

# **Phylogenetic perspective on host plant use, colonization and speciation in butterflies**

**Elisabet Weingartner**



**Department of Zoology**  
Stockholm University  
2008

**Phylogenetic perspective on host plant use, colonization and speciation in butterflies**

Doctorial dissertation 2008

Elisabet Weingartner  
Department of Zoology,  
Stockholm University,  
SE-106 91 Stockholm,  
Sweden

© Elisabet Weingartner, Stockholm 2008

Cover illustration:

*Polygonia c-album* and plant.

Original picture from Wilkes, Benjamin, The English Moths and Butterflies ...

London, [1747-1749]

Vetenskapsakademiens boksamling på Stockholms universitetsbibliotek

ISBN 978-91-7155-705-6

Printed in Sweden by US-AB, Stockholm 2008

## ABSTRACT

Ecological factors could be a driving force in the speciation process. In butterflies, host plant use might be such a factor as there seems to exist a relationship between host plant range and butterfly diversity. According to the oscillation hypothesis, a broad host plant repertoire could enhance establishment in new areas, in which local adaptations can take place, and hence facilitate diversification within species. In this thesis we have studied three butterfly genera *Polygonia* (Nymphalidae, Nymphalini), *Pararge* (Nymphalidae, Satyrinae) and *Celastrina* (Lycaenidae: Polyommatainae). In the first paper a dated phylogeny, based on nuclear DNA (nDNA) and mitochondrial DNA (mtDNA), of *Polygonia* was constructed. We found strong conflict between the nDNA and mtDNA datasets. Possibly this can be explained by ancestral and recent hybridizations between contemporary taxa. The results point to the importance in using different markers when we try to construct the evolution of taxa. In the second paper a sister group comparison was made in order to discover whether host plant range has had an effect on species diversity. *Polygonia* includes both species that are specialists on the ancestral host plant ("urticalean rosids") as well as species with a broad host plant range. Our result indicated higher diversification rates in clades which included species with larvae feeding on different, or additional, plants compared to the "urticalean rosids" specialists. In the third paper our focus was on the colonization abilities in polyphagous butterflies. The haplotype structure of the mtDNA *cytochrome oxidase I* (COI) within the Nearctic species of *Celastrina* was analysed in a haplotype network. We also contrasted the haplotype structure in the two polyphagous and closely related species *P. c-album* (Palearctic) and *P. faunus* (Nearctic). We found little variation in *Celastrina* and the results from the analyses imply that the genus has expanded recently and rapidly. There are indications of differentiation in COI and, possibly, host plant use is involved. In *P. c-album* we also found little variation and the species seems to have expanded recently and rapidly, whereas in *P. faunus* we found structure among the haplotypes. We believe that several different haplotypes of this species have been preserved during glaciations in the Nearctic. In the fourth paper the evolution of the grassfeeding *Pararge* was analysed. The phylogeny was based on the mtDNA COI and the nDNA *wingless* (*wgl*) and times of divergences were calculated. We found a deep divergence between the European and Moroccan populations of *P. aegeria* which indicates the importance of the Mediterranean as a barrier for gene flow. The establishment of populations on the Atlantic islands might serve as examples of good colonization abilities in grass feeding butterflies.

## CONTENTS

List of papers	5
Introduction	6
What is a species?	7
The speciation process	7
Host plant use and diversification in insects	10
Introgression	11
A presentation of the butterflies	11
<i>Polygonia</i>	12
<i>Pararge</i>	13
<i>Celastrina</i>	13
Genes and methods	14
Genes	14
Analyses	15
Dating phylogenies	16
Mismatch distribution	16
Sister group comparison	17
Papers in short	17
Paper I	17
Paper II	18
Paper III	19
Paper IV	21
Conclusion	22
Acknowledgements/Tack	22
References	25

## PAPERS I-IV

The thesis is based on the following papers, which will be referred to by their roman numbers (I-IV).

- I. **Wahlberg N., Weingartner E., Warren A. D. and Nylin S.** Timing major conflict between mitochondrial and nuclear genes in species relationships of *Polygonia* butterflies (Nymphalidae: Nymphalini). *Submitted manuscript*.
- II. **Weingartner E., Wahlberg N. and Nylin S.** Dynamics of host plant use and species diversity in *Polygonia* butterflies (Nymphalidae). 2006. *Journal of Evolutionary Biology* **19**:483-491.
- III. **Weingartner E., Dalén L., Leski M., Wright D., Warren A. D., Pavulaan H. and Nylin S.** Can diversity of host use drive the diversification of phytophagous insects? Phylogeography of three butterfly taxa with wide host plant ranges (Lepidoptera: Lycaenidae and Nymphalidae). *Manuscript*.
- IV. **Weingartner E., Wahlberg N. and Nylin S.** 2006. Speciation in *Pararge* (Satyrinae: Nymphalidae) butterflies – North Africa is the source of ancestral populations of all *Pararge* species. *Systematic Entomology* **31**: 621-632.

Papers II and IV are reprints with kind permission of the original publishers, which own the copyrights:

- II: © European Society for Evolutionary Biology  
IV: © The Royal Entomological Society

## INTRODUCTION

The diversity of plants, fungi and animals is a result of the process of speciation. This process is complex and spans the research areas of evolution, genetics, systematics, ethology, ecology and biogeography. Several studies have suggested that herbivory (the ability to feed on plants) has increased diversification rates in insects (Farrell, 1998; Mitter, *et al.*, 1988). Plants are a diverse resource and offer an abundance of different niches. In many insect groups the majority are specialists on one plant family (Novotny & Basset, 2005; Scott, 1986; Ward & Spalding, 1993). However, there is, as yet, no support for a universal tendency of transitions from a generalist feeding habit towards specialization (Janz, *et al.*, 2001; Nosil, 2002; Stireman, 2005). If we do not believe that it is common to randomly shift from one plant species to another there must be a process leading to specializations on different plants. Janz *et al.* (2001) proposed that species might oscillate between phases of specialization and a more general feeding habit (Janz & Nylin, 2008; Janz, *et al.*, 2006; Nylin & Janz, 2008). The ability to feed on plants that are not included in the natural repertoire, has been confirmed in laboratory studies (Futuyma, *et al.*, 1993; Janz, *et al.*, 2001; Pratt & Ballmer, 1991; Wiklund, 1975) which support the hypothesis of host plant use as a plastic trait (Nylin & Janz, 2008). A phase in which an insect is able to feed on different plants will then work as a "platform" from which specialization to ancient or novel host plants can occur. The term "different plants" refers to host plants that are evolutionarily distantly related, e. g. that belong to different plant families or groups of families. In order to feed on a plant the insect must be able to recognize that plant and also be able to metabolize its chemical compounds. Although it cannot be ruled out that even if plants are distantly related they can still share traits important for butterfly use, similarities between different plants species is, likely, often a consequence of shared ancestry. There is an evolutionary relationship between butterfly species and larval host plant where phylogenetically related butterflies tend to use related plant species (Ehrlich & Raven, 1964; Janz & Nylin, 1998). In british longhorn beetles it has been shown that in about half of the genera, containing both specialists and generalists, the host plant of the specialists are also included in the diet of the generalists (Futuyma, 1989). The purpose of this thesis has been to gain further understanding of the relationships between host plant use, colonization and speciation in butterflies. The priority has been to present a framework in which further details about diversification in polyphagous (generalist feeding) butterflies can be studied. In order to interpret the evolution of a group of taxa a phylogenetic background is essential. The focus has been on the butterfly genera *Polygonia* (Nymphalidae: Nymphalinae), *Celastrina* (Lycaenidae: Polyommatinae) and *Pararge* (Nymphalidae: Satyrinae). The first two genera include both species specialized with larvae feeding on one or a few closely related plants as well as species with the ability to use several different plants. The last genus belongs to the subfamily Satyrinae which mainly feed on grasses. Although *Pararge* is a specialist on the grass family Poaceae, this resource is common and widespread and thus offers ample opportunities for colonizations. The starting point has been that the abilities to feed on several different plants (e.g to be polyphagous) will enhance the possibilities to colonize new areas, adapt to the local resources and, in some cases, respecialize or specialize on new host plants according to the oscillation hypothesis (Janz & Nylin, 2008; Janz, *et al.*, 2006; Nylin & Janz, 2008). This assumption have not yet been possible to test but our results are in line with the predictions. Throughout this thesis the use of "host plant" is synonymous with "larval food plant".

## *What is a species?*

In order to discuss speciation, a few words are needed to explain different views of how to define a species. The perspective of what a species is has changed through time. At the time of Linnaeus species were defined based on their appearance, e.g shared morphological characters (see Futuyma, 1998). This was then questioned, as some species apparently are polymorphic. The definition changed to include common descent. Even if two organisms differed morphologically they should be considered conspecific if they are offspring of the same parents. The most widely used definition is the Biological species concept that was put forward by Dobzhansky (1937), Muller (1942) and Mayr (1942). This concept was defined by Mayr as "species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups" (Mayr, 1942 p.120). However, this concept, although commonly used, has severe limitations. It is neither applicable to asexual organisms nor to populations separated in time (Futuyma, 1998) and fertile hybrids can in some cases be produced between obvious separate species. A well known example is the edible frog *Pelophylax kl. esculentus* which is a hybrid between *P. lessonae* and *P. ridibundus*. Other concepts have been put forward in order to adjust the definition to different perspectives (for an overview see Endler, 1989; for an overview see Futuyma, 1998; Mayr, 1966; Templeton, 1989). De Queiroz (2007) made an attempt to unify the wide range of species concepts and removed the details in the different concepts but kept what unified them. This unifying concept was defined as "a species is a separately evolving metapopulation lineage" (de Queiroz, 2007 p.881). The definitions of how to separate species, for example, the ecological species concept (Andersson, 1990), the genealogical species concept (Avice & Ball, 1990) or the genotypic cluster species concept (Mallet, 1995) then serve to diagnose a species along a continuum. Indeed, speciation is a process in which morphological and ecological characters as well as different genes evolve at different speed. This makes it hard to always be able to delimit a universal species concept.

## *The speciation process*

Speciation has fascinated people for hundreds of years. The process is central in ecological, biogeographical and, of course, evolutionary biology. It is affected by interactions both between species (gene flow) as well as within species (divergent selection, assortative mating), but also by the abilities to disperse and colonize new areas. If we understand how animals and plants have evolved and reacted to different environmental factors, such as climate change and fragmentation, it is also possible to get an idea what the consequences will be on the diversity when climate and habitats are changing in the future.

The classical explanation for the speciation process is a spatial view in which reproductive isolation is achieved either as a cause of allopatry (Mayr, 1942; Mayr, 1966) or a cause of the mechanisms that operate in sympatry (Maynard Smith, 1966). Allopatric speciation has been considered a common mode of speciation in animals and refers to when a population is separated in space. This can be due to a historical vicariant event when a barrier, such as the emergence of a mountains or a water divide, splits up an area. It could also be due to changes in the habitat, such as fragmentation caused by predation (Menge & Sutherland, 1976) or humans, in which an area becomes unfavourable in such a way that individuals in two populations will no longer be able to enter the space of each other. Dispersal to new, more and less isolated, areas by individuals from a population is also a cause for allopatric distribution that eventually can lead to diversification (Allmon, 1992). Speciation is a result of lack of gene flow, where there is no genetic exchange between populations. Given enough time, that

probably will vary from case to case, two allopatric populations will diverge to an extent that reproduction between them will no longer be possible, due to either drift (random mutations), selection (natural or sexual) or pleiotropy (multiple effects of a single gene which affects more than one phenotypic character). In a smaller population this will probably be a more rapid process compared to a larger population. If however, speciation occurs between two populations within the same geographic area, the process is called sympatric speciation. In this case, divergence is achieved as a direct result of selection, natural or sexual, as well as assortative mating (preference for a mate with a particular morphology or mating site etc.). Mayr described sympatric speciation as when "the splitting of the gene pool itself is caused by ecological factors" (Mayr, 1966 p. 451) but he doubted that it was of any real importance in the splitting of populations into new species. Sympatric speciation is still a debated subject (Felsenstein, 1981; Futuyma & Mayer, 1980) and over the years there has been a lively debate about the evidence of sympatric speciation in nature (for a review see Via, 2001). In theory new species can arise without geographic barriers via several different mechanisms; polyploidy, habitat race systems (Diehl & Bush, 1989) or selection against intermediate phenotypes (Kondrashov, *et al.*, 1998). This will divide the population into two reproductively isolated populations, although they persist in the same space. An example of a system that could be applied to sympatric speciation is the habitat race system in plant feeding insects, ("host races"). In such a system where individuals mate and feed at the same spot, a preference can, possibly, evolve for a new host plant and thus these divergent individuals may become isolated from the original individuals, as they rarely meet for mating.

However, speciation is probably not that straightforward, as to be classified as either allopatry or sympatry, and should instead be considered as a continuum where allopatry and sympatry are the ends of that continuum. Between these two processes there exist many variants where divergence for example starts in allopatry but later populations come in contact again (secondary contact) and divergence becomes even stronger because of selection against hybridization (reinforcement). Contemporary distribution pattern cannot be used without caution, as many animals are mobile and their distributions could have been different at the time of divergence. The histories of species and populations are, in temperate areas, also affected to a large extent by the climatic changes that have occurred throughout the last million years (Hewitt, 2004). Thus besides very clear cut examples, such as island speciation, geographically based speciation models might not be the best way to explain the process.

The impact of ecology on the speciation process has been brought into the light again in the last ten years. Rundle and Nosil (2005) defined ecological speciation as "the process by which barriers to gene flow evolve between populations as a result of ecologically-based divergent selection" (p. 336). Divergent selection refers to cases in which selection in two populations act in different directions. Especially ecological factors, such as environmental differences in resource use, mate recognition in different habitats (Boughman, 2001; Leal & Fleishman, 2004; Naisbit, *et al.*, 2001; Schluter & Price, 1993) and predation (Jiggins, *et al.*, 2001; Nosil, 2004; Vamosi & Schluter, 2002), have been studied. Divergent selection can occur in two populations that are allopatrically distributed as well as populations that are sympatric. Rundle and Nosil (2005) suggested that rather than using the geographically based view of speciation, we should focus on the mechanisms that inhibit gene flow between the populations and thus cause populations to evolve into different species. The reproductive isolating barriers could either be those operating before mating (pre-zygotic), after mating but before fertilization (post-mating pre-zygotic) or after fertilization (post-zygotic). I will briefly present these mechanisms which have ecological components but for a more thorough discussion, see Coyne & Orr (2004) and Rundle & Nosil (2005).



*Pre-zygotic barriers:*

Habitat isolation. This applies to the situation when two populations are separated due to divergent selection for different habitats. This situation can be applied to both allopatry and sympatry. In allopatry it requires local adaptation to the different habitats. Dobzhansky (1940) emphasized that isolating mechanisms are formed only incompletely in allopatry, and will fully develop when populations meet again in secondary contact. Mating between the previously separated populations are expected to produce inferior hybrids and thus lead to a selection against hybridizations (Marshall, et al., 2002) through assortative mating (non-random mating within a population where individuals tend to mate with individuals resembling themselves). This is called reinforcement (Blair, 1955; Dobzhansky, 1940). Reinforcement could be operating in secondary contact (Barton & Hewitt, 1985), as well as in an environmental cline with different selection in different ends of the cline (Endler, 1977) or in sympatry with polymorphism in a heterogenous environment (Maynard Smith, 1966). When two populations exist in sympatry, selection for niche divergence or strong selection in females or males for characters in the opposite sex (Lande, 1981) is required. A common example is plant feeding insects that mate and lay their eggs in or close to the host plant of the larvae (Johnson, et al., 1996), such as pea aphids *Acyrtosiphon pisum* (Hawthorne & Via, 2001). Divergent selection for different host plants can alter changes in morphology and mate recognition systems which lead to assortative mating. If the insect is dependent on the phenology of the plant, temporal isolation between populations can also be the result (see Croft, et al., 2008 regarding an example with treehoppers (Hemiptera)).

Natural selection against immigrants. This could be seen as a mechanism for reproductive isolation in itself (Rundle & Nosil, 2005). Individuals that migrate into the habitat of the other population will be poorly adapted and thus will produce few offspring.

Temporal isolation. This refers to situations such as when two populations of plant feeding insects are selected for plants that occur during separate times of the years. The mating periods of two insect populations will then never coincide thus the populations are reproductive isolated in time, but not in space.

Sexual isolation. Divergent selection for different visual and behavioural cues may cause individuals from one population to be less attractive or not recognized as mates in the other population (Boughman, 2001; Boughman, 2002). One example is the lizard *Anolis cristatellus* in which different mate signal characters have evolved in mesic and xeric habitats. The characters have evolved differently due to the light intensity and spectral quality in the two habitats (Leal & Fleishman, 2004).

*Post-zygotic barriers:*

Post-zygotic isolation. This refers to when fitness of the hybrids will be low because they are poorly adapted to both parental habitats (Coyne & Orr, 2004; Rice & Hostert, 1993; Rundle & Whitlock, 2001). Genetic incompatibilities might evolve more rapidly under divergent selection (Rundle & Nosil, 2005).

Sexual selection against hybrids. Selection against hybrids can appear when the sexual display in the hybrid is maladapted to the environment (Schluter, 2000). It could also be a consequence of divergent selection in two populations in which two phenotypes have

evolved. An intermediate phenotype of the hybrid will then be unattractive in both parental environments (Rundle & Nosil, 2005). The butterflies *Heliconius cydno* and *H. melpomene* (Nymphalidae) are flying in sympatry in parts of their distribution. They are unpalatable and warningly coloured. Both species are mimics of two other *Heliconius* butterflies. There is strong assortative mating between sympatric populations and the hybrids are, due to the intermediate phenotype being both poorly adapted and not attractive as mates for the wild type individuals (Jiggins, *et al.*, 2001; Naisbit, *et al.*, 2001).

### *Host plant use and diversification in insects*

So far, the emphasis on host plant use in insects have been on the advantages of being a specialist. However, when these hypotheses have been tested the results have been ambiguous (see Singer, 2008). Specialization can be a physiological adaptation to the chemical compounds of a plant (Dethier, 1954; Ehrlich & Raven, 1964), an adaptation to avoid predation (Jeffries & Lawton, 1984; Singer & Stireman, 2005), or to optimize information-processing (Bernays & Wcislo, 1994) in, for example search and egg laying behavior (Janz & Nylin, 1997). On the other hand, being restricted to one or a few plants also makes the species vulnerable to any changes in the availability of that plant. Abundance and distribution will probably vary according to season and climatic changes. However, laboratory experiments on butterflies (Janz, *et al.*, 2001; Pratt & Ballmer, 1991; Wiklund, 1975) and beetles (Futuyma, *et al.*, 1994) have shown that species may also be able to feed on host plants not ordinarily used by the insect, thus the potential range of host plants is probably wider than the range actually used by the insect.

The entomologist Benjamin Walsh in 1864 raised the idea (see Brues, 1924) that host-plant shifts can be factors involved in the speciation process. He noted the existence of populations that prefer different host-plants but do not seem to have any structural morphological differences. Walsh suggested that those groups were incipient species. The occurrence and mechanisms of host plants shifts have continued to fascinate during the years and is no less discussed today. For example, Carroll *et al.* (1997) reported rapid diversification in the soapberry bug (*Jadera hematoloma*) due to shifts to introduced host plants, Via (1999; Via, *et al.*, 2000) found reproductive isolation between populations of pea aphids (*Acyrhosiphon pisum*) using different host plants and the work by Wood have shown that host shift in the treehoppers *Enchenopa binotata* (Hemiptera) has resulted in temporal isolation and divergent mate recognition systems (see Cocroft, *et al.*, 2008 and references therein). The idea that polyphagy might be a path leading to adaptation to different host plants and host shifting was suggested by Ward and Spalding (1993). However, how an insect population shift from one host plant to another was not explained. Janz *et al.* (2001) came up with the theory that there might be a relationship between host range expansion and speciation. There might be a continuous process of polyphagy and monophagy where phases of specialization can change to a phase where more host plants are incorporated and then back to a phase of specialization on the ancient or a novel host plant. The ability to feed on ancestral host plants in the butterfly tribe Nymphalini was shown in laboratory feeding experiment (Janz *et al.*, 2001), thus there exists a potential to broaden the host plant range. This plasticity (Nylin & Janz, 2008) for utilizing plants used earlier in the history of the lineage could be activated in the butterflies which would then be able to rapidly broaden their host plant range if necessary. Being able to use a set of different host plants should also enhance colonization in a phase of spatial expansion, as the insect will then be released from the limitations of a particular host plant, e.g the abiotic requirements and limitation in time (Janz, *et al.*, 2006). When an insect has colonized a new area, adaptation to local host plants might lead to different scenarios. Either it

could respecialize on the ancestral host plant, specialize on a novel local host plant or continue to use a range of host plants. This process of polyphagous and monophagous transition states has been called the oscillation hypothesis (Janz & Nylin, 2008; Janz, *et al.*, 2006; Nylin & Janz, 2008). In Paper II we showed that using more or additional host plants than the ancestral ones in *Polygonia* indicated more butterfly species. This relationship was statistically significant in a study on family level of Nymphalidae (Janz, *et al.*, 2006). An expansion of host plant range was followed by an increase in the rate of net diversification although, turning this around, clades with more butterfly species did not necessarily use more host plants.

### *Introgression*

Even though two populations have reached the level of being separate species, it happens that they mate and produce offspring (hybrids). At least 6% of the *Papilio* butterfly species are involved in hybridization (Sperling, 1990). In a hybridization event, genes are transferred from one population to the other. If one gene from one population gets established in another population this is termed introgression. When species hybridize the outcome could be either a viable hybrid and thus a zone of hybrids can appear, or the hybrids will be inferior. The width of the hybrid zone will be dependent on the rate of dispersals and strength of the selection against hybrids (Barton & Hewitt, 1985).

Haldane (1922 p. 101) wrote that "[w]hen in the F1 offspring of two different animal races one sex is absent, rare, or sterile, that sex is the heterozygous sex". This has been called Haldane's rule. Although a reverse to this phenomenon was found in a study of water striders (Gerridae: Hemiptera) (Spence, 1990), the rule is common (Orr, 1997; Presgraves, 2002). The heterozygotic sex (or heterogametic sex see Coyne & Orr, 1989) is males in mammals and most insects but females in birds and butterflies. In organisms where females are the heterogametic sex this will result in sterile or inviable females.

Genes have different ability to introgress (Baack & Rieseberg, 2007) and in general genes involved with ecological, sexual or post-mating isolation are assumed to be more reluctant to introgress. Several studies have shown that mtDNA introgresses relatively easily (Bachtrog, *et al.*, 2006; see Ballard & Whitlock, 2004 and references therein). However, as the mitochondrial genes are inherited from mother to offspring it will be uncommon that mitochondrial genes will be transferred from one species to another in species where females are the heterogametic sex, which is the case for birds and butterflies, according to Haldane's rule (Sperling, 2003). Still, our interpretation of the results in Paper I is that mtDNA might have introgressed in *Polygonia* butterflies. Mitochondrial introgression has been suggested as a potential explanation for haplotype relationship between *Papilio rutulus* and *P. eurymedon* (Sperling, 1993). Presgraves (2002) showed that sterility/inviability evolves gradually in Lepidoptera, thus recently speciated species might not show Haldane's rule.

## A PRESENTATION OF THE BUTTERFLIES

I have studied three different butterfly genera; *Polygonia* (Nymphalidae, Nymphalini), *Pararge* (Nymphalidae, Satyrini) and *Celastrina* (Lycaenidae, Polyommataini). Here I will give a closer presentation of these butterflies.

## *Polygonia*

*Polygonia*, a genus within the Nymphalidae, is the subject of Papers I, II and III. The genus generally is considered to include five Palearctic species (*P. c-album*, *P. c-aureum*, *P. egea*, *P. gigantea* and *P. interposita*), seven species in North America (*P. comma*, *P. faunus*, *P. gracilis*, *P. interrogationis*, *P. oreas*, *P. progne* and *P. satyrus*) (Opler & Warren, 2002) and two species in Mexico (*P. g-argenteum* and *P. haroldii*). Although a few phylogenetic analysis have been made (Nylin, *et al.*, 2001; Wahlberg, *et al.*, 2005; Wahlberg & Nylin, 2003) species status of some of the taxa is still uncertain. In Nylin *et al.* (2001) and Wahlberg & Nylin (2003) the phylogenies were constructed based on both morphological and molecular data (mtDNA and nDNA, total of 1074 and 2926 basepairs respectively) whereas the study in Wahlberg *et al.* (2005) was based on molecular data only (mtDNA and nDNA, total 2930 basepairs).

The species generally called *Kaniska canace* is sometimes considered to belong in the genus *Polygonia*. In the phylogenies where morphological characters were included *K. canace* grouped together with *Polygonia* (Nylin, *et al.*, 2001; Wahlberg & Nylin, 2003) but in the phylogenies which have been based on molecular markers *K. canace* is instead basal in the *Nymphalis* clade, which is the sister clade to *Polygonia* (Wahlberg, *et al.*, 2005). Another controversy is *P. interposita* that sometimes has been treated as a subspecies of *P. c-album* (e.g. Gorbunov, 2001) and sometimes as a species (e.g. Churkin, 2003; Tuzov, *et al.*, 2000). The range of *P. interposita* is probably not within the range of *P. c-album* (see Churkin, 2003). *Polygonia c-album* occur in the forest belt of the Palearctic region whereas *P. interposita* has been found in the mountain areas, possibly from Ghissar to Altai. Several individuals that were sent to us turned out instead to be *P. undina*, which also occur in these mountain areas (Tuzov, *et al.*, 2000). The latter species has been treated as a subspecies of *P. egea* but our results in paper I strongly suggest that *P. undina* should have species status. The species status of *P. gracilis zephyrus* has also been unclear. *Polygonia gracilis* is distributed in the eastern part of southern Canada and the north eastern US and *P. g. zephyrus* is a more western taxon. In these two outermost areas the species are morphologically distinguishable but in the area in between, intermediate phenotypes are flying (N. Kondla pers. comm.). In Layberry *et al.* (1998) *P. g. zephyrus* is treated as a subspecies to *P. gracilis* while in Guppy and Shepard (2001) they are treated as two separate species. These two taxa could be an example of incipient speciation due to divergent selection in allopatry or it could be an example of secondary contact after divergence in allopatry. In order to catch any genetic differences between these taxa we have provisionally designated them as separate species. Last, *P. oreas* was earlier considered a subspecies of *P. progne* (Scott, 1986), but Wahlberg *et al.* (2005) found it to be more closely related to *P. gracilis*. A phylogenetic analysis including all, currently recognized, species has not been performed earlier.

The ancestral species of *Polygonia* was most probably Palaeartic and the Nearctic was colonized in two separate events (Wahlberg, *et al.*, 2005). There are species (*P. c-album* and *P. faunus*) which are feeding on a variety of host plants from different plant families (see Table I in Paper I) as well as species (*P. c-aureum*, *P. egea*, *P. comma* and *P. interrogationis*) which are restricted to plants from the urticalean rosid clade (e.g. Urticaceae, Ulmaceae and Cannabaceae (Chase, *et al.*, 1993; Soltis, *et al.*, 2000; Sytsma, *et al.*, 2002). There are also species (*P. gracilis* and *P. zephyrus*) that use plants from the families Grossulariaceae and Ericaceae but not the urticalean rosids. The use of urticalean rosids seems to be ancestral within the Nymphalini butterfly group (Janz & Nylin, 1998), probably also for the subfamily

Nymphalinae (Nylin & Wahlberg, 2008). One of the "true" polyphagous species is the comma butterfly, *P. c-album*, a model organism in studies of interactions between butterflies and host plants (Janz & Nylin, 1997; Janz & Nylin, 1998; Nylin, 1988), phenotypic plasticity in life-history traits (Nylin, 1992; Wiklund, *et al.*, 1992), and effects of higher temperature on host plant use and distribution (Braschler & Hill, 2007; Bryant, *et al.*, 2002).

### *Pararge*

The genus *Pararge*, also belonging to Nymphalidae, is the focus in Paper IV. *Pararge* belongs to the species rich subfamily Satyrinae, including approximately 2500 species, where the majority feed on the grass family Poaceae. The radiation of the Satyrini (tribe within the Satyrinae that feed on Poaceae) appears to be correlated with the spread of grass and hence the evolution of the ability to feed on grass seems to have been involved in the diversification in this butterfly group (Peña & Wahlberg, 2008). Grass is a widespread resource and will possibly enhance establishment in new areas. Three species within *Pararge* are recognized: *P. aegeria*, *P. xiphia* and *P. xiphioides*. The speckled wood butterfly, *P. aegeria*, is widespread from the southern slopes of the Atlas mountains in North Africa, throughout Europe up to 64°N and eastwards to Turkey, Israel, Syria, Transcaucasus and Urals. Since 1976 it has also been recorded from Madeira and is now well established there (Owen, *et al.*, 1986). Two subspecies are described, *P. a. aegeria* and *P. a. tircis*, distributed in southern and northern Europe respectively. They are distinguished by the ground colour which is orange in *P. a. aegeria* but more brownish in *P. a. tircis* (Tolman & Lewington, 1997). In France there is a zone of intermediate forms which also occur in south west Britain, Greece and the island Samos (Tolman & Lewington, 1997). *Pararge aegeria* has been frequently used in studies of behaviour and territorial defense (Bergman, *et al.*, 2007; Davies, 1978; Jones, 1992; Kemp, *et al.*, 2006a; Kemp, *et al.*, 2006b; Shreeve, 1984; Shreeve, 1987; Wickman & Wiklund, 1983), life history traits (Berger, *et al.*, 2008; Gotthard, *et al.*, 2007; Gotthard, *et al.*, 1994; Gotthard, *et al.*, 2000; Karlsson & Van Dyck, 2005; Sibly, *et al.*, 1997; Stevens, 2004; Van Dyck, *et al.*, 1997; van Dyck & Wiklund, 2002), habitat finding in fragmented areas (Merckx & Van Dyck, 2007) and the impact of climate change on species distribution (Hill, *et al.*, 2006; Hill, *et al.*, 1999). *Pararge xiphia* is endemic to Madeira and *P. xiphioides* occurs on four of the Canary islands (Gran Canaria, Tenerife, La Gomera and La Palma). The three species have, most probably, diversified in an allopatric process. The phylogenetic relationships of the genus had not been studied earlier.

### *Celastrina*

In Paper III, we have also studied the genus *Celastrina* within the family Lycaenidae. This genus include approximately 25 described species in the Palaearctic and Nearctic regions as well as in Mexico and Central America. This is not a well studied group and the exact number of species is uncertain. There are nine described species in the Nearctic (Pelham, 2008). In the eastern portion of the continent, *ladon* is distributed from New England to northern Florida, *neglecta* from southeastern Canada to Florida and west to Colorado and Montana, *nigra* and *neglectamajor* in the Appalachian and Ozarks regions, *idella* along the Atlantic coast from New Jersey to Georgia (Wright & Pavulaan, 1999), and *serotina* in New England (Pavulaan & Wright, 2005). In the western portion of the continent, *echo* is found from British Columbia to Baja California and east to Colorado, and *humulus* from Colorado to Montana. *Celastrina lucia* is the most polymorphic species and is found in the northern region of the United states and throughout Canada to the arctic. In several regions two to three species occur in sympatry or parapatry. In many cases, these species are morphologically very distinct (e.g., *ladon* and

*nigra* in the Appalachian Mountains in which *ladon* is blue but *nigra* is dark brown) and also differ in host plant choice.

The genus as a whole utilizes a broad range of host plants (Ranunculaceae, Rosaceae, Fabaceae, Ericaceae, Caprifoliaceae, Oleaceae, Cannabaceae, Betulaceae, Cornaceae, Rhamnaceae and Anacardiaceae) and includes highly polyphagous species, such as the Palearctic *C. argiolus*. However, there are also several Nearctic species which are monophagous. *Celastrina nigra* uses Goat's beard *Aruncus dioicus*, *C. neglectamajor* uses bugbane *Cimicifuga racemosa*, and *C. idella* uses hollies (*Ilex*), and possibly *Itea virginica*. Many but not all *Celastrina* larvae feed on flowering parts of herbs and scrubs, but for example, *C. nigra* feeds on the leaves of Goat's beard and *C. serotina* frequently uses mite-induced cherry leaf galls. The Nearctic *Celastrina* species can be found in diverse habitats such as woodlands, mountain regions and deserts.

## GENES AND METHODS

### *Genes*

Mitochondrial markers are widely used in phylogenetic studies due to several characteristics. They are easy to amplify and there is almost a lack of recombination (at least in animals). Mitochondrial genes are haploid and are inherited from mother to offspring whereas nDNA, which is diploid, is inherited from mother and father. Thus the effective population size  $N_e$  for mtDNA is one fourth of autosomal genes (Avice, 1994). This will generally lead to alleles being fixed by drift more rapidly in the mtDNA compared to autosomes (Moore, 1995), although in populations with strong sexual selection the rate of drift will be higher in autosomes (Ballard & Whitlock, 2004). However, the mitochondrial genome is inherited as one linked gene and different parts of this genome should generally be considered not independent as the whole molecule has the same genealogy. Gene trees are not species trees (Maddison, 1997) and information from other genes are needed to reconstruct species history more correctly (Maddison, 1997), at least at inter-species levels.

In Paper I, III and IV the mtDNA *cytochrome oxidase I* (COI) was used. For Paper I sequences of mtDNA *NADH dehydrogenase I* (ND1) was included. We have also used nuclear DNA (nDNA). For Paper IV we used *wingless* (*wgl*) and for Paper I the additional genes *Elongation factor-1 alpha* (EF1- $\alpha$ ), *Glyceraldehyde 3-phosphate dehydrogenase* (GAPDH) and *ribosomal protein S5* (RpS5) were used.

Why does different genes give different trees? "[P]hylogenetic history is in more than one place at once; it is a composite of all the varied histories of all the genes, some of which might place species A next to B, others might place A next to C, etc. Just as an electron can be depicted as a cloud, we might want to view phylogeny as a diffuse cloud of gene histories" (Maddison, 1997 p. 533). A phylogeny based on sequences is just a glance at a specific moment in the genes history. Every gene has its own history, genealogy, due to drift and selection. For some genes not enough time has passed since the populations started to diverge to sort out the polymorphism that exist. For another gene there might have been occasions of hybridization between two populations in which the gene has been transferred between the populations and thus obscure the real divergence between them. Genes involved with speciation, in which there is strong selection, should be more differentiated between species than autosomes (Ting, *et al.*, 2000) as "speciation genes" are, potentially, less affected by shared polymorphism or introgression.

## *Analyses*

A phylogeny is a reconstruction of the evolutionary relationships among taxa (families, genera, species or populations). It is often constructed as a tree with branches representing time or mutations since two taxa diverged, and nodes representing a common ancestor. The phylogeny could be constructed by analyzing morphological characters or molecular sequence data. There are several schools of how to analyse the evolution, all with their pros and cons. The most common methods are Maximum Parsimony (MP), Maximum Likelihood (ML) and Bayesian Inference (BI).

In Paper I, II and IV we used a MP method to obtain a hypothesis of the evolution. In MP analyses it is assumed that evolution has occurred in a way that minimizes the amount of character changes. For evolutionary tree construction this implies that the tree with least changes in character states, the minimum-length trees, is the one preferred. Finding the shortest tree is often done by a heuristic search in which trees are built by adding taxa in a random order by stepwise addition. Then parts of the tree are broken and rearrangements are done in order to see if this leads to a shorter tree. This process is repeated as many times as needed and the shortest trees are kept. This method will not try all possible tree combinations as this would take an excessive amount of time. To evaluate the strength of the hypothesis (the tree), support values can be calculated. With Bootstrap support values (Felsenstein, 1985a) (which was used in Paper IV) characters are sampled and a new matrix is constructed with the same amount of taxa and characters but in which some characters may have been sampled several times while others may not be included. This is replicated several times (Felsenstein, 2004). If a branch gets a support value of 50 then that group (leading from the particular branch) has been recovered in 50% of the replicates. For Paper I and IV we also used Bremer Support (BS) values (Bremer, 1988; Bremer, 1994). This is an index value and it shows how many more changes are needed to collapse a particular node. A low value will in this case signify low support for a clade. In datasets that include several different genes it is possible to evaluate how much each gene contributes to the support of each node. This is done using Partition Bremer support (PBS) values (Baker & DeSalle, 1997; Gatesy, *et al.*, 1999). The MP method is simple and easy to use but has some drawbacks. There is a risk that the method underestimates the amount of changes. For example, in comparing two taxa (species 1 and 2), an A at position 10, could either be due to a common ancestor with that character in that position (0 changes) or, it could be due to that species 1 had an A which was substituted by a T and then back to an A (2 changes). In rapidly evolving lineages, this could result in that long branches are interpreted as closely related, although in reality they are not. In order to take into consideration that two identical basepairs are not necessarily the same (homologous) some kind of model of molecular evolution is needed.

In Bayesian analysis the aim is to calculate the posterior distribution, that is the probability of a hypothesis given the data (the sequences) or more straightforwardly, that the tree is correct. To do this, some parameters of substitution model and nucleotide frequencies must be set. A plausible tree with tree lengths is also needed as well as a prior distribution. In order to calculate the probability a Markov Chain Monte Carlo (MCMC) method has been implemented in many programs. In this method samples are drawn randomly from the posterior distribution of hypotheses and from these samples, which are the trees, the probabilities of the nodes are calculated (Felsenstein, 2004). With Bayesian MCMC it is possible to find new optimal trees via less optimal trees in the tree space. This method was used in Paper I and IV when calculating the time of divergence between the taxa.

Instead of visualizing the relationships among individuals in a tree fashion way, it could be constructed as a network (Posada & Crandall, 2001). This approach was used in Paper I and III in order to get a better understanding of unresolved groups. A haplotype network will show if all the unresolved haplotypes in a tree is due to the fact that they are identical or due to ambiguous changes caused by homoplasies or recombinations.

### *Dating phylogenies*

In Paper I and IV times of divergence between lineages were estimated. Dating phylogenies could be performed either with the substitution rate held constant over time or with rates that vary independently among the lineages. The program BEAST v1.4.6 (Drummond & Rambaut, 2007) was used in Paper I. This program, which is a Bayesian method, estimates a tree and times of divergences at the same time. We used a model with a relaxed clock in which rates among lineages are permitted to vary in an autocorrelated manner (Drummond, *et al.*, 2006). This can be done in a way so that the evolutionary rate continuously change along the branch (Drummond, *et al.*, 2006). For Paper IV the times of divergence were analysed with the program R8s (Sanderson, 2004). In estimating times of divergencies a topology with branch lengths is required. We used a MP tree and the branch lengths were estimated with a model in a maximum likelihood framework, both with clock-like substitution rate and with rates that could vary between lineages (Sanderson, 1997), although with the constraints that the rate could vary too much across the phylogeny (Sanderson, 2002).

For calibration of nodes fossil records or ages of vicariant events can be used. These ages will constrain the time limits and can either indicate the maximum or the minimum age of a clade. Dating nodes is complicated by the fact that any error in where on a branch the calibration point is set, will be magnified through the phylogenetic tree as branches are scaled to the calibration point (Baldwin & Sanderson, 1998). Thus, all ages should be interpreted with care. Fossil records of butterflies are scarce. However, several fossils exist that are close relatives to extant species within the subfamily Nymphalinae (Wahlberg, 2006). The ages of these fossils, e. g. species belonging to the genera *Vanessa* and *Hypanartia*, were used to calibrate the phylogeny of Nymphalinae (Wahlberg, 2006). The estimated age of the split between *Aglais* and *Nymphalis-Polygonia* was used in Paper I to root the tree. For Paper IV we employed a geographic event to calibrate the phylogeny. The age of the youngest colonized Canary Island, La Palma, dated to be 2 million years old was used to calibrate the split between *Pararge xiphioides* population on that island and the sister population on another island.

### *Mismatch distribution*

Expansions in populations size, due to spatial (Excoffier, 2004; Ray, *et al.*, 2003) or demographic expansions (Excoffier, 2004; Rogers, 1995; Rogers & Harpending, 1992), could be detected by analysing the frequencies of pairwise differences, a mismatch distribution (Rogers & Harpending, 1992; Slatkin & Hudson, 1991). If the population has experienced long term demographic stability the distribution will be multimodal and ragged (Rogers & Harpending, 1992), whereas a population that has experienced a sudden expansion will have a unimodal wavelike mismatch distribution (Slatkin & Hudson, 1991). The time of the observed expansion can be calculated from the equation  $\tau = 2\mu t$  (Rogers, 1995), where  $\mu$  is the mutation rate for the sequence and  $t$  is the time since expansion. We have used this method in Paper III .



## *Sister taxa comparisons*

One of the difficulties in interpreting anything about a speciation event is to distinguish between the differences at the initial state of speciation and those added by post speciation evolution (Berlocher, 1998). Phylogenies are essential when we would like to compare species richness or find out which factors that could have been important at the speciation event. Generalizations about a pattern can however only be made when several cases of sister taxa are compared. There are several advantages with sister taxa comparisons. Sister taxa are by definition of equal age, they share an ancestor not shared by any member of another pair, thus shared evolution will eliminate some potential "noise", and the contrast between two sister groups can be assumed to be independent (Felsenstein, 1985b). However, sister group comparison cannot tell us about the direction of speciation rate change, that is, if a more diverse sister group is due to a net rate increase in this clade or rate decrease in the sisterclade (Sanderson & Donoghue, 1996). More complete phylogenetic information on related clades is necessary to resolve such issues. In Paper II we used sister taxa comparisons to see if host plant range affect species diversity in *Polygonia*.

## PAPERS IN SHORT

### *I. Timing major conflict between mitochondrial and nuclear genes in species relationships of *Polygonia* butterflies (Nymphalidae: Nymphalini)*

A phylogeny is needed in order to know the relationships between species within a genus or between genera in a family and this "pattern of relationships" can be used to apply evolutionary questions such as character state evolution, biogeography and evolution of host plant use etc. In this paper we made a phylogenetic analysis, with datings, of the genera *Polygonia*. In earlier analyses (Wahlberg & Nylin, 2003) it was noted that some positions have been incongruent between the mtDNA and nDNA datasets. Such conflicts have been found in several recent papers (Bachtrog, *et al.*, 2006; Bull, *et al.*, 2006; Chan & Levin, 2005; Kronforst, *et al.*, 2006; Linnen & Farrell, 2007; McCracken & Sorenson, 2005; Sota, 2002; Sota & Vogler, 2001). In those studies the conflicts have been explained by hybridization (ancient or recent) (Bachtrog, *et al.*, 2006; Bull, *et al.*, 2006; Chan & Levin, 2005; Kronforst, *et al.*, 2006; Linnen & Farrell, 2007; Sota, 2002), or incomplete lineage sorting (McCracken & Sorenson, 2005). By adding more nDNA and analysing these and mtDNA as separate genomes we intended to get more information about the incongruence in our study organism. Our phylogeny was also the first in which all recent known species are included, as well as several subspecies. Three taxa *P. interposita*, *P. gigantea* and *P. g-argenteum* have, to our knowledge, never before been analysed.

The phylogenetic results was mainly in accordance with (Wahlberg, *et al.*, 2005) in which morphological characters were not included. Out of fourteen nodes ten showed incongruence between the mtDNA and nDNA genomes. In the mtDNA dataset most species were monophyletic (the deviations were *P. interposita* which was nested in *P. c-album*, *P. g-argenteum* which was nested within *P. comma* and the most apical clade with *P. gracilis*, *P. zephyrus* and *P. oreas* which was unresolved). However, few species were monophyletic in the nDNA datasets. We believe that this was due to ancient polymorphism which have yet not been sorted out, but also due to recent or ongoing gene flow. The position of *K. canace* was still uncertain. Based on our results *K. canace* should not be included in *Polygonia* as it grouped together with *Nymphalis* (in both mtDNA and nDNA). As this branch was long and it

still tended to move around (when morphological characters are added see Nylin, *et al.*, 2001; Wahlberg & Nylin, 2003) we believe it should still be kept in a separate genus. The host plant use in *Kaniska* is differing as it is considered a specialist on *Smilax* (Smilacaceae), which is not used by any of the other species within *Nymphalis* and *Polygonia*. The genus *Polygonia* (without *K. canace*) was estimated to have split from *Nymphalis* for 21 million years ago (95% credibility interval of 13-29 million years ago) and began to diversify, with regard to extant species, 18 million years ago (95% credibility interval of 10-25 million years ago). *Polygonia. e. undina* was clearly differentiated from *P. egea* for all genes and should be considered as a species on its own. The most apical clade could not be resolved. Included was *P. oreas* which earlier was considered a subspecies to *P. progne* (Scott, 1984; Scott, 1986). This relationship could not be detected in the phylogeny, instead *P. oreas* grouped together with *P. gracilis* and *P. zephyrus* (and with *P. progne* in the RpS5 dataset). The latter two taxa are probably conspecific but in order to find details a more thorough study is needed. According to the records, *P. gracilis* is a specialist on Grossulariaceae whereas *P. zephyrus* uses Grossulariaceae and Ericaceae. It is interesting to note that this clade, which also includes *P. progne* and *P. haroldii*, has probably shifted hosts to these latter plant families. Host plant for *P. haroldii* is unconfirmed but possibly *Ribes*, Grossulariaceae (Krogen, 2000). The incongruence between the datasets regarding the position of *P. satyrus* could be due to an ancient introgression event. In the nDNA dataset *P. satyrus* was more related to *P. interrogationis*, *P. comma* and *P. g-argenteum*, whereas in mtDNA it was closer to the clade including *P. gracilis* and *P. oreas*. Morphologically, *P. satyrus* and *P. comma* are similar and their larvae make similar nests out of altered host plant leaves (Scott, 1986). They use host plants from the urticalean rosids, which is also used by *P. interrogationis*. The host plant of *P. g-argenteum* is unknown. These similarities support the evolutionary relationship along with the nDNA phylogeny. The distribution of *P. satyrus* is sympatric with *P. gracilis* (+ *P. zephyrus*) and we believe that an ancient introgression event can be the cause of the pattern seen in the mtDNA. If so, then this shows an example where mtDNA has crossed species barriers in a species where females are the heterogametic phase, despite Haldane's rule.

## II. Dynamics of host plant use and species diversity in *Polygonia* butterflies (*Nymphalidae*)

In this paper we applied the theory of the oscillation hypothesis (Janz & Nylin, 2008; Janz, *et al.*, 2006; Nylin & Janz, 2008) on a phylogeny of the genus *Polygonia*. In this theory it is suggested that the phase of a broad host plant range could enhance establishment in new areas and thus more opportunities for diversification by allopatry. If this is so, then there should be an asymmetric pattern in the phylogeny where clades including species with a broadened host plant range should be more species-rich than the sister clade including only specialists on ancestral host plants.

The main purpose of this paper was to find out if there were any differences in number of species between clades including species that utilize the ancestral host plants, the "urticalean rosids", compared to their sister clade, including species with a broadened host plants repertoire or species that have shifted to novel plants. At the time this paper was produced we did not have access to *P. gigantea*, *P. interposita*, *P. e. undina* or *P. g-argenteum* which we have today. The comparisons were made on a preliminary phylogeny based on mtDNA COI as well as nDNA EF1- $\alpha$  and wgl. However, the results in Paper I will not change our conclusion in this paper. We searched for clades with "urticalean rosids" specialists and compared species number with their sister clade. The latter clade sometimes included other "urticalean rosids" specialists, which were used in another comparisons, hence the contrasts

were not phylogenetically or statistically independent. The comparisons were done by hand. No statistics could be done as the dataset was too small.

In the preferred tree, four comparisons were found, all showed that species that use other or additional host plants than the ancestral urticalean rosids, are more species-rich than their sister clades that exclusively use the ancestral host plants. However, there are two interpretations for the host plant evolution in our phylogeny (Fig. 3). If we consider the butterflies using the ancestral host plants, the urticalean rosids, as being specialists, then the most parsimonious explanation must be that host plant range have been broadened at two occasions (in the ancestor of *P. faunus*+*P. c-album* and in the ancestor of *P. haroldii*+*P. progne*+*P. oreas*+*P. zephyrus*+*P. gracilis*). However, if we have a look outside the *Polygonia* butterflies we find that the closest relative, the genus *Nymphalis*, beside the "urticalean rosids" also use some of the plant families (e.g Betulaceae and Salicaceae) that reappear in the clade *P. faunus*+*P. c-album* and *P. progne*+*P. haroldii*+*P. gracilis*+*P. oreas*+*P. zephyrus*. Also, the plant families Grossulariaceae and Ericaceae are important parts of the repertoire for both of the clades within *Polygonia* that have broadened their host plant range. If this indicates that host plant use is a plastic trait, then the interpretation of the phylogeny must be that the first broadening of host plant range actually happened in the common ancestor of *Polygonia* and *Nymphalis*, followed by respecializations back to the "urticalean rosids" several times as well as further broadening of the range on an early *Polygonia* node. This ability to rapidly broaden the host plant range could have enhanced colonization of new areas as well as the persistence in old areas under change.

Even if the comparisons of species richness are too few to be reliable, our results show a pattern that we must take into consideration. In no case we found the opposite pattern. The phylogeny of *Polygonia* shows that there have been two colonization events to the Nearctic region. In one case, the clade *P. interrogationis*+*P. comma*+*P. satyrus*+*P. haroldi*+*P. progne*+*P. oreas*+*P. gracilis*+*P. zephyrus*, net diversification rate seems to have accelerated. We cannot decide if the Nearctic was colonized by an "urticalean rosids" specialist or a polyphagous population.

### *III. Can diversity of host use drive the diversification of phytophagous insects? Phylogeography of three butterfly taxa with wide host plant ranges (Lepidoptera: Lycaenidae and Nymphalidae)*

In this paper we compared the haplotype structure in COI (mtDNA) of *Polygonia c-album*, *P. faunus* and the Nearctic *Celastrina* species. All are able to use several different species as larval host plants. The ability to use different plants as food, as well as the ability to use both herbs and trees, as in the case of *Polygonia*, would be an advantage in a new area where climate and local habitat might be different from the original area of distribution. An expansion phase, spatial or demographic, can be detected by molecular and phylogeographical methods. However, an expansion phase is not evidence that the use of different host plants per se is the explanation. Monophagous butterfly species might also show an expansion phase given that the one host plant is distributed and suitable as a food plant in the area of expansion. A recent expansion would however indicate a good ability to colonize new areas and host plant is then, most probably, one important characteristic. Our aim is to determine if the observed patterns of haplotype structure are consistent with the idea that oscillations in host plant range have acted as a driver of diversification in phytophagous insects. Besides that these butterflies are distributed in different continents (*Celastrina* and *P. faunus* in the Nearctic versus *P. c-album* in the Palearctic), this study also offer example of sedentary

(*Celastrina*) versus more mobile (*P. c-album* and *P. faunus*) butterfly species as well as flower feeders (*Celastrina*), and thus constrained in flowering time, versus tree feeders (*P. c-album*).

*Polygonia c-album* and *P. faunus* are the two most polyphagous species in the genus *Polygonia*, using host plants from seven different plant families (Paper II). They are closely related (Paper I), but distributed on different continents. Thus, although of equal age (see Paper I regarding the position of *P. interposita*) and similar feeding abilities, the history of their distribution areas could have shaped the butterfly species history in different ways.

In order to establish that all the Nearctic *Celastrina* species have evolved from the same lineage we first constructed a phylogeny based on mtDNA COI. The Palearctic species *C. argiolus* and *C. ladonides* were also included. *Celastrina argiolus* is a widespread species in Europe and North Africa and considered to be closely related to the Nearctic species. Our phylogeny showed that all Nearctic species have evolved from the same ancestor but the whole clade is unresolved, thus we cannot draw any conclusion about the interspecific relationships within the clade.

The haplotype structure of the mtDNA COI indicated recent and rapid expansion for the Nearctic *Celastrina* species and the Palearctic *P. c-album*. Both these species are distributed over a large area the nucleotide diversity was in both cases lower than for *P. faunus*. The haplotype network for *P. faunus* showed several central haplotypes, indicating that different populations survived in isolation during glaciations. The Palearctic *P. c-album* was probably more affected by the last glaciations than *P. faunus*. While *P. c-album* probably was extinct in most areas in Europe and Siberia during the last glaciations, several populations of *P. faunus* possibly survived in different refugia in North America. Other studies have found that the Rocky Mountains have been an important refugium region during cold periods (DeChaine & Martin, 2005a; DeChaine & Martin, 2005b).

According to our results, the Nearctic *Celastrina* species have expanded throughout North America during the last 14 -39000 years ago, and diversified into 9 proposed species. There is little variation in the mtDNA, with 68% of the individuals having the same haplotype. However, in two cases the haplotypes are shared between two sympatric species with different host plant use. The first case is the sympatric *C. nigra*, which is brown, and *C. ladon*, which is blue. *C. ladon* is distributed from New England to Florida and the larvae feed on *Cornus* (Cornaceae), *Prunus* (Rosaceae) and *Viburnum* (Caprifoliaceae). *Celastrina nigra* on the other hand has a distribution limited to the Appalachian and Ozarks regions, and the larvae feed on the leaves of of goatsbeard *Aruncus dioicus* (Rosaceae). The second case is *C. lucia*, distributed in the northern region of the United states and throughout Canada to the Arctic, and *C. idella*, distributed along the Atlantic coast from New Jersey to Georgia (Wright & Pavulaan, 1999). *Celastrina lucia* has larvae that are polyphagous although this sympatric population prefers *Vaccinium* (Ericaceae) while the larvae of *C. idella* feed on the flowers of the holly *Ilex* (Aquifoliaceae). Diversification could potentially be a consequence of divergent selection in host plant use for sympatric populations, or it could be the consequence of two formerly allopatric populations in which one population has specialized on a host plant. In the latter scenario the sympatric distribution today is due to secondary contact between species. More studies are needed to draw any conclusions about speciation in the group. Our results imply that the radiation has been recent, and probably occurred during the final stages of the Late Pleistocene or the early Holocene. Thus, the result could either be due to a founder event during a Late Pleistocene colonisation of the Nearctic via the Bering land bridge, or from

post-glacial expansion from a single refugium. If this result has bearing upon reality, then this species complex is one of the fastest examples of adaptive radiation.

*Polygonia c-album* seems to have radiated from a single refugium during the last 8000-36000 years. Although they are polyphagous, the populations in England and Spain are more specialized than populations from Sweden and Norway (Nylin, *et al.*, 2008). Radiation from a southern refuge would likely favour individuals that have a strategy of being more general in host plant use compared with specialists. A reasonable conclusion is that the northern populations are more recent than the southern populations, which in that case have had more time for local adaptations. However, a more important factor might be latitudinal adaptations. Several studies have shown that univoltine Swedish and Norwegian populations of *P. c-album* prefer host plants other than urticalean rosids to a greater extent than populations from England, which are bivoltine, and Spain, which are trivoltine (Nygren, *et al.*, 2006; Nylin, 1988; Nylin, *et al.*, 2000). The time stress for butterflies in areas where there can be two generations per year might select for females to oviposit on plants that yield short development times (Janz, *et al.*, 1994; Nylin, 1988; Scriber & Lederhouse, 1992).

#### *IV. Speciation in Pararge (Satyrinae: Nymphalidae) butterflies – North Africa is the source of ancestral populations of all Pararge species*

Speciation in *Pararge* has, most probably, been an allopatric process (*P. xiphia* and *P. xiphioides*, are endemic to islands). Our aim was to find out if the island species have been colonized from the European mainland or North Africa and when these colonizations occurred. Further, we wanted to find out if there is molecular support for the two subspecies in *P. aegeria*. We analysed molecular sequence data from the COI (mtDNA) and *wingless* (nDNA). In order to get an estimate of the age of the clades the nodes were dated. We used the age of the La Palma, which is the youngest Canary Island inhabited by *P. xiphioides*, to calibrate the phylogenetic tree.

*Pararge* forms a strongly supported monophyletic group were all three species are recognized as monophyletic. *Pararge xiphia* seems to have diverged first from the common ancestor. According to our analysis the age of the split between *P. xiphia* and the ancestor of *P. aegeria*+*P. xiphioides* is estimated to 5 million years. *Pararge xiphioides* and *P. aegeria* are sister species and diverged about 3 million years ago. The two subspecies *P. a. tircis* and *P. a. aegeria* were not distinguishable based on DNA haplotypes, instead our data clearly distinguish between the European specimens and those from North Africa. This split seems to have occurred about 1 million years ago.

The population of *P. aegeria* on Madeira had North African haplotypes and thus originate from there rather than Europe. *Pararge xiphioides* on the Canary Islands, probably have an African origin, as these areas are relatively close. Our result indicate that the colonization of the Canary Islands has been rapid and occurred from older to younger island or, from west to east. However, haplotypes from Gran Canaria are needed and could potentially shed some light on this scenario.

We hypothesized that the Mediterranean Sea forms a strong barrier to dispersal for *Pararge* butterflies, and has done so for at least the past 1 million years. We suggest, based on the results from the *Pararge* phylogeny, that the ancestor of the genus colonized North Africa from the Palearctic when the Mediterranean Sea was desiccated (5.57-4.93 million years ago).

The populations in Europe might well have been extinct during cold periods, but recolonized from North Africa less than 1 million years ago. We suggest that the Mediterranean has been, and still is, a barrier to gene flow, although the distance between the continents is only 14 km at the Strait of Gibraltar (Dobson & Wright, 2000). We have not been able to sequence genes from individuals distributed in the south of Spain, thus we can at this point not conclude that the barrier is the Mediterranean and not in fact in Spain. Although butterflies are flying organisms and "should" be able to cross the Mediterranean, the same pattern has also been found in the Nymphalidae butterflies *Melitaea cinxia* (Wahlberg & Saccheri, 2007), in various species of *Coenonympha* (Kodandaramaiah & Wahlberg, accepted pending revision) and *Melanargia galathea* (Habel, *et al.*, 2008). However, in the latter case, colonization from North Africa to Europe seem to have been a much later event.

The ability to feed on grass might have been a factor enhancing the colonization of the Atlantic islands. Of the ten endemic butterfly species of Canary Islands five are satyrines and one is a hesperid also feeding on grass. The, probably late, colonization of *P. aegeria* in Madeira and successful establishment might also indicate good ability to settle down. However, if this is due to the grass feeding ability or not, we cannot conclude.

## CONCLUSION

This thesis has aimed at getting a deeper understanding of the relationship between host plant use and speciation in the butterfly genera *Polygonia*, *Pararge* and *Celastrina*. This work is a framework in which further studies can be performed.

The main focus have been on the influence of host plant range and speciation. This work has also given support to the oscillation hypothesis in that host plant range is important in the speciation process. Broadening the host plant range by adding host plants to the repertoire seems to be involved with higher net diversification rate in *Polygonia*. However, it does not need to be the host plant per se that enhance diversification. It could as well be that the potential to disperse, colonize and get established in new areas will be larger if a species or population is able to expand its host plant range. The evolution of the temperate biota has been influenced by the climatic changes during the last million years (Hewitt, 2004). Being able to use several different host plants will likely also improve the chances to survive during unfavourable conditions. In such situations, the possibilities to survive are to a large extent dependent on the biogeographic conditions, such as the availability of refugia and the possibilities to migrate. The Nearctic *Celastrina* species offer a good opportunity to study local adaptation in host plant use, as well as other ecological factors, in recently diverged polyphagous butterflies. The sympatric populations of *C. idella* and *C. lucia* in New Jersey could potentially be especially interesting in this respect.

## ACKNOWLEDGEMENTS/TACK

Först av alla vill jag tacka Sören och Niklas. Sören har alltid tagit sig tid att svara på frågor, diskutera problem som jag funderat över, såväl angående forskning som annat, och lugnat mig när jag varit övertygad om att det absolut inte varit lugnt. Niklas har lärt mig nästan allt jag vet om molekylärt arbete och hur man spänner fjärilar som varit frysta. Berit, Anette och Siv – tack för all hjälp med diverse ärenden som måste fixas och som ofta fixas i senaste minuten. Trots sena (och bortglömda) rapporter som ska lämnas in om föräldraledighet och annat så är ni ändå alltid trevliga och hjälpsamma och än har jag inte stött på något som inte gått att lösa.

Ni är värda mycket! Anette har även introducerat mig i persisk matlagning och en helt ny kulinarisk värld har öppnats. Tack, Ulf för hjälp men datorn när det har strulat eller när jag fått konstiga mail.

Jag vill tacka alla dem som jag delat rum med genom åren för mycket trevligt sällskap – Johanna (en kort period då vi båda var helt vilsna), Minna, Bodil, Dan, Anna, Love, Helena och Thomas. Minna och Bodil blev de som gav mig den första inblicken i hur det fungerar på zoologen. Ett särskilt varmt tack till Minna för alla pratstunder och allt stöd jag fått av dig alltsedan jag kom hit. Man känner sig alltid lite bättre efter att ha pratat med dig. Du är en ängel. När jag flyttade ihop med Anna och Love blev vi ett ”molekylär-arbetarrum”. Att sitta tillsammans med er har varit jättekul och det kändes verkligen tomt och tråkigt när jag blev kvar ensam, även om utbytet av erfarenheter och hjälp med molekylära frågor nog blev mer frekvent efter att Anna flyttat till Malmö och Love till England. Extra tack till Love som ”hållit mig i handen” när jag gav mig in i fylogeografien och delat med sig av goda råd och entusiasm! När jag kom tillbaka från den senaste föräldraledigheten hade Helena och Thomas flyttat in. Tack till er för mycket trevligt sällskap!

Det finns många som förtjänar ett tack för de här åren på zootis, och försöker jag nämna alla kommer jag ofelbart att glömma hälften. Men jag vill nämna några: Anders B, Ullasa, Rei, Jossan, Karin E, Karin N, Jessica, Anna Palmé och Mariannne för hjälp på lab och/eller för alla diskussioner och engagemang kring forskning och andra viktiga frågor. Niklas J, som har läst, reflekterat och föreslagit förbättringar till några av mina manus, Bertil, som alltid omsorgsfullt läser och kommenterar alla manus, Kalle, som visade mig hur man letar upp och fångar kvickgräsfjärilar på Madeira och Bengt som anställde mig som assistent på faunistikursen och som tagit lite extra väl hand om oss ibland.

I also want to thank my coauthors David Wright, Andy Warren, Mike Leski and Harry Pavulaan who have provided me with a lot of butterflies. You have also been very encouraging and generous in sharing your knowledge. I learnt a lot from you!

Malin, tack för allt stöd. För mig är du som ”Hanna från Arlööv”, och du och Micke har hjälpt mig förstå den akademiska världen lite bättre.

Jag har gjort min undervisningsdel hos Magnus och Erland. Det har varit jättekul, lärorikt och trevligt att få vara assistent till er. Erland, du finns med på listan över personer som inte gärna går i närkontakt med kor, där jag själv också är med (men vi kan ju alltid ringa Stefan eller kalla på Fredrik) och på listan över personer som betytt mycket för mig under de här åren. Dessutom är du med på listan över personer som jag haft knasigt kul ihop med. Stort tack till de andra faunisterna: Rasmus, Helena S, Tobias, Magne, Fredrik och Martin. Det har varit väldigt kul. Jag hoppas att det kommer bli många fler strapatser till alla möjliga ställen.

Tack även till Clas-Ove Strandberg på raritetsbiblioteket (SU) som fixade och trixade i sista sekunden med att ta fram bilden till framsidan.

De här åren har varit de kanske mest omtumlande åren i mitt liv. Jag vill tacka alla som funnits utanför universitetet och som stöttat på det psykologiska eller praktiska planet (läs: hämtat, lämnat, passat barn...): Annika, Thomas, Annelie, Marianne, Ralph, Leni, Anna, Tobias, Cattis, Marre, Espen med familj, Lena och Carl.

De näst sista jag vill nämna - och som haft stor betydelse för att mitt arbete blev klart - är

L-G, mamma och Lasse. Tusen och åter tusen tack för att ni alltid ställer upp! Ni är helt ovärderliga. Det saknas en person som, näst efter mig, blev gladast av alla när jag kom in på forskarutbildningen. Jag önskar att du var med nu, pappa – det blev klart!

Till sist - Erik – du vet hur mycket jag uppskattar vad du gjort för mig, allt stöd och alla terapier du gett, om och om igen. Tack för att ni finns hos mig, Erik, Viktor och Sara.



## REFERENCES

- Allmon, W. D. 1992. A causal analysis of stages in allopatric speciation. *Oxford Surveys in Evolutionary Biology* **8**: 219-257.
- Andersson, L. 1990. The Driving Force: Species Concepts and Ecology. *Taxon* **39**: 375-382.
- Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman & Hall, New York, USA.
- Avise, J. C. & Ball, R. M., Jr. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology* **7**: 45-67.
- Baack, E. J. & Rieseberg, L. H. 2007. A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development Genomes and evolution* **17**: 513-518.
- Bachtrog, D., Thornton, K., Clark, A. & Andolfatto, P. 2006. Extensive introgression of mitochondrial DNA relative to nuclear genes in the *Drosophila yakuba* species group. *Evolution* **60**: 292-302.
- Baker, R. H. & DeSalle, R. 1997. Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Systematic Biology* **46**: 654-673.
- Baldwin, B. G. & Sanderson, M. J. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences of the United States of America* **95**: 9402-9406.
- Ballard, J. W. O. & Whitlock, M. C. 2004. The incomplete natural history of mitochondria. *Molecular Ecology* **13**: 729-744.
- Barton, N. H. & Hewitt, G. M. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**: 113-148.
- Berger, D., Walters, R. & Gotthard, K. 2008. What limits insect fecundity? Body size- and temperature-dependent egg maturation and oviposition in a butterfly. *Functional Ecology* **22**: 523-529.
- Bergman, M., Gotthard, K., Berger, D., Olofsson, M., Kemp, D. J. & Wiklund, C. 2007. Mating success of resident versus non-resident males in a territorial butterfly. *Proceedings of the Royal Society B-Biological Sciences* **274**: 1659-1665.
- Berlocher, S. H. 1998. Can sympatric speciation via host or habitat shift be proven from phylogenetic and biogeographic evidence? In: *Endless forms species and speciation* (Howard, D. J. & Berlocher, S. H., eds.), Oxford University Press, USA.
- Bernays, E. A. & Wcislo, W. T. 1994. Sensory capabilities, information-processing, and resource specialization. *Quarterly Review of Biology* **69**: 187-204.
- Blair, W. F. 1955. Mating call and stage of speciation in the *Microhylla olivacea* - *M. carolinensis* complex. *Evolution* **9**: 469-480.
- Boughman, J. W. 2001. Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**: 944-948.
- Boughman, J. W. 2002. How sensory drive can promote speciation. *Trends in Ecology and Evolution* **17**: 571-577.
- Braschler, B. & Hill, J. K. 2007. Role of larval host plants in the climate-driven range expansion of the butterfly *Polygonia c-album*. *Journal of Animal Ecology* **76**: 415-423.
- Bremer, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* **42**: 795-803.
- Bremer, K. 1994. Branch support and tree stability. *Cladistics* **10**: 295-304.
- Brues, C. T. 1924. The specificity of food-plants in the evolution of phytophagous insects. *The American Naturalist* **58**: 127-144.
- Bryant, S. R., Thomas, C. D. & Bale, J. S. 2002. The influence of thermal ecology on the distribution of three nymphalid butterflies. *Journal of Applied Ecology* **39**: 43-55.

- Bull, V., Beltran, M., Jiggins, C. D., McMillan, W. O., Bermingham, E. & Mallet, J. 2006. Polyphyly and gene flow between non-sibling *Heliconius* species. *BMC Biology* **4**.
- Carroll, S. P., Dingle, H. & Klassen, S. P. 1997. Genetic differentiation of fitness-associated traits among rapidly evolving populations of the soapberry bug. *Evolution* **51**: 1182-1188.
- Chan, K. M. A. & Levin, S. A. 2005. Leaky prezygotic isolation and porous genomes: Rapid introgression of maternally inherited DNA. *Evolution* **59**: 720-729.
- Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., Duvall, M. R., Price, R. A., Hills, H. G., Qiu, Y.-L., Kron, K. A., Rettig, J. H., Conti, E., Palmer, J. D., Manhart, J. R., Sytsma, K. J., Michaels, H. J., Kress, W. J., Karol, K. G., Clark, W. D., Hedrén, M., Gaut, B. S., Jansen, R. K., Kim, K.-J., Wimpee, C. F., Smith, J. F., Furnier, G. R., Strauss, S. H., Xiang, Q.-Y., Plunkett, G. M., Soltis, P. S., Swensen, S. M., Williams, S. E., Gadek, P. A., Quinn, C. J., Eguiarte, L. E., Golenberg, E., Learn Jr., G. H., Graham, S. W., Barrett, S. C. H., Dayanandan, S. & Albert, V. A. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* **80**: 528-580.
- Churkin, S. V. 2003. Taxonomic notes on *Polygona* Hubner, [1818] (Lepidoptera, Nymphalidae) with the description of a new subspecies. *Helios (Moscow)* **4**: 132-147.
- Cocroft, R. B., Rodrigues, R. L. & Hunt, R. E. 2008. Host shifts, the evolution of communication, and speciation in *Enchenopa binotata* species complex of treehoppers. In: *Specialization, speciation and radiation: the evolutionary biology of herbivorous insects* (Tilmon, K. J., ed.), University of California Press, Berkeley, California, USA.
- Coyne, J. A. & Orr, H. A. 1989. Patterns of speciation in *Drosophila*. *Evolution* **43**: 362-381.
- Coyne, J. A. & Orr, H. A. 2004. Speciation. Sinauer Associates, Inc. Publishers Sunderland, Massachusetts, USA.
- Davies, N. B. 1978. Territorial defence in the speckled wood butterfly (*Pararge aegeria*): the resident always wins. *Animal Behaviour* **26**: 138-147.
- de Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology*. **56**: 879-886.
- DeChaine, E. G. & Martin, A. P. 2005a. Historical biogeography of two alpine butterflies in the Rocky Mountains: broad-scale concordance and local-scale discordance. *Journal of Biogeography* **32**: 1943-1956.
- DeChaine, E. G. & Martin, A. P. 2005b. Marked genetic divergence among sky island populations of *Sedum lanceolatum* (Crassulaceae) in the Rocky Mountains. *American Journal of Botany* **92**: 477-486.
- Dethier, V. G. 1954. Evolution of feeding preferences in phytophagous insects. *Evolution* **8**: 33-54.
- Diehl, S. R. & Bush, G. L. 1989. The role of habitat preference in adaptation and speciation. In: *Speciation and its consequences* (Otte, D. & Endler, J. A., eds.), Sinauer Associates Inc., Sunderland, Massachusetts, USA.
- Dobson, M. & Wright, A. 2000. Faunal relationships and zoogeographical affinities of mammals in north-west Africa. *Journal of Biogeography* **27**: 417-424.
- Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York, USA.
- Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. *The American Naturalist* **74**: 312-321.
- Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**: 699-710.
- Drummond, A. J. & Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**.

- Ehrlich, P. R. & Raven, P. H. 1964. Butterflies and plants: a study in coevolution. *Evolution* **18**: 586-608.
- Endler, J. A. 1977. Geographic variation, speciation, and clines. Princeton University Press, Princeton, New Jersey, USA.
- Endler, J. A. 1989. Conceptual and other problems in speciation. In: *Speciation and its consequences* (Otte, D. & Endler, J. A., eds.), Sinauer Associates Inc, Sunderland, Massachusetts, USA.
- Excoffier, L. 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology* **13**: 853-864.
- Farrell, B. D. 1998. "Inordinate fondness" explained: why are there so many beetles? *Science* **281**: 555-559.
- Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* **35**: 124-138.
- Felsenstein, J. 1985a. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783-791.
- Felsenstein, J. 1985b. Phylogenies and the comparative method. *The American Naturalist* **125**: 1-15.
- Felsenstein, J. 2004. Inferring phylogenies. Sinauer Associates Inc., Sunderland, Massachusetts, USA.
- Futuyma, D. J. 1989. Macroevolutionary consequences of speciation: Inferences from phytophagous insects. In: *Speciation and its consequences* (Otte, D. & Endler, J. A., eds.), Sinauer Associates, Inc, Sunderland, Massachusetts, USA.
- Futuyma, D. J. 1998. Evolutionary Biology. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- Futuyma, D. J., Herrmann, C., Milstein, S. & Keese, M. C. 1993. Apparent transgenerational effects of host plant in the leaf beetle *Ophraella notulata* (Coleoptera: Chrysomelidae). *Oecologia* **96**: 365-372.
- Futuyma, D. J. & Mayer, G. 1980. Non-allopatric speciation in animals. *Systematic Zoology* **29**: 254-271.
- Futuyma, D. J., Walsh Jr., J. S., Morton, T., Funk, D. J. & Keese, M. C. 1994. Genetic variation in a phylogenetic context: responses of two specialized leaf beetles (Coleoptera: Chrysomelidae) to host plants of their congeners. *Journal of evolutionary biology* **7**: 127-146.
- Gatesy, J., O'Grady, P. & Baker, R. H. 1999. Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. *Cladistics* **15**: 271-313.
- Gorbunov, P. Y. 2001. The butterflies of Russia: classication, genitalia, keys for identification (Lepidoptera: Hesperioidea and Papilionoidea). Russian Academy of Sciences Institute of Plant and Animal Ecology. Thesis, Ekaterinburg, Russia.
- Gotthard, K., Berger, D. & Walters, R. 2007. What keeps insects small? Time limitation during oviposition reduces the fecundity benefit of female size in a butterfly. *The American Naturalist* **169**: 768-779.
- Gotthard, K., Nylin, S. & Wiklund, C. 1994. Adaptive variation in growth rate: life history costs and consequences in the speckled wood butterfly, *Pararge aegeria*. *Oecologia* **99**: 281-289.
- Gotthard, K., Nylin, S. & Wiklund, C. 2000. Mating opportunity and the evolution of sex-specific mortality rates in a butterfly. *Oecologia* **122**: 36-43.
- Guppy, C. S. & Shepard, J. H. 2001. Butterflies of British Columbia. UBC Press in collaboration with the Royal British Columbia Museum, Canada.

- Habel, J. C., Meyer, M., El Mousadik, A. & Schmitt, T. 2008. Africa goes Europe: The complete phylogeography of the marbled white butterfly species complex *Melanargia galathea*/*M. lachesis* (Lepidoptera : Satyridae). *Organisms Diversity & Evolution* **8**: 121-129.
- Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics* **12**: 101-109.
- Hawthorne, D. J. & Via, S. 2001. Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* **412**: 904-907.
- Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society London Ser. B* **359**: 183-195.
- Hill, J. K., Hughes, C. L., Dytham, C. & Searle, J. B. 2006. Genetic diversity in butterflies: interactive effects of habitat fragmentation and climate-driven range expansion. *Biology Letters* **2**: 152-154.
- Hill, J. K., Thomas, C. D. & Huntley, B. 1999. Climate and habitat availability determine 20th century changes in a butterfly's range margin. *Proceedings of the Royal Society of London Series B Biological Sciences* **266**: 1197-1206.
- Janz, N., Nyblom, K. & Nylin, S. 2001. Evolutionary dynamics of host-plant specialization: a case study of the tribe Nymphalini. *Evolution* **55**: 783-796.
- Janz, N. & Nylin, S. 1997. The role of female search behaviour in determining host plant range in plant feeding insects: a test of the information processing hypothesis. *Proceedings of the Royal Society of London Series B Biological Sciences* **264**: 701-707.
- Janz, N. & Nylin, S. 1998. Butterflies and plants: a phylogenetic study. *Evolution* **52**: 486-502.
- Janz, N. & Nylin, S. 2008. The oscillation hypothesis of hostplant range and speciation. In: *The evolutionary biology of herbivorous insects: Specialization, speciation and radiation* (Tilmon, K. J., ed.), University of California Press, Berkeley, California, USA.
- Janz, N., Nylin, S. & Wahlberg, N. 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects. *BMC Evolutionary Biology* **6**:4.
- Janz, N., Nylin, S. & Wedell, N. 1994. Host plant utilization in the comma butterfly: sources of variation and evolutionary implications. *Oecologia* **99**: 132-140.
- Jeffries, M. J. & Lawton, J. H. 1984. Enemy free space and the structure of ecological communities. *Biological Journal of the Linnean Society* **23**: 269-286.
- Jiggins, C. D., Naisbit, R. E., Coe, R. L. & Mallet, J. 2001. Reproductive isolation caused by colour pattern mimicry. *Nature* **411**: 302-305.
- Johnson, P. A., Hoppensteadt, F. C., Smith, J. J. & Bush, G. L. 1996. Conditions for sympatric speciation: a diploid model incorporating habitat fidelity and non-habitat assortative mating. *Evolutionary Ecology* **10**: 187-205.
- Jones, M. J. 1992. The speckled wood butterflies *Pararge xiphia* and *P. aegeria* (Satyridae) on Madeira: distribution, territorial behaviour and possible competition. *Biological Journal of the Linnean Society* **46**: 77-89.
- Karlsson, B. & Van Dyck, H. 2005. Does habitat fragmentation affect temperature-related life-history traits? A laboratory test with a woodland butterfly. *Proceedings of the Royal Society B-Biological Sciences* **272**: 1257-1263.
- Kemp, D. J., Wiklund, C. & Gotthard, K. 2006a. Life history effects upon contest behaviour: Age as a predictor of territorial contest dynamics in two populations of the speckled wood butterfly, *Pararge aegeria* L. *Ethology* **112**: 471-477.
- Kemp, D. J., Wiklund, C. & van Dyck, H. 2006b. Contest behaviour in the speckled wood butterfly (*Pararge aegeria*): seasonal phenotypic plasticity and the functional significance of flight performance. *Behavioral Ecology and Sociobiology* **59**: 403-411.

- Kodandaramaiah, U. & Wahlberg, N. accepted pending revision. Biogeography of *Coenonympha* butterflies (Nymphalidae - Satyrinae) - how important has vicariance been? *Systematic Entomology*.
- Kondrashov, A. S., Yampolsky, L. Y. & Shabalina, S. A. 1998. On the sympatric origin of species by means of natural selection. In: *Endless forms species and speciation* (Howard, D. J. & Berlocher, S. H., eds.), Oxford University Press, New York Oxford, USA.
- Krogen, R. 2000. Records of *Polygonia haroldi* (Dewitz, 1877) in Sonora, Mexico (Lepidoptera, Nymphalidae). *Atalanta* **31**: 67-70.
- Kronforst, M. R., Young, L. G., Blume, L. M. & Gilbert, L. E. 2006. Multilocus analyses of admixture and introgression among hybridizing *Heliconius* butterflies. *Evolution* **60**: 1254-1268.
- Lande, R. 1981. Models of speciation by sexual selection on polygenic traits (mating preferences/sexual dimorphism/genetic correlation/runaway process). *Proceedings of the National Academy of Sciences of the United States of America* **78**: 3721-3725.
- Layberry, R. A., Hall, P. W. & Lafontaine, J. D. 1998. The Butterflies of Canada. University of Toronto Press, Toronto, Canada.
- Leal, M. & Fleishman, L. J. 2004. Differences in visual signal design and detectability between allopatric populations of *Anolis* Lizards. *The American Naturalist* **163**: 26-39.
- Linnen, C. R. & Farrell, B. D. 2007. Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera : Diprionidae) sawflies. *Evolution* **61**: 1417-1438.
- Maddison, W. P. 1997. Gene trees in species trees. *Systematic Biology* **46**: 523-536.
- Mallet, J. 1995. A species definition for the Modern Synthesis. *Trends in Ecology and Evolution* **10**: 294-299.
- Marshall, J. L., Arnold, M. L. & Howard, D. J. 2002. Reinforcement: the road not taken. *Trends in Ecology and Evolution* **17**: 558-563.
- Maynard Smith, J. 1966. Sympatric speciation. *The American Naturalist* **100**: 637-650.
- Mayr, E. 1942. Systematics and the Origin of Species. Columbia University Press, New York, USA.
- Mayr, E. 1966. Animal species and evolution. The Belknap press of Harvard University press, Cambridge, Massachusetts, USA.
- McCracken, K. G. & Sorenson, M. D. 2005. Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (*Nomonyx-Oxyura*)? *Systematic Biology* **54**: 35-55.
- Menge, B. A. & Sutherland, J. P. 1976. Species diversity gradients: synthesis of the roles of predation, competition and temporal heterogeneity. *The American Naturalist* **110**: 351-369.
- Merckx, T. & Van Dyck, H. 2007. Habitat fragmentation affects habitat-finding ability of the speckled wood butterfly, *Pararge aegeria* L. *Animal Behaviour* **74**: 1029-1037.
- Mitter, C., Farrell, B. & Wiegmann, B. 1988. The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? *The American Naturalist* **132**: 107-128.
- Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49**: 718-726.
- Muller, H. 1942. Isolating mechanisms, evolution, and temperature. *Biol Symp* **6**: 71-125.
- Naisbit, R. E., Jiggins, C. D. & Mallet, J. 2001. Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. *Proceedings of the Royal Society of London serie B* **268**: 1849-1854.
- Nosil, P. 2002. Transition rates between specialization and generalization in phytophagous insects. *Evolution* **56**: 1701-1706.

- Nosil, P. 2004. Reproductive isolation caused by visual predation on migrants between divergent environments. *Proceedings of the Royal Society of London Series B Biological Sciences*. **B 271**.
- Novotny, V. & Basset, Y. 2005. Host specificity of herbivores in tropical forests. *Proceedings of the Royal Society of London Series B Biological Sciences* **272**: 1083-1090.
- Nygren, G. H., Nylin, S. & Stefanescu, C. 2006. Genetics of host plant use and life history in the comma butterfly across Europe: varying modes of inheritance as a potential reproductive barrier. *Journal of Evolutionary Biology* **19**: 1882-1893.
- Nylin, S. 1988. Host plant specialization and seasonality in a polyphagous butterfly, *Polygonia c-album* (Nymphalidae). *Oikos* **53**: 381-386.
- Nylin, S. 1992. Seasonal plasticity in life-history traits - growth and development in *Polygonia c-album* (Lepidoptera, Nymphalidae). *Biological Journal of the Linnean Society* **47**: 301-323.
- Nylin, S., Bergström, A. & Janz, N. 2000. Butterfly host plant choice in the face of possible confusion. *Journal of Insect Behavior* **13**: 469-482.
- Nylin, S. & Janz, N. 2008. Butterfly host plant range: an example of plasticity as a promoter of speciation? *Evolutionary Ecology* **on-line early**.
- Nylin, S., Nyblom, K., Ronquist, F., Janz, N., Belicek, J. & Källersjö, M. 2001. Phylogeny of *Polygonia*, *Nymphalis* and related butterflies (Lepidoptera: Nymphalidae): a total-evidence analysis. *Zoological Journal of the Linnean Society* **132**: 441-468.
- Nylin, S., Nygren, G. H., Söderlind, L. & Stefanescu, C. 2008. Geographical variation in host plant utilization in the comma butterfly: the roles of time constraints and plant phenology. *Evolutionary Ecology* **on-line early**.
- Nylin, S. & Wahlberg, N. 2008. Does plasticity drive speciation? Host-plant shifts and diversification in nymphaline butterflies (Lepidoptera: Nymphalidae) during the tertiary. *Biological Journal of the Linnean Society* **94**: 115-130.
- Opler, P. A. & Warren, A. D. 2002. Butterflies of North America. 2. Scientific Names List for Butterfly Species of North America, North of Mexico. Gillette Museum Publications, Fort Collins, Colorado, USA.
- Orr, H. A. 1997. Haldane's rule. *Annual Review of Ecology & Systematics* **28**: 195.
- Owen, D. F., Shreeve, T. G. & Smith, A. G. 1986. Colonization of Madeira by the speckled wood butterfly, *Pararge aegeria* (Lepidoptera: Satyridae), and its impact on the endemic *Pararge xiphia*. *Ecological Entomology* **11**: 349-352.
- Pavulaan, H. & Wright, D. M. 2005. *Celastrina serotina* (Lycaenidae: Polyommatainae): a new butterfly species from the Northeastern United States and Eastern Canada. *The Taxonomic Report* **6**: 1-18.
- Pelham, J. 2008. A catalogue of the butterflies of the United States and Canada, with a complete bibliography of the descriptive and systematic literature. *The Journal of Research on the Lepidoptera* **40**.
- Peña, C. & Wahlberg, N. 2008. Prehistorical climate change increased diversification of a group of butterflies. *Biology Letters* **4**: 274-278.
- Posada, D. & Crandall, K. A. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution* **16**: 37-45.
- Pratt, G. F. & Ballmer, G. R. 1991. Acceptance of *Lotus scoparius* (Fabaceae) by larvae of Lycaenidae. *Journal of the Lepidopterists' Society* **45**: 188-196.
- Presgraves, D. C. 2002. Patterns of postzygotic isolation in Lepidoptera. *Evolution* **56**: 1168-1183.
- Ray, N., Currat, M. & Excoffier, L. 2003. Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution* **20**: 76-86.

- Rice, W. R. & Hostert, E. E. 1993. Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* **47**: 1637-1653.
- Rogers, A. R. 1995. Genetic evidence for a Pleistocene population explosion. *Evolution* **49**: 608-615.
- Rogers, A. R. & Harpending, H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**: 552-569.
- Rundle, H. D. & Nosil, P. 2005. Ecological speciation. *Ecology Letters* **8**: 336-352.
- Rundle, H. D. & Whitlock, M. C. 2001. A genetic interpretation of ecologically dependent isolation. *Evolution* **55**: 198-201.
- Sanderson, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* **14**: 1218-1231.
- Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* **19**: 101-109.
- Sanderson, M. J. 2004. R8s, Version 1.70. *Published by author*. URL <http://ginger.ucdavis.edu/r8s/>.
- Sanderson, M. J. & Donoghue, M. J. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. *Trends in Ecology and Evolution* **11**: 15-20.
- Schluter, D. 2000. The ecology of adaptive radiation. Oxford University Press, Oxford, USA.
- Schluter, D. & Price, T. 1993. Honesty, perception and population divergence in sexually selected traits. *Proceedings of the Royal Society of London Series B-Biological Sciences* **253**: 117-122.
- Scott, J. A. 1984. A review of *Polygonia progne (oreas)* and *P. gracilis (zephyrus)* (Nymphalidae), including a new subspecies from the southern Rocky Mountains. *Journal of Research on the Lepidoptera* **23**: 197-210.
- Scott, J. A. 1986. The Butterflies of North America. Stanford University Press, Stanford, USA.
- Scriber, J. & Lederhouse, R. 1992. The thermal environment as a resource dictating geographic patterns of feeding specialization of insect herbivores. In: *Effects of resource distribution on animal-plant interactions* (Hunter, M., Ohgushi, T. & Price, P., eds.), Academic Press, New York, USA.
- Shreeve, T. G. 1984. Habitat selection, mate location, and microclimatic constraints on the activity of the speckled wood butterfly *Pararge aegeria*. *Oikos* **42**: 371-377.
- Shreeve, T. G. 1987. The mate location behaviour of the male speckled wood butterfly. *Animal Behaviour* **35**: 682-690.
- Sibly, R. M., Winokur, L. & Smith, R. H. 1997. Interpopulation variation in phenotypic plasticity in the speckled wood butterfly, *Pararge aegeria*. *Oikos* **78**: 323-330.
- Singer, M. S. 2008. Evolutionary ecology of polyphagy. In: *Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects* (Tilmon, K. J., ed.), University of California Press, Berkeley, California, USA.
- Singer, M. S. & Stireman, J. O. 2005. The tri-trophic niche concept and adaptive radiation of phytophagous insects. *Ecology Letters* **8**: 1247-1255.
- Slatkin, M. & Hudson, R. R. 1991. Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. *Genetics* **129**: 555-562.
- Soltis, D. E., Soltis, P. S., Chase, M. W., Mort, M. E., Albach, D. C., Zanis, M., Savolainen, V., Hahn, W. H., Hoot, S. B., Fay, M. F., Axtell, M., Swensen, S. M., Prince, L. M., Kress, W. J., Nixon, K. C. & Farris, J. S. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* **133**: 381-461.
- Sota, T. 2002. Radiation and reticulation: extensive introgressive hybridization in the carabid beetles *Ohomopterus* inferred from mitochondrial gene genealogy. *Population Ecology* **44**: 145-156.

- Sota, T. & Vogler, A. P. 2001. Incongruence of mitochondrial and nuclear gene trees in the carabid beetles *Ohomopterus*. *Systematic Biology* **50**: 39-59.
- Spence, J. R. 1990. Introgressive hybridization in Heteroptera - the example of *Limnporus Stål* (Gerridae) species in Western Canada. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **68**: 1770-1782.
- Sperling, F. 2003. Butterfly species and molecular phylogenies. In: *Butterflies: Evolution and Ecology Taking Flight* (Boggs, C., Watt, W. & Ehrlich, P., eds.), University of Chicago Press, Chicago, USA.
- Sperling, F. A. H. 1990. Natural hybrid of *Papilio* (Insecta: Lepidoptera): poor taxonomy or interesting evolutionary problem? *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **68**: 1790-1799.
- Sperling, F. A. H. 1993. Mitochondrial-DNA variation and Haldane rule in the *Papilio glaucus* and *P. troilus* species groups. *Heredity* **71**: 227-233.
- Stevens, D. J. 2004. Pupal development temperature alters adult phenotype in the speckled wood butterfly, *Pararge aegeria*. *Journal of Thermal Biology* **29**: 205-210.
- Stireman, J. O. 2005. The evolution of generalization? Parasitoid flies and the perils of inferring host range evolution from phylogenies. *Journal of Evolutionary Biology* **18**: 325-336.
- Sytsma, K. J., Morawetz, J., Pires, J. C., Nepokroeff, M., Conti, E., Zjhra, M., Hall, J. C. & Chase, M. W. 2002. Urticalean rosids: circumscription, rosid ancestry, and phylogenetics based on *rbcL*, *trnL-F*, and *ndhF* sequences. *American Journal of Botany* **89**: 1531-1546.
- Templeton, A. R. 1989. The meaning of species and speciation: a genetic perspective. In: *Speciation and its consequences* (Otte, D. & Endler, J. A., eds.), Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- Ting, C.-T., Tsauro, S.-C. & Wu, C.-I. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proceedings of the National Academy of Sciences USA* **97**: 5313-5316.
- Tolman, T. & Lewington, R. 1997. Butterflies of Britain and Europe. Harper Collins Publishers, London, United Kingdom.
- Tuzov, V. K., Bogdanov, P. V., Churkin, S. V., Dantchenko, A. V., Devyatkin, A. L., Murzin, V. S., Samodurov, G. D. & Zhdanko, A. B. 2000. Guide to the butterflies of Russia and adjacent territories (Lepidoptera, Rhopalocera). Pensoft Publishers, Sofia, Bulgaria.
- Vamosi, S. M. & Schluter, D. 2002. Impacts of trout predation on fitness of sympatric sticklebacks and their hybrids. *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**: 923-930.
- Van Dyck, H., Matthysen, E. & Dhondt, A. A. 1997. Mate-locating strategies are related to relative body length and wing colour in the speckled wood butterfly *Pararge aegeria*. *Ecological Entomology* **22**: 116-120.
- Van Dyck, H. & Wiklund, C. 2002. Seasonal butterfly design: morphological plasticity among three developmental pathways relative to sex, flight and thermoregulation. *Journal of Evolutionary Biology* **15**: 216-225.
- Via, S. 1999. Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution* **53**: 1446-1457.
- Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology and Evolution* **16**: 381-390.
- Via, S., Bouck, A. C. & Skillman, S. 2000. Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution* **54**: 1626-1637.
- Wahlberg, N. 2006. That awkward age for butterflies: Insights from the age of the butterfly subfamily Nymphalinae (Lepidoptera : Nymphalidae). *Systematic Biology* **55**: 703-714.



- Wahlberg, N., Brower, A. V. Z. & Nylin, S. 2005. Phylogenetic relationships and historical biogeography of tribes and genera in the subfamily Nymphalinae (Lepidoptera: Nymphalidae) *Biological Journal of the Linnean Society* **86**: 227-251.
- Wahlberg, N. & Nylin, S. 2003. Morphology versus molecules: resolution of the positions of *Nymphalis*, *Polygonia*, and related genera (Lepidoptera: Nymphalidae). *Cladistics* **19**: 213-223.
- Wahlberg, N. & Saccheri, I. 2007. The effects of Pleistocene glaciations on the phylogeography of *Melitaea cinxia* (Lepidoptera: Nymphalidae). *European Journal of Entomology* **104**: 675-684.
- Ward, L. K. & Spalding, D. F. 1993. Phytophagous British insects and mites and their food-plant families: total numbers and polyphagy. *Biological Journal of the Linnean Society* **49**: 257-276.
- Wickman, P. O. & Wiklund, C. 1983. Territorial defence and its seasonal decline in the speckled wood butterfly (*Pararge aegeria*). *Animal Behaviour* **31**: 1206-1216.
- Wiklund, C. 1975. The evolutionary relationship between adult oviposition preferences and larval host plant range in *Papilio machaon* L. *Oecologia* **18**: 185-197.
- Wiklund, C., Wickman, P. O. & Nylin, S. 1992. A sex difference in the propensity to enter direct/diapause development: a result of selection for protandry. *Evolution* **46**: 519-528.
- Wright, D. M. & Pavulaan, H. 1999. *Celastrina idella* (Lycaenidae: Polyommatainae): a new butterfly species from the Atlantic coastal plain. *The Taxonomic Report* **1**: 1-11.