACAGTAATTAATATACGGTCTCCAGGAATT 1945

Query:   221 ATATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGAATTACAGC 267
               | ||| |||| |||||| |||||| ||||| | ||||  ||||||||
Sbjct:  1946 ACATTAGATCGAATACCATTATTTGTTTGATCGGTAGTTATTACAGC 1992
```

homologies of 83% and 74%, respectively. The search also revealed that the sequence data obtained from *C. piatrix* was from the plus strand, as indicated in figure 24.

Searches using *C. grynea* sequence data and the same BLAST parameters as above (section VI), resulted in CO-I gene region comparisons similar to that of *C. piatrix* (figure 25). The similarity percentages are more variable, however, which is dependent on the length of the query sequence. All of the matched sequences were again from the CO-I gene region. *Spodoptera frugiperda* was 89% homologous. *Anopheles quadrimaculatus* (Insecta: Diptera) and *Ophrealla notulata* (Insecta: Coleoptera) were 90% and 91% homologous, respectively. These results also reveal that the sequence data from *C. grynea* was from the plus strand, as in *C. piatrix.*

Comparisons with *C. sordida* revealed that sequence data from this specimen was obtained from the minus strand, as shown in figure 26. Homologous CO-I gene regions from other organisms were also more diverse than comparisons using *C. piatrix* and *C. grynea* sequence data. Again, these results further indicate that the amount of sequence data affects homologous percentages between sequence data, leading to increasingly spurious matches. *Glossiphonia complanata* (Metozoa: Annelida) and *Galapaganus galapagoensis* (Insecta: Coleoptera) were 86% and 82% homologous, respectively. The amount of sequence homology for *Spodoptera frugiperda* ,was again the highest observed (92%). It is also evident from these results that this region is highly conserved as both the Annelid specie and the Coleoptera specie show homology to the same region.

Other searches which used less contrained parameters resulted in sequence comparisons to a wide variety of organisms. Taxonomic representatives from Fungi (*Podospora anserina*), Chordata (*Homo sapiens*), Mollusca (*Katharina tunicata*), and Nematoda (*Caenorhabditis elegans*) were among the many mitochondrial DNA sequence comparisons obtained using sequence data from *C. piatrix* (figure 27).

Figure 25) Results of non-redundant sequence comparisons of *C. grynea*, using the M13/pUC forward primer, performed by BLAST. Query sequence represents *C. grynea*, Sbjct represents the known sequence listed above each comparison. The lines between each base represent homologous sites.

gb|U72974|SFU72974 Spodoptera frugiperda strain corn haplotype 1 cytochrome oxidase subunit I gene, mitochondrial gene encoding mitochondrial protein, partial cds
Length = 683

Plus Strand HSPs:

Score = 334 (92.3 bits), Expect = 8.4e-30, Sum P(2) = 8.4e-30
Identities = 74/83 (89%), Positives = 74/83 (89%), Strand = Plus / Plus

```
Query:    42 TTTTTTCTTACATTTAGCTGGAATTTCTTCAATTTTAGGAGCTATTAATTTTATTACTAC 101
             ||| |   | |||||||||||||||||| || |||||||||||||||| ||||||||||
Sbjct:   432 TTTCTCACTTCATTTAGCTGGAATTTCATCTATTTTAGGAGCTATTAACTTTATTACTAC 491

Query:   102 AATTATTAATATACGATTAAATA 124
             |||||||||||||||||||||||
Sbjct:   492 TATTATTAATATACGATTAAATA 514
```

Score = 159 (43.9 bits), Expect = 8.4e-30, Sum P(2) = 8.4e-30
Identities = 35/39 (89%), Positives = 35/39 (89%), Strand = Plus / Plus

```
Query:    10 TTCTCATAGAGGTAGTTCAGTAGATTTAGCTATTTTTTC 48
             | ||||| | ||||||||||||||||||||||||||| ||
Sbjct:   399 TGCTCATGGTGGTAGTTCAGTAGATTTAGCTATTTTCTC 437
```

gb|L04272|MSQNCATR Anopheles quadrimaculatus NADH dehydrogenase subunits (1-4, 4L, 5-6); cytochrome oxidase subunits (1-3); adenosine triphosphatase subunits (6,8); cytochrome b; transfer RNA; ribosomal RNA (large and small subunits).
Length = 15,455

Plus Strand HSPs:

Score = 317 (87.6 bits), Expect = 2.4e-25, Sum P(2) = 2.4e-25
Identities = 69/76 (90%), Positives = 69/76 (90%), Strand = Plus / Plus

```
Query:    42 TTTTTTCTTACATTTAGCTGGAATTTCTTCAATTTTAGGAGCTATTAATTTTATTACTAC 101
             |||||  |||||||||||| |||||||||||||||||||||||  |||||||||||| ||
Sbjct:  1857 TTTTTCATTACATTTAGCAGGAATTTCTTCAATTTTAGGAGCAGTTAATTTTATTACAAC 1916

Query:   102 AATTATTAATATACGA 117
             | ||||||||||||||
Sbjct:  1917 AGTTATTAATATACGA 1932
```

gb|U20707|ONU20707 Ophraella notulata isolate 126 cytochrome oxidase gene, mitochondrial gene encoding mitochondrial protein, partial cds.
Length = 420

Plus Strand HSPs:

Score = 286 (79.0 bits), Expect = 8.9e-23, Sum P(2) = 8.9e-23
Identities = 62/68 (91%), Positives = 62/68 (91%), Strand = Plus / Plus

```
Query:    50 TACATTTAGCTGGAATTTCTTCAATTTTAGGAGCTATTAATTTTATTACTACAATTATTA 109
             ||||||||||||||| || |||||||||| || |||||||||||| ||||||||||
Sbjct:   155 TACATTTAGCTGGAATCTCCTCAATTTTAGGTGCAATTAATTTTATTACAACAATTATTA 214

Query:   110 ATATACGA 117
             |||| |||
Sbjct:   215 ATATGCGA 222
```

Score = 117 (32.3 bits), Expect = 8.9e-23, Sum P(2) = 8.9e-23
Identities = 29/36 (80%), Positives = 29/36 (80%), Strand = Plus / Plus

```
Query:    10 TTCTCATAGAGGTAGTTCAGTAGATTTAGCTATTTT 45
             | | ||| ||||    ||||||||||||||||||
Sbjct:   114 TGCCCATGGAGGATCATCAGTAGATTTAGCTATTTT 149
```

Figure 26) Results of non-redundant sequence comparisons of *C. sordida* using the M13/pUC forward primer, performed by BLAST. Query sequence represents *C. sordida*, Sbjct represents the known sequence listed above each comparison. The lines between each base represent homologous sites.

gb|U72977|SFU72977 Spodoptera frugiperda strain rice haplotype 1
cytochrome oxidase subunit I gene, mitochondrial gene encoding
mitochondrial protein, partial cds
Length = 683

Minus Strand HSPs:

Score = 219 (60.5 bits), Expect = 2.5e-23, Sum P(2) = 2.5e-23
Identities = 47/51 (92%), Positives = 47/51 (92%), Strand = Minus / Plus

```
Query:   132 AATTTAATATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGAATTAAC 82
             ||||||  ||||||||||||||||||||||||||||||||||||||||| |
Sbjct:   514 AATTTATCATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGTATTACC 564
```

Score = 196 (54.2 bits), Expect = 2.5e-23, Sum P(2) = 2.5e-23
Identities = 44/50 (88%), Positives = 44/50 (88%), Strand = Minus / Plus

```
Query:    51 CCAGTTTTAGCTGGAGCTATTACTATACTTTTAACTGATCGAAATTTAAA 2
             || |||||||||||||||||||||| |  | ||||||||||| ||||
Sbjct:   589 CCTGTTTTAGCTGGAGCTATTACTATATTACTTACTGATCGAAATCTAAA 638
```

gb|AF003258|AF003258 Glossiphonia complanata cytochrome c oxidase

subunit I (COI) gene, partial cds
Length = 651

Minus Strand HSPs:

Score = 187 (51.7 bits), Expect = 1.7e-16, Sum P(2) = 1.7e-16
Identities = 43/50 (86%), Positives = 43/50 (86%), Strand = Minus / Plus

```
Query:    51 CCAGTTTTAGCTGGAGCTATTACTATACTTTTAACTGATCGAAATTTAAA 2
             || || ||||| | ||||||||||||| | |||||||||||||| |||||
Sbjct:   550 CCTGTATTAGCAGCAGCTATTACTATATTATTAACTGATCGAAACTTAAA 599
```

Score = 145 (40.1 bits), Expect = 1.7e-16, Sum P(2) = 1.7e-16
Identities = 37/47 (78%), Positives = 37/47 (78%), Strand = Minus / Plus

```
Query:   130 TTTAATATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGAATTA 84
             | ||| ||| || | |||||||||||||||||||||||    ||||
Sbjct:   477 TATAACATTAGAACGTTTACCTTTATTTATTTGAGCTGTATTTATTA 523
```

gb|AF015914|AF015914 Galapaganus galapagoensis cytochrome c oxidase
subunit 1 gene, partial cds
Length = 1233

Minus Strand HSPs:

Score = 221 (61.1 bits), Expect = 1.3e-16, Sum P(2) = 1.3e-16
Identities = 53/64 (82%), Positives = 53/64 (82%), Strand = Minus / Plus

```
Query:    65 TACTCTCACTCTATCCAGTTTTAGCTGGAGCTATTACTATACTTTTAACTGATCGAAATT 6
             |||| | |   | ||||||||||||| ||||||||||||||||||| |||||||||
Sbjct:   343 TACTTTTATCTTTACCAGTTTTAGCTGGTGCTATTACTATACTTTTAACAGATCGAAATA 402

Query:     5 TAAA 2
             | ||
Sbjct:   403 TCAA 406
```

Score = 117 (32.3 bits), Expect = 1.3e-16, Sum P(2) = 1.3e-16
Identities = 33/45 (73%), Positives = 33/45 (73%), Strand = Minus / Plus

```
Query:   128 TAATATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGAATTA 84
             ||| |  ||| |  |||||||||||| | ||||| || | ||| |
Sbjct:   286 TAACACCTGACCGTATACCTTTATTTGTATGAGCAGTTGAAATCA 330
```

Figure 27) Results of less contrained sequence comparisons of *C. piatrix*, using the M13/pUC forward primer, performed by BLAST. Query sequence represents *C. piatrix*, Sbjct represents the known sequence listed above each comparison. The lines between each base represent homologous sites.

gb|J01415|HUMMTCG Human mitochondrion, complete genome
Length = 16,569

Plus Strand HSPs:

Score = 329 (90.9 bits), Expect = 5.0e-20, P = 5.0e-20
Identities = 141/235 (60%), Positives = 141/235 (60%), Strand = Plus / Plus

```
Query:    37 TTCAATTGTAGAAAATGGAGCAGGAACTGGATGAACAGTATCACCCCCTCTTTCTTCCAA 96
             | | || || ||    |||||||||||| || |||||||| |  || ||  |  |    ||
Sbjct:  6248 TGCTATAGTGGAGGCCGGAGCAGGAACAGGTTGAACAGTCTACCCTCCCTTAGCAGGGAA 6307

Query:    97 TATTCGTCATAGAGGTAGTTCAGTAGATTTAGCTATTTTTTCATTACATTTAGCTGGAAT 156
                 ||    ||    || |||||  || | || || || |||||  |||| ||  |
Sbjct:  6308 CTACTCCCACCCTGGAGCCTCCGTAGACCTAACCATCTTCTCCTTACACCTAGCAGGTGT 6367

Query:   157 TTCTTCAATTTTAGGAGCTATTAATTTTATTACCACAATTATTAATATAGCATTAAATAA 216
              || || || ||||| || || ||||| || || |||||||| ||||||  |      |
Sbjct:  6368 CTCCTCTATCTTAGGGGCCATCAATTTCATCACAACAATTATCAATATAAAACCCCCTGC 6427

Query:   217 CTTAATATTTGATCAAATACCTTTATTTATTTGAGCTGTAGGAATTACAGCTTTC 271
             | |||       | ||||  ||  | ||  | ||| | ||   ||| |||||  ||
Sbjct:  6428 CATAACCCAATACCAAACGCCCCTCTTCGTCTGATCCGTCCTAATCACAGCAGTC 6482
```

gb|U09810|KTU09810 Katharina tunicata mitochondrial genome.
Length = 15,532

Plus Strand HSPs:

Score = 541 (149.5 bits), Expect = 1.2e-39, P = 1.2e-39
Identities = 161/227 (70%), Positives = 161/227 (70%), Strand = Plus / Plus

```
Query:    44 GTAGAAAATGGAGCAGGAACTGGATGAACAGTATCACCCCCTCTTTCTTCCAATATTCGT 103
             || || |  || || || ||||| ||||| || |  || ||| |  |    ||| |  |
Sbjct:   343 GTTGAGAGGGGGGCTGGGACTGGGTGAACTGTGTATCCTCCTTTGGCGGGGAATGTGGGG 402

Query:   104 CATAGAGGTAGTTCAGTAGATTTAGCTATTTTTTCATTACATTTAGCTGGAATTTCTTCA 163
             |||   ||| | || || || ||||||||||||||| |||||||||||||| | || ||
Sbjct:   403 CATGCTGGTGGATCTGTTGACTTAGCTATTTTTTCTTTACATTTAGCTGGAGTATCGTCT 462

Query:   164 ATTTTAGGAGCTATTAATTTTATTACCACAATTATTAATATAGCATTAAATAACTTAATA 223
             |||||||| ||| |||||||||||||| |||||| | ||||||  |  | |     ||  |
Sbjct:   463 ATTTTAGGGGCTGTTAATTTTATTACTACAATTGTAAATATACGAAGAGAAGGGATACAA 522

Query:   224 TTTGATCAAATACCTTTATTTATTTGAGCTGTAGGAATTACAGCTTT 270
             || || | | ||||||| ||| ||||  |||||  |||||| ||| |
Sbjct:   523 TTAGAGCGATTACCTTTGTTTGTTTGGTCTGTAAAAATTACTGCTAT 569
```

emb|X55026|MTPACG P.anserina complete mitochondrial genome
gb|M61734|PANMTPACGA Podospora anserina, mitochondrial complete genome, races (s+) and (A+).
Length = 100,314

Plus Strand HSPs:

Score = 158 (43.7 bits), Expect = 0.00018, Sum P(2) = 0.00018
Identities = 54/82 (65%), Positives = 54/82 (65%), Strand = Plus / Plus

```
Query:   102 GTCATAGAGGTAGTTCAGTAGATTTAGCTATTTTTTCATTACATTTAGCTGGAATTTCTT 161
             |||||||  ||    |   || |||||||||||  |||  | |||||  ||  |  ||   |
Sbjct: 47975 GTCATAGTGGACCTAGTGTGGATTTAGCTATATTTGCTTTACACCTATCAGGGGTAAGTA 48034

Query:   162 CAATTTTAGGAGCTATTAATTT 183
                |  |||| ||||| |||||
Sbjct: 48035 GTTTACTAGGGGCTATGAATTT 48056
```

The results from the database searches show that the PCR primers 78X-1 and 78X-2 successfully amplified the 500 bp CO-I gene region of *C. piatrix*, *C, grynea*, and *C. sordida*. They also demonstrate that these gene regions are homologous to a wide variety of species.

VII. *Analysis of the DNA Sequence Data from Catocala Species*

Before the DNA sequence data can be analyzed and used for estimating the phylogeny among *Catocala* species, it is necessary to sequence the entire 500 bp region from both of the majority (heavy) and minority (light) strands. Although this has not yet been accomplished, the sequence data from *C. piatrix* and *C. grynea* were entered into the program MacClade, version 3.04 (Maddison and Maddison, 1992). Similar to other tree constructing software packages such as PHYLIP (Felsenstein 1991) and PAUP (Swofford 1991), this software is used to analyze character data and for constructing phylogenetic trees. Sequence data from the CO-I gene of the honeybee (*Apis mellifera ligustica*) (Crozier et. al., 1989) has also been included as a point of reference.

Figure 28 shows the data matrix generated using this software. The honeybee sequence begins with the primer binding site of 78X-2 and ends with the primer binding site of 78X-1. This region represents the amount of sequence data that is expected to be sequenced from the *Catocala* species (approximately 523 bp in length).

The sequence data was aligned by locating conserved regions within the data matrix. Alignment resulted in an indel (insertion/deletion) at site 184 of *C. grynea* which must be confirmed by sequencing from the other strand. Interestingly, however, alignment of *C. piatrix* to the honeybee sequence resulted in the loss of what appears to be two amino acid sites. Sites 83-85 and 98-100 of the honeybee sequence appear to have been deleted from that of *C. piatrix*. These results are again preliminary, as the amino acid sequence must be determined as well as the remaining sequence data for *Catocala* species.

Figure 28) Data matrix generated by MacClade, version 3.04, using Apis, *C. piatrix*, and *C. grynea* CO-I sequence data. Unsequenced bases are indicated as (?), except for sites 83-85 and 98-100, which indicate gaps. Bases which are homologous to the first sequence (Apis) are indicated as (·).

A total of six nucleotide substitutions are evident in the corresponding *C. piatrix* and *C. grynea* CO-I gene regions (sites 172-297). Although preliminary, these results demonstrate that the amount of nucleotide sequence divergence is 4.8% (95.2% homologous). Transitions occur at sites 176 and 272 which result in C -> T replacements at both sites of *C. grynea*. Transversion events occur at sites 208 (A -> T), 221 (A -> T), 288 (G -> C), and 289 (C -> G). The sequence data also shows that this region is A+T rich. The percentage of A+T content for corresponding sites (172-297) are comparable for all of the sequences: honeybee (79%), *C. grynea* (77%), and *C. piatrix* (75%).

In order to determine whether these mutations cause amino acid replacement, the corresponding sequences were entered into DNA Strider, version 1.0. This software contains capabilities that allow the amino acid sequence to be determined from DNA sequence data. The "3-phase translation" option was chosen for this search which determines the amino acid sequence in three possible open reading frames. As explained above, alignment of these sequences resulted in an indel at site 184 (see figure 28). To convert these sequences for direct amino acid comparisons, a guanine (g) was added to the DNA sequence of *C. grynea*. Removing this base from *C. piatrix* resulted in a stop codon within the amino acid sequence which is not a likely event.

Figure 29 shows the amino acid sequences (41 codons) for the three open reading frames of *C. piatrix* and *C. grynea*. Replacement at codon 39 from arginine (*C. grynea*) to alanine (*C. piatrix*) is shown in the third open reading frame. The ratio of silent to replacement substitutions for the regions compared is 95.2: 2.4. These observations are again preliminary and subject to further analysis. DNA sequence data from each species in its entirety, along with sequence data from other species of *Catocala*, will ensure a more comprehensive and less arbitrary characterization of the nucleotide and amino acid sequences. Therefore, facilitating the evolutionary histories of the members of this large genus to be estimated.

Figure 29) 3-phase amino acid translation of corresponding nucleotide sequences of *C. piatrix* and *C. grynea*, performed by DNA Strider, version 1.0. Top line represents the DNA sequence to be translated. Remaining lines represent amino acid translations in three possible open reading frames. Top two amino acid translations result in premature stop codons in the sequences. The third amino acid sequence does not contain premature stop codons and is thought to be the accurate translated sequence. Underlined codons represent replacement.

C. piatrix 126 -> 3-phase Translation

DNA sequence 126 b.p. cttccaatattc ... tagcattaaata linear

```
1    /    1                                31   /   11
ctt cca ata ttc gtc ata gag gta gtt cag tag att tag cta ttt ttt cat tac att tag
leu pro ile phe val ile glu val val gln AMB ile AMB leu phe phe his tyr ile AMB
 phe gln tyr ser ser AMB arg AMB phe ser arg phe ser tyr phe phe ile thr phe ser
  ser asn ile arg his arg gly ser ser val asp leu ala ile phe ser leu his leu ala
61   /   21                               91   /   31
ctg gaa ttt ctt caa ttt tag gag cta tta att tta tta cca caa tta tta ata tag cat
leu glu phe leu gln phe AMB glu leu leu ile leu leu pro gln leu leu ile AMB his
 trp asn phe phe asn phe arg ser tyr OCH phe tyr tyr his asn tyr OCH tyr ser ile
  gly ile ser ser ile leu gly ala ile asn phe ile thr thr ile ile asn ile ala leu
121  /   41
taa ata
OCH ile
 lys
  asn
```

C. grynea 126 -> 3-phase Translation

DNA sequence 126 b.p. cttctaatattc ... tacgattaaata linear

```
1    /    1                                31   /   11
ctt cta ata ttc gtc ata gag gta gtt cag tag att tag cta ttt ttt ctt tac att tag
leu leu ile phe val ile glu val val gln AMB ile AMB leu phe phe leu tyr ile AMB
 phe OCH tyr ser ser AMB arg AMB phe ser arg phe ser tyr phe phe phe thr phe ser
  ser asn ile arg his arg gly ser ser val asp leu ala ile phe ser leu his leu ala
61   /   21                               91   /   31
ctg gaa ttt ctt caa ttt tag gag cta tta att tta tta cta caa tta tta ata tac gat
leu glu phe leu gln phe AMB glu leu leu ile leu leu leu gln leu leu ile tyr asp
 trp asn phe phe asn phe arg ser tyr OCH phe tyr tyr tyr asn tyr OCH tyr thr ile
  gly ile ser ser ile leu gly ala ile asn phe ile thr thr ile ile asn ile arg leu
121  /   41
taa ata
OCH ile
 lys
  asn
```

DISCUSSION

Phylogenetic studies are important areas of systematic research. The evolutionary relationships (phylogeny) of organisms may serve as taxonomic markers for the proper arrangement of taxa into hierarchical categories and for the naming of new species as they are discovered. The early work of Linnaeus, coupled with Darwinian theory and the re-discovery of Mendelian genetics (in 1900), led to a synthesis of new ideas and philosophies for comparative studies (Mayr et. al., 1953).

The holarctic genus *Catocala* Schrank is an interesting group for phylogenetic study. Totaling more than 200 species, the amount of diversity that is found among sympatric populations of *Catocala* is undoubtedly the product of apostatic selection. The relative success of these moths is delicately balanced between diversity and the frequency with which individual species occur. Bark-like cryptic forewings, coupled with colorful banded and unbanded hindwings, are thought to account for their ability to evade detection and introduce anomaly (the unexpected) into the predator-prey system (Sargent, 1976). Chromatic and achromatic forms, particularly those which display similar forewings, pose many interesting questions regarding the evolution of these moths. Although controversial, sympatric speciation has been proposed by Sargent (1981) as a plausible explanation with respect to chromatic and achromatic *Catocala*. Unfortunately, however, factors which influence mating preferences, courtship behaviors, and pheromones of the underwings are unkown, as matings in captivity have met with little success (Sargent, 1976).

The majority of the literature on the underwings address functional and behavioral characteristics. Hypotheses have been proposed by Sargent (1976, 1981) regarding the evolutionary events and selective pressures that have given rise to such diversity. Multivariate analysis supported the view that anomaly has been an important factor in the evolution of the achromatic hindwing group among Juglandaceae feeding species (MacLean, 1984). Ingalls (1993) has also provided evidence to support the function of

hindwings as anomolous stimuli. The evolutionary relationships between these moths have yet to be established however. Perhaps a comprehensive phylogenetic assessment of *Catocala* species would shed new light on this subject; offering alternative hypotheses in an effort to explain the enormous amount of diversity within this genus.

Recently, the comparison of DNA sequence data has become a powerful tool in such endeavors, offering an enormous amount of data that can be analyzed and included for estimating these relationships at many different taxonomic levels. Unfortunately, however, not all sequences are appropriate for inferring the evolutionary histories of taxa. Gene trees may differ from species (true) trees if certain precautionary measures are not taken into consideration.

Appropriate gene sequences must be identified and analyzed to ensure an accurate estimation. Systematic studies using genetically determined characters should be congruent to other such studies based on different sets of characters in the same taxa (Hillis, 1987). Homoplasy due to ancestral polymorphism and multiple hits may account for alternative topologies (branching patterns) that do not corroborate previous hypotheses of phylogeny based on phenotypic characters (Brower, 1994). These characteristics lead to cladograms that are not well supported, indicating that the sequences have become randomized with respect to phylogenetic history (Hillis et. al., 1993).

Evolutionary rates and the phylogenetic utility of DNA sequences are well documented (Brower and DeSalle, 1994; Brown et. al., 1979; Felsenstein, 1988; Friedlander et. al., 1992; Halanych, 1991; Hillis and Dixon, 1991; Quicke, 1993; Simon et. al., 1994; Wolstenholme, 1992; Wolstenholme and Clary, 1985). The genes and the nucleotide regions within them vary in their phylogenetic usefulness because not all genes evolve at equal rates. Phylogenetic studies of distantly related taxa are best done using slowly evolving genes. Due to their functional significance in transcription and translation, rRNA and tRNA genes are among the slowest evolving sequences found among extant taxa. Mutations which alter their ability to synthesize protein would be a lethal event for

any organism. Due to this constraint, these regions are therefore highly conserved across all genomes and have been used for phylogenetic analyses of the deepest branches of *the tree of life*.

Rapidly evolving genes, on the other hand, are useful for comparisons of closely related (recently diverged) taxa. Mitochondrial DNA (mtDNA) has been estimated to evolve 5-10 times faster than single-copy nuclear DNA in animals, and possess several attractive characteristics for inferring such evolutionary relationships. Unlike nuclear DNA, mtDNA is usually inherited through the maternal (macrogametic) line, and therefore, not subject to independent assortment or homologous recombination. This charactersitic ensures that the integrity of the molecule is maintained from generation to generation. Any changes that may occur, would therefore, be due entirely to mutational events. MtDNA is also relatively easy to isolate and analyze because it is relatively small (approximately 16 kb) and mitochondria are abundant in most tissues.

The genes which comprise mtDNA include 2 rRNA genes, 22 tRNA genes, and approximately 13 protein coding genes. Similar to their nuclear counterparts, the rRNA and tRNA genes are highly conserved. Protein coding regions evolve at a much faster rate, however, which is thought to be due to the degenerate nature of the genetic code. Base substitutions called silent mutations, do not necessarily cause amino acid replacements within codons. Therefore, mutation rates may be high in some protein coding sequences without altering the primary structure of the amino acid sequence.

Cytochrome c oxidase, complex IV of the electron transport system, is a transmembrane protein composed of 6-13 subunits. It participates in the oxidation of NADH and $FADH_2$ to produce ATP by catalyzing the oxidation reactions of cytochrome c. Subunits I, II, and III are the largest of the proteins and are encoded by mtDNA. In comparison to other mitochondrial protein coding genes, they are the most conserved.

In this study, the mitochondrial cytochrome oxidase subunit I (CO-I) gene was chosen in an attempt to infer the phylogenetic relationships among the *Catocala* species.

Primers 78X-1 and 78X-2 were used to successfully amplify a 500 bp CO-I gene region from DNA isolated from *C. piatrix, C. grynea, C. sordida, C. cara, C. neogama,* and *Eupathenos nubilis*. PCR reactions using DNA isolated from *C. mira* failed to produce PCR products however. These results suggest that the primer binding sites of *C. mira* may be sufficiently dissimilar, thereby preventing the amplification of the CO-I gene region. These results also suggest that *E. nubilis* may share a common evolutionary history with the *Catocala*, as both genera belong to the same subfamily Catocalinae. This evidence, however, is somewhat ambiguous in that PCR reactions have the potential of being problematic. Further analysis regarding *C. mira* may provide evidence to support or refute the results found in this study.

DNA sequence data was obtained from *C. piatrix, C. grynea, and C. sordida.* Sequence analysis performed by DNA strider revealed that data from *C. sordida* was obtained from the complementary strand to that of *C. piatrix* and *C. grynea.* DNA sequence comparisons from corresponding sites of *C. piatrix* and *C. grynea* showed a total of six out of 125 sites were informative. The divergence between these two sequences was calculated to be 4.8%. Amino acid sequence comparisons between these two regions also showed replacement at one site from arginine (*C. grynea*) to alanine (*C. piatrix*). Arginine contains the charged polar side chain: $-(CH_2)_3-NH-CNH-NH_2$. Alanine, on the other hand, contains a non-polar side chain: $-CH_3$. As these results are preliminary, the effect that this substitution would have on the protein cannot be assessed accurately. This codon may not be structurally significant. Furthermore, amino acid comparisons of the entire sequence may reveal other substitutions which may complement and therefore, correct structural alterations resulting from this change.

The A+T content for corresponding sites was also calculated from honeybee (*Apis mellifera*) (79%), *C. grynea* (77%), and *C. piatrix* (75%). These results, again premature, agree with other mtDNAs examined. A+T content from Drosophila mtDNAs have been calculated to range from 74-80% (Wolstenholme and Clary, 1985). Brown et. al. (1994)

found the CO-I and CO-II A+T content among *Greya* (Lepidoptera: Prodoxidae) to be 77%.

GenBank search capabilities revealed that the sequence data from *C. piatrix, C. grynea,* and *C. sordida* was homologous to CO-I gene regions from a wide variety of organisms. These comparisons also confirmed that the sequence data from *C. sordida* was obtained from the minus strand, and therefore, complementary to the sequence data from *C. piatrix* and *C. grynea.*

Preliminary analysis of the CO-I gene sequence data from *C. sordida*, *C. piatrix* and *C. grynea* show that this region contains sites that are phylogenetically informative. DNA sequence data from each species in its entirety will determine whether the phylogenetic signal of this gene region is strong enough to minimize any homoplastic noise. These characteristics are necessary to ensure an accurate phylogenetic estimation.

Intraspecific DNA sequence comparisons may also be useful in future experiments. Such information may provide clues as to the divergence between conspecific specimens, and between male and female specimens. Future experiments which compare DNA sequence data from chromatic and achromatic species, particularly, those which share similar forewings. Although achromatic specimens have yet to be collected for this study, sequence comparisons from the CO-I gene region from such species may delineate the evolutionary relationships that exists between them.

BIBLIOGRAPHY

1. Arnett, R. H. 1993. American Insects: A Handbook of the Insects of America North of Mexico. The Sandhill Crane Press, Inc., Florida.

2. Ashlock, P. D. 1979. An Evolutionary Systematist's View of Classification. Syst. Zool. **28**: 441-450.

3. Baverstock, P. R. and C. Moritz. 1996. Project Design, 17-27. *In* D.M. Hillis, C.Moritz, B.K. Mable (eds), Molecular Systematics, second edition. Sinauer Associates, Inc., Sunderland Mass.

4. Brower, A. V. 1994. Phylogeny of *Heliconius* Butterflies Inferred from Mitochondrial DNA Sequences (Lepidoptera: Nymphalidae). Molecular Phylogenetics and Evolution. **3** (2): 159-174.

5. Brower, A. V. and R. DeSalle. 1994. Practical and Theoretical Considerations of Choice of a DNA Sequence Region in Insects Molecular Systematics, with a Short Review of Published Studies Using Nuclear Gene Regions. Ann. Entomol. Soc. Am. **87** (6): 702-716.

6. Brown, W. M., M. George, and A. C. Wilson. 1979. Rapid Evolution of Animal Mitochondrial DNA. Proc. Natl. Acad. Sci. USA. **76** (4): 1967-1971.

7. Brown, J. M., O. Pellmyr, J. N. Thompson, and R. G. Harrison. 1994. Phylogeny of *Greya* (Lepidoptera: Prodoxidae), Based on Nucleotide Sequence Variation in Mitochondrial Cytochrome Oxidase I and II: Congruence with Morphological Data. Mol. Biol. Evol. **11** (1): 128-141.

8. Brown, G. G., and M. V. Simpson. 1982. Novel Features of Animal mtDNA Evolution as shown by Sequences of two Rat Cytochrome Oxidase Subunit II Genes. Proc. Natl. Acad. Sci. USA. **79**: 3246-3250.

9. Common, I. F. B. 1990. Moths of Australia. Melbourne University Press, Melbourne, Australia.

10. Covell, C. V. 1984. A Field Guide to the Moths of Eastern North America. Houghton Mifflin Company, Boston.

11. Croze, H. 1970. Searching Image in Carrion Crows. Z. Tierpschychology, supplement 5, 1-86.

12. Crozier, R. H. 1993. Molecular methods for insect phylogenetics. 164-121. *In* J. Oakeshott and M.J. Whitten (eds), Molecular approaches to fundamental and applied entomology. Springer-Verlag New York Inc., New York.

13. Crozier, R. H., Y. C. Crozier, and A. G. Mackinlay. 1989. The CO-I and CO-II Region of Honeybee Mitochondrial DNA: Evidence for Variation in Insect Mitochondrial Evolutionary Rates. Mol. Biol. Evol. **6** (4): 399-411.

14. Donoghue, M. J., J. A. Doyle, J. Gauthier, A. G. Kluge, T. Rowe. 1989. The importance of fossils in phylogeny reconstruction. Ann. Rev. Ecol. Syst., **20**, 431-460.

15. Doran, G. H., D. N. Dickel, W. E. Ballinger Jr., O. F. Agee, P. J. Laipis, W. W. Hanswirth. 1986. Anatomical, cellular and molecular analysis of 8,000-yr-old human brain tissue from the Windover archaeological site. Nature, **323**, 803-806.

16. Dowling, T. E., C. Moritz, J. D. Palmer, and L. H. Rieseberg. 1996. Nucleic Acids III: Analysis of Fragments and Restriction Sites, 249-320. *In* D.M. Hillis, C.Moritz, B.K. Mable (eds), Molecular Sytematics, second edition. Sinauer Associates, Inc., Sunderland Mass.

17. Edmunds, M. 1974. Defence in Animals: A survey of anti-predator defences. Essex: Longman. 357.

18. Edmunds, M. 1990. The evolution of cryptic coloration. 3-16. *In* D.L. Evans and J.O. Schmidt (eds.), Insect Defenses: Adaptive mechanisms and strategies of prey and predators. State University of New York Press, Albany.

19. Evans, H. E. 1984. Insect Diversity, 31-51. *In* H. E. Evans (ed.), Insect Biology: A Textbook of Entomology. Addison-Wesley Publishing Co. Inc., Mass.

20. Felsenstein, J. 1988. Phylogenies from Molecular Sequences: Inference and Reliability. Annu. Rev. Genet. **22**: 521-565.

21. Felsenstein, J. 1991. PHYLIP (Phylogenetic Inference Package). Distributed by the author, University of Washington, Seattle.

22. Fitch, W. M. 1984. Cladistic and other methods: problems, pitfalls, and potentials. 221-252. *In* T. Duncan and T.F. Stuessy (eds), Cladistics: perspectives on the reconstruction of evolutionary history. Columbia University Press, New York.

23. Forey, P. L. 1992. Fossils and cladistic analysis, 124-136. *In* P. L. Forey, C. J. Humphries, I. J. Kitching, R.W. Scotland, D. J. Siebert, and D. M. Williams (eds.), Cladistics: A practical course in systematics. Oxford University Press Inc., New York.

24. Friedlander, T. P., J. C. Regier, and C. Mitter. 1992. Nuclear Gene Sequences for Higher Level Phylogenetic Analysis: 14 Promising Candidates. Syst. Biol. **44**: 483-490.

25. Gingerich, P. D. 1979. Paleontology, Phylogeny, and Classification: An Example from Mammalian Fossil Record. Syst. Zool. **28**: 451-464.

26. Gillespie, J. H. 1991. Protein Evolution, 3-44. *In* J. H. Gillespie (ed.), The Causes of Molecular Evolution. Oxford University Press, New York.

27. Haeckel, E. 1866. Generelle Morphologie der Organismen. G. Reimer, Berlin.

28. Halanych, K. M. 1991. 5S ribosomal RNA Sequences inappropriate for Phylogenetic Reconstruction. Mol. Biol. Evol. **8**: 249-253.

29. Hennig, W. 1966. Phylogenetic Systematics. University of Illinois Press, Urbana.

30. Hillis, D. M. 1987. Molecular Versus Morphological Approaches to Systematics. Ann. Rev. Ecol. Syst. **18**: 23-42.

31. Hillis, D. M., M. W. Allard, M. M. Miyamoto. 1993. Analysis of DNA Sequence Data: Phylogenetic Inference. Methods in Enzymology. **224**: 456-487.

32. Hillis, D. M. and M. T. Dixon. 1991. Ribosomal DNA: Molecular Evolution and Phylogenetic Inference. Quarterly Review of Biology. **66** (4): 411-453.

33. Hillis, D. M. and J. P. Huelsenbeck. 1992. Signal, Noise, and Reliability in Molecular Phylogenetic Analyses. Journal of Heredity. **83** (3): 189-195.

34. Hillis, D. M., B. K. Mable, A. Larson, S. K. Davis, and E. A. Zimmer. 1996. Nucleic Acids IV: Sequencing and Cloning, 321-381. *In* D.M. Hillis, C.Moritz, B.K. Mable (eds), Molecular Sytematics, second edition. Sinauer Associates, Inc., Sunderland Mass.

35. Hull, D. L. 1970. Contemporary systematic philosophies. Ann. Rev. Syst. Evol. **1**: 19-54.

36. Hull, D. L. 1979. Limits of Cladism. Syst. Zool. **28**: 416-440.

37. Ingalls, V. 1993. Startle and Habituation Response of Blue Jays (*Cyanocitta cristata*) in a Laboratory Simulation of Anti-Predator Defenses of *Catocala* Moths (Lepidoptera: Noctuidae). Behavior. **126** (1-2): 77-95.

38. *Kettlewell, H. B. D. 1955. Selection Experiments on Industrial Melansim in the Lepidoptera. Heredity. **9**: 323-342.

39. *Kettlewell, H. B. D. 1956. Further selection Experiments on Industrial Melansim in the Lepidoptera. Heredity. **10**: 287-301.

40. Li, W. H. and D. Graur. 1991. Molecular Phylogeny, 99-135. *In* W. H. Li and D. Graur (eds), Fundamentals in Molecular Evolution, Sinauer Associates, Inc., Sunderland, Mass.

41. Liu, H. and A. T. Beckenbach. 1992. Evolution of the Mitochondrial Cytochrome Oxidase II Gene among 10 Orders of Insects. Molecular Phylogenetics and Evolution. **1**: 41-52.

42. MacLean, D. B. 1984. Evaluation of Population Attributes and Life History Traits of *Catocala* (Lepidoptera: Noctuidae) by Means of Multivariate Analysis. Am. Midl. Nat. **112**: 67-75.

43. Maddison, W. P. and D. R. Maddison. 1992. MacClade: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Inc., Sunderland, Mass.

* unread material cited in Sargent, 1976.

44. Mayr, E., E. G. Linsley, and R. L. Usinger. 1953. Taxonomy, Its History and Functions, 3-22. *In* E. J. Boell (ed.), Methods and Principles of Systematic Zoology, McGraw-Hill Book Company Inc., New York.

45. McNeill, J. 1979. Purposeful Phenetics. Syst. Zool. **28**: 465-482.

46. Messing, J., and J. Vieira. 1982. The pUC Plasmids, an M13mp7-Derived System for Insertion Mutagenesis and Sequencing with Synthetic Universal Primers. Gene. **19**: 259-268.

47. Miller, J. S. 1991. Cladistics and Classification of Notodonitidae (Lepidoptera: Noctuidae) based on Larval and Adult Morphology. Bull. Am. Mus. Nat. Hist. **204**: 1-230.

48. Mitter, C. and E. Silverfine. 1988. On the Systematic Position of *Catocala* Schrank (Lepidoptera: Noctuidae). Systematic Entomology. **13**: 67-84.

49. Moritz, C. and D. M. Hillis. 1996. Molecular Systematics: Context and Controversies, 1-4. *In* D.M. Hillis, C.Moritz, B.K. Mable (eds), Molecular Sytematics, second edition. Sinauer Associates, Inc., Sunderland Mass.

50. New, T. R. 1996. The Diversity, Structure and Classification of Insects, 1-13. *In* David Meagher (ed.), Name That Insect: A Guide to the Insects of Southeastern Australia. Oxford University Press, Melbourne, Australia.

51. Novacek, M. J. 1992. Fossils as Critical Data for Phylogeny, 46-88. *In* Extinction and Phylogeny, M. J. Novacek and Q. D. Wheeler (eds.), Columbia University Press, New York.

52. Pääbo, S. 1985. Molecular cloning of ancient Egyptian mummy DNA. Nature, **314**, 644-645.

53. Patterson, C. 1981. The significance of fossils in determining evolutionary relationships. Ann. Rev. Ecol. Syst., **12**, 195-223.

54. Platnick, N. I. 1979. Philosophy and the transformation of cladistics. Syst. Zool., **28**: 537-546.

55. Polumbi, S. R. 1996. Nucleic Acids II: The Polymerase Chain Reaction, 205-248. *In* D.M. Hillis, C.Moritz, B.K. Mable (eds), Molecular Sytematics, second edition. Sinauer Associates, Inc., Sunderland Mass.

56. Quicke D. L. J. 1993. Principles and Techniques in Contemporary Taxonomy. Blackie Academic and Professional, Glasgow, UK.

57. Ross, H. H. 1982. Arthropoda: Insects and their Allies, 27-58. *In* H. H. Ross, C. A. Ross, and J. R. P. Ross (eds.), A Textbook of Entomology. John Wiley and Sons, Inc., New York.

58. Sargent, T. D. 1968. Cryptic Moths: Effects on Background Selections of Painting the Circumocular Scales. Science. **159**: 100-101.

59. Sargent, T. D. 1969. Behavioral Adaptations of Cryptic Moths. II. Experimental Studies on Bark-like Species. J. N. Y. Entomol. Soc. **77**: 75-79.

60. Sargent, T. D. 1973. Studies on the *Catocala* (Noctuidae) of Southern New England. IV. A Preliminary Analysis of Beak-damage Specimens, with Discussion on Anomoly as a Potential Anti-Predator Function on Hindwing Diversity. J. Lepid. Soc. **27**: 175-192.

61. Sargent, T. D. 1976. Legion of Night: The Underwing Moths. University of Massachusetts Press, Amherst, Mass.

62. Sargent, T. D. 1981. On the Achromatic *Catocala*. Journal of the Lepidopterists' Society. **35** (2): 81-93.

63. Sargent, T. D. 1990. Startle as an Anti-Predator Mechanism, with Special Referance to the Underwing Moths, (*Catocala*), 229-249. *In* D.L. Evans and J.O. Schmidt (eds.), Insect Defenses: Adaptive mechanisms and strategies of prey and predators. State University of New York Press, Albany.

64. Scotland, R. W. 1993. Cladistic theory, 3-13. *In* P.L. Forey, C. J. Humphries, I. J. Kitching, R. W. Scotland, D. J. Siebert, and D. M. Williams (ed.), Cladistics: A practical course in systematics. Oxford University Press Inc., New York.

65. Scott-Ram, N. R. 1990. Evolutionary Systematics and Theoretical Information, 41-99. *In* N. R. Scott-Ram (ed.), Transformed Cladistics, Taxonomy and Evolution. Cambridge University Press, Cambridge.

66. Selden, P. A. and J. C. Gall. 1992. A triassic mygalomorph spider from the Northern Vosges, France. Palaeontology, **35**, 211-235.

67. Simon, S., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers. Ann. Entomol. Soc. Am. **87** (6): 651-701 (1994).

68. Smith, A. B. 1994. Systematics and the fossil record: documenting evolutionary patterns. Blackwell Scientific Publications, Inc., Cambridge Mass.

69. Sokal, P. H. A. and R. R. Sneath. 1973. Numerical Taxonomy: The Principles and Practice of Numerical Classification. W. H. Freeman and Company, San Francisco.

70. Swofford, D. L. 1991. Phylogenetic Analysis Using Parsimony (PAUP). Illinois Natural History Survey, Champaign.

71. Swofford, D. L., G. J. Olsen, P. J. Waddell, D. M. Hillis. 1996. Phylogenetic Inference, 407-514. *In* D.M. Hillis, C. Moritz, B. K. Mable (eds), Molecular Sytematics, second edition. Sinauer Associates, Inc., Sunderland Mass.

72. Wake, D. B. 1996. Introduction, xvii-xxv. *In* M. J. Sanderson and L. Hufford (eds), Homoplasy: The recurrence of similarity in evolution. Academic Press, Inc., San Diego.

73. Williams, D. M. 1993. DNA analysis: theory, 89-101. *In* P. L. Forey, C. J. Humphries, I. J. Kitching, R. W. Scotland, D. J. Siebert, and D. M. Williams (ed.), Cladistics: A practical course in systematics. Oxford University Press Inc., New York.

74. Wilson, A. C., R. L. Cann, S. M. Carr, M. George, U. B. Gyllensten, K. M. Helm-Bychowski, R. G. Higuchi, S. R. Palumbi, E. M. Prager, R. D. Sage, and M. Stoneking. 1985. Mitochondrial DNA and Two Perspectives on Evolutionary Genetics. Biol. J. Linn. Soc. **26**: 375-400.

75. Wootton, R. J. 1981. Palaeozoic Insects, 319-344. *In* T. E. Mittler, F. J. Radovsky, and V. H. Resh (eds.), Annual Review of Entomology, vol. 26. Annual Reviews Inc., California.

76. Wolstenholme, D. R. 1992. Animal Mitochondrial DNA: Structure and Evolution, 173-216. *In* K. W. Jeon and D. R. Wolstenholme (eds.), Mitochondrial Genomes. Int. Rev. Cytol. **141**: 173-216.

77. Wolstenholme, D. R. and D. O. Clary. 1985. Sequence Evolution of Drosophila Mitochondrial DNA. Genetics. **109**: 725-744.