University of Gothenburg
Faculty of Natural Sciences
2008

UNIVERSITY OF GOTHENBURG

THE EVOLUTION OF FLAT PERIWINKLES *LITTORINA FABALIS* AND *L. OBTUSATA* EMPHASIZING MITOCHONDRIAL INTROGRESSION AND RESTRICTED RECOMBINATION

PETRI KEMPPAINEN

Doctoral Thesis

Department of Marine Ecology - Tjärnö
University of Gothenburg
Tjärnö
SE-452 96, Strömstad
Sweden

Avhandlingen för filosofie doktorsexamen i marin ekologi vid Göteborgs universitet (examinator: Per Jonsson) som enligt beslut av naturvetenskapliga fakulteten kommer att försvaras offentligt fredagen den 3:e oktober 2008 kl. 15.00 i stora föreläsningssalen i Botanhuset, Carl Skottbergsgatan 22b, Göteborg.

Fakultetsopponent: Prof. Roger Butlin, Animal and Plant Sciences, University of Sheffield, United Kingdom

Published by the Department of Marine Ecology, University of Gothenburg, Sweden

ISBN: 9189677-42-0

Printed by Intellecta DocuSys, Västra Frölunda, Göteborg

Kemppainen, Petri (2008) The evolution of flat periwinkles *Littorina fabalis* and *L. obtusata* emphasizing mitochondrial introgression and restricted recombination. Department of Marine Ecology, Gothenburg University, Sven Lovén Centre for Marine Sciences - Tjärnö. SE-45296, Strömstad, Sweden.

ABSTRACT: The evolution of species takes in general very long time and different mechanisms are likely to operate during the various stages of this process. Accordingly speciation should be studied at different levels of species divergence. In this thesis I have studied ecological and genetical differentiation between two ecotypes of *Littorina fabalis* as well as between *L. fabalis* and *L. obtusata* - two closely related, directly developing, marine, intertidal gastropods. In *L. fabalis* size is about 25% larger in moderately exposed habitats compared to sheltered habitats and in this thesis I present data showing that this genetically inherited size different is maintained by an interaction of several selective forces, including life history optimisation, size selective crab predation, fucoid algae functioning as refuges from crab predation and wave-induced dislodgement.

The two ecotypes of *L. fabalis* differ also in the protein arginine kinase (*Ark*) and a Randomly Amplified Polymorphic DNA (RAPD) locus and this linkage disequilibrium persists in locations where both ecotypes are present suggesting that recombination is strongly suppressed between *Ark*, the RAPD locus and one or several loci influencing size. Chromosomal rearrangement, in particular inversions are very effective in restricting recombination and if locally adapted alleles in at least two loci on the same chromosome occur in heterogeneous environments, an inversion may immediately protect these from being mixed up with alleles (introduced by migration) that are locally adapted for other microhabitats. This model predicts that differential selection on these alleles exist before an inversion appears and I have tested this by sequencing an intron of *Ark*. The SS-ecotype was nearly fixed for one haplotype while the diversity among LM-ecotypes was much higher supporting a scenario where a recently derived inversion (or other kind of chromosomal rearrangement) restricts recombination between *Ark* and one or several loci that influence size. In this thesis a novel method used for the sequencing of the *Ark* intron that does not require the cloning of each sample individually (which is both time consuming and expensive) is also presented.

Littorina fabalis and *L. obtusata* are considered as well defined species with clear differences in ecology, morphology and nuclear DNA (allozymes) and with microsatellites I could show that hybridisation between these species has not been occurring at least during the last 10,000 years (they are easily identified in the field by both size and coloration). Despite this they show no consistent differences in the mitochondrial cyt-b gene, which could either be due to incomplete lineage sorting or introgression. The idea that mitochondrial DNA can be used as a barcode in species identification is attractive but has in recent years gained criticism because the nature of the mitochondrial molecule makes it specifically prone for introgression between species. Locally restricted mitochondrial introgressions are common among closely related species but the flat periwinkle case study in this thesis clearly shows that a lack of mitochondrial divergence can also exist throughout the whole distribution range for a geographically wide spread species (*L. fabalis* and *L. obtusata* occur sympatrically from Spain to Iceland and the White Sea, Russia).

Keywords: Local adaptations, chromosomal rearrangements, speciation, mitochondrial introgression

TO MY FAMILY AND FRIENDS

LIST OF PAPERS

This thesis is based on the following papers referred to in the text by their Roman numerals. The papers are appended at the end of the thesis

I Kemppainen P., Panova M., Hollander J., Johannesson K. (submitted manuscript). Reticulate evolution in marine snails – Evidence from morphology, microsatellites and mitochondrial DNA

II Kemppainen P., Van Nes S., Ceder K., Johannesson K. 2005. Refuge function of marine algae complicates selection in an intertidal snail. Oecologia 143: 402–411

III Bierne N., Tanguy A., Faure M., Faure B., David E., Boutet I., Boon E., Quere N., Plouviez S., Kemppainen P., Jollivet D., Moraga D., Boudry P., David, P. 2007. Mark–recapture cloning: a straightforward and cost effective cloning method for population genetics of single copy nuclear DNA sequences in diploids. Molecular Ecology Notes 7 (4): 562-566

IV Kemppainen P., Lindskog T., Johannesson K. (manuscript). A test of the local adaptation mechanism for ecotype linked chromosomal rearrangement in the marine gastropod, *Littorina fabalis*.

A doctoral thesis at a university in Sweden is produced either as a monograph or as a collection of papers. In the latter case, the introductory part constitutes the formal thesis, which summarises the accompanying papers. These have already been published or are manuscripts at different stages (in press, accepted, submitted or in preparation).

CONTENTS

AIMS AND SCOPE OF THIS THESIS

The study of natural sciences is driven by an urge to explain everything in as simple theories as possible. Since speciation is such a slow process, can take countless different paths and involve so many different mechanisms, there are still numerous aspects of this where scientists have not reached a consensus, despite over 150 years of work. In this thesis I have studied speciation at two levels; between two ecotypes of *Littorina fabalis* and between *Littorina fabalis* and *Littorina obtusata* (two closely related direct developing marine intertidal gastropods), using ecological experiments as well as phylogenetic and population genetic approaches with several types of genetic markers. At the beginning of my work two facts were clear: *i*) in *L. fabalis* size increases with about 25% going from sheltered to moderately exposed habitats but it was uncertain in what way this difference is maintained by natural selection and *ii*) the protein arginine kinase (*Ark*; as revealed by allozyme electrophoresis) and a Randomly Amplified Polymorphic DNA (RAPD) locus varies predictably between these ecotypes, despite no differentiation at almost 50 other allozyme, RAPD and microsatellite loci. In fact, this association is so strong that if you go to a moderately exposed shore anywhere from the Britannic peninsula (in France) to Norway (including UK), you will almost exclusively find large snails, which will contain a specific set of exposed *Ark* alleles and an RAPD band that is not present in small individuals from sheltered locations. Accordingly in **paper II** I started out with investigating how natural selection shapes the size variation of *L. fabalis* (which is genetically determined) through a series of manipulative field and laboratory experiments. The linkage between micro-habitat and *Ark* was thereafter studied by sequencing an intron of *Ark* for both ecotypes in several locations throughout the distribution range (**paper IV**). For this study I used a novel sequencing procedure (described in **paper III**) that does not require individual cloning of PCR products (which is time consuming and expensive). Our hypothesis was that if recombination is strongly restricted in the region of the genome where *Ark* is located, possibly due to a chromosomal rearrangement, the individual genotypes would group according to ecotype and not by geographic location. Since loci responsible for species-specific differences often map to chromosomal rearrangements, especially in sympatric species-pairs, their potential roles in differentiation and speciation are particularly interesting. However most others studies have been conducted in closely related species, which are fixed for different rearrangements and the two ecotypes of *L. fabalis* gives therefore a rare opportunity to study the segregation of a possible rearrangement and its early dynamics.

The idea that mitochondrial DNA can be used as a barcode in species identification is attractive but has in recent years gained criticism because the nature of the mitochondrial molecule makes it specifically prone for introgression between species. These introgressions are however mostly confined to local populations or to species that are otherwise geographically restricted (i.e endemic species). *Littorina fabalis* and *L. obtusata* are considered as well defined species with clear differences in ecology, morphology and nuclear DNA (allozymes). In a phylogeographic study, which was initially planned with special reference to the ecotypes of *L. fabalis,* the mitochondrial cytochrome b gene did not (highly unexpectedly) show any consistent differences between *L. fabalis* and *L. obtusata* throughout their whole pan-European distribution range (**paper I**). This prompted me to complement the study with additional data from morphology and microsatellites to make sure that *i*) I indeed identified the species correctly and *ii*) that they really are different species. This also changed the subject of the study to one about reticulate evolution and mitochondrial introgression rather than phylogeography.

INTRODUCTION

Ever since the publication of "The Origin of Species" by Charles Darwin (1859), biologists have been challenged by two core questions: what are species? and how do they form? Most of the ideas about speciation emerged already during the first half of the last century, but not until the second half did it become possible to test and fine-tune many of these theories with the advance of molecular genetic techniques and computer simulations. Despite the existence of a wealth of different species concepts (de Queiros 1998), they virtually all conform to the general idea that species are segments of population level evolutionary lineages that at some time during their divergence cross a threshold beyond which their separation becomes irreversible (Templeton 1998, de Queiros 1998). If geographic barriers to gene flow are lacking, isolating mechanisms do not typically provide a complete barrier to gene flow between closely related species. Instead, species boundaries are often semi-permeable, allowing neutral or advantageous alleles to move between species unless they are tightly linked to loci that contribute in some way to isolation (Rieseberg et al. 1999, Navarro and Barton 2003, Panithanarak et al. 2004, Martinsen et al. 2001, Barton and Bengtsson 1986). This does not necessarily counteract species divergence, because new mutations that are either contributing to adaptive differences between species or tightly linked to isolating loci may continue to accumulate (Morjan and Rieseberg 2004). Thus, for many taxa, the unit of isolation is not the entire genome, but rather genomic regions that harbor isolation loci. This implies a balance between gene flow and natural selection, where selection overrides the effects of gene flow in the course of speciation (Templeton 1998).

The Evolution of Genic Incompatibilities

The most straightforward way by which evolutionary lineages may split up is when two populations become isolated from each other by a geographic barrier (allopatric speciation). Each population experiences random mutations that will become fixed (due to chance or selection) over long periods of time until they become distinct genetically. Since alleles in the populations have evolved independently during the separation time, they have not been tested with each other and there is thus a substantial chance that a diverged allele will not work well in the genetic background of the other species. This forms the basis of the Dobzhansky-Muller incompatibility (which was originally proposed by Bateson already at 1909, see Coyne and Orr 2004) and it appear to be the most common way by which negative selected gene interactions in hybrids arise in natural populations (Coyne and Orr 2004). Even if each individual incompatibility is likely to have only small effects on hybrid fitness the cumulative effect of many may result in complete reproductive isolation. However, until this has happened, species boundaries are semi-permeable, allowing neutral or advantageous alleles to introgress unless they are linked to loci that are involved in the hybrid incompatibilities. This can be illustrated by a simple example (Box 1): consider two populations with the initial genotype aabbcc (for simplicity all loci are on the same chromosome) that become isolated for a period of time and allele A will be fixed in population 1 and allele B in population 2. These substitutions cause an incompatibility in the hybrid, and we are interested in the possibility of the neutral allele C (in population 2) to introgress into population 1. Through recombination, the hybrid may form eight different gametes and when it is back-crossed with an individual from population 1 (which only produces Abc gametes) there are subsequently also eight possible zygotes of which two include the neutral allele C but not the allele B that causes incompatibility together with A. The introgression of C into

population 1 is therefore ultimately dependent of the magnitude of selection on the combination of A and B and the likelihood that recombination will separate B and C (Martinsen et al. 2001). From this example two additional conclusions may be drawn; all chromosomes which do not include alleles that cause incompatibilities are free to introgress (given that reproductive isolation is not already completed) and this is particularly true for the mitochondrial DNA molecule, as it's exceptionally small genome size (at lest compared to genomic chromosomes) greatly decreases the probability of negative gene interactions with nuclear DNA (Martinsen et al. 2001). The second conclusion is that if recombination is somehow restricted between B and C, the allele C cannot escape the incompatibility caused by the combination of A and B (Noor et al. 2001b). This thesis is strongly related to these conclusions as during the divergence of *L. fabalis* and *L. obtusata* mitochondrial introgression appears to have been very common despite clear differences in the nuclear DNA (**paper I**) and in *L. fabalis* the genetic differences between locally adapted ecotypes are restricted to a chromosomal region where recombination is strongly suppressed (**paper II** and **IV**).

MITOCHONDRIAL INTROGRESSION

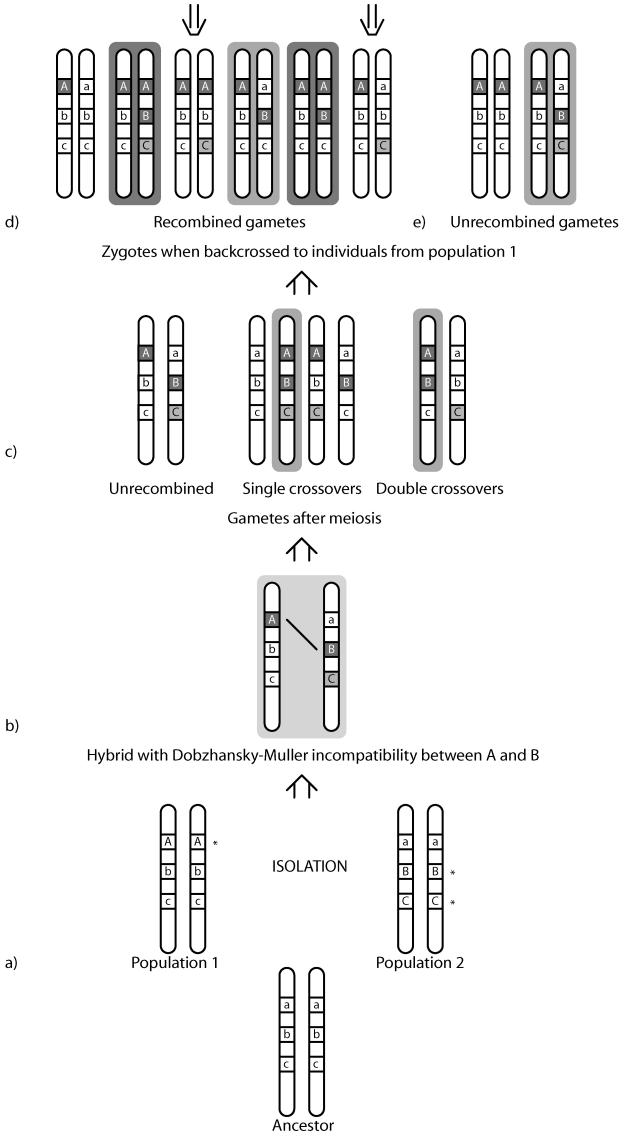
Mitochondrial DNA has been used extensively in the last three decades as a tool for inferring the evolutionary and demographic past of both populations and species. The foremost reason for this is that the mitochondrial DNA molecule has a potentially higher probability of being congruent with the species tree in closely related species compared to nuclear DNA. This is because it is maternally inherited and haploid, which reduces the effective population size to a fourth that of a nuclear-autosomal gene (Moore 1995, see also Bazin 2006a, 2006b and Berry 2006). One definition of the effective population size is "the number of breeding individuals in an idealised population that would show the same amount of genetic variation under random genetic drift as the population under consideration" (Wright 1931, 1938). Random genetic drift is in turn the evolutionary process of change in the allele frequencies of a population from one generation to the next due chance and the larger the effective population size the smaller the effect of random genetic drift will be and thus the more genetic variation (originally introduced to the population by mutations) a population may hold (Hartl and Clark 1997).

All chromosomes in a species segregate independently and due to recombination most genes within a specific chromosome will also represent independent samples of the genome (unless recombination is somehow restricted; Box 1). If the gene copies share a most common ancestor further back in history than the actual splitting of two species, the gene tree does not accurately represent the species tree (Box 2). If no introgression occurs between two species the gene copies will nevertheless accumulate unique mutations (which cannot spread to the other species) and given enough time the ancestral lineages will eventually be completely lost due to random genetic drift (this is called lineages sorting). The larger the population size of the species, the more slowly this process will be (e.g. Avise 2000, Hartl and Clark 1997) but if introgression *does* occur it will reintroduce previously lost allelic lineages into one or both gene pools and it is sometimes difficult to separate this phenomenon from incomplete lineage sorting.

Once a mitochondrial DNA molecule has spread from one species to the other, it can only be lost by natural selection or random genetic drift and because of this a signal of introgression can be present longer in mtDNA than in nDNA, where recombination as well can obscure the genetic signal of past hybridisation (Funk and Omland 2003). In addition, once mtDNA has leaked through a species boundary it may be driven to fixation by directional selection (i.e. selective sweeps; e.g. Bachtrog et al. 2006, Wilson and

Box 1. Introgression between two species with Dobzhansky-Muller incompatibilities

Schematic outline of the possibility of introgression between two populations after initial divergence and secondary contact. (a) When populations are isolated for a period of time, different mutations (indicated by asterisk) will be fixed in the different populations. (b) Dobzhansky-Muller (D-M) incompatibility between A and B (Coyne and Orr 2004, see text for details) occur in the hybrid individual. Background shade indicates impaired vigour or fertility, which can occur att all stages, b to d (and also in the adult backcross individual) and is cumulative. "C" is a neutral or positively selected allele and we are interested if it can introgress from population 2 to population 1 despite the D-M incompatibility. (c) Gametes produced by the hybrid. (d) Arrows indicate backcross zygotes that have full vigour and fertility and include the C allele from population B – introgression is possible. (e) If recombination does not occur - introgression of C from population 2 to population 1 is impaired (Noor 2001b). Introgression of loci in chromsomes that do not involve D-M incompatibilites (including mitochondrial DNA) are free to introgress.



Bernatchez 1998). Accordingly, among closely related taxa some discordance between nuclear DNA and mitochondrial DNA is almost the rule rather than the exception (see Chan and Levin 2005 for a review). However so far extensive mitochondrial introgression, to the point where mtDNA from one species completely replaces that of another, have been observed in local populations (e.g. Melo-Ferreira et al. 2005, Ruber et al. 2001, Wilson and Bernatchez 1998, Berthier et al. 2006, Roca et al. 2005) or in species that are geographically restricted (Bachtrog et al. 2006, Carson and Dowling 2006).

The Flat Periwinkle Case Study

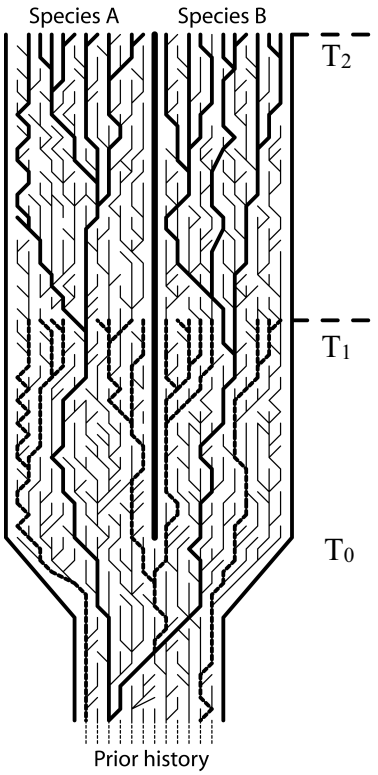
The genus *Littorina* comprises 19 living species of marine rocky-intertidal herbivorous gastropods that are widely distributed throughout the northern hemisphere (Reid, 1996). After the opening of the Bering Strait 3.5-7 million years ago, two independent lineages of *Littorina* invaded the North Atlantic from the Pacific (Reid 1990, Zaslavskaya et al. 1992, Marincovich and Gladenkov 1999). One lineage gave rise to the flat periwinkles (*L. fabalis*, *L. obtusata),* and the rough periwinkles (*L. saxatilis, L. arcana* and *L. compressa*) and the other is still today represented by a single species, *L. littorea* (Reid, 1996). The major distinction between these two lineages is that *L. littorea* has a pelagic larval stage, whereas the other species have direct development (Reid, 1996). As a result, both flat and rough periwinkles have very restricted gene flow and show high degree of ecotypic variation on scales down to just a few meters (Reid, 1996). This has most likely also contributed to the higher number of species in the flat/rough periwinkle lineage compared to the *L. littorea* lineage, which makes this group of species an exceptionally interesting model for the study of speciation.

When I started sequencing the mitochondrial cytochrome b gene (cyt-b; 350 base pairs) for *L. fabalis* and *L. obtusata* it was obvious that there was no clear difference between them (most haplotypes were shared) and the more populations I sequenced the more it also became clear that it was not a local phenomenon (Fig. 1). In fact the only deep divergence was found between a clade of Spanish *L. fabalis* and *L. obtusata* haplotypes and the remaining haplotype network. This was surprising since *L. fabalis* and *L. obutsata* differ in a number of ecological and morphological characters (Table 1) and in addition eight out of 30 allozyme loci have been found to be fixed between these species and the overall allozyme differences (Nei's D = 2.25) translates to a divergence of about 1.25 million years ago (Tatarenkov 1995). According to Coyne and Orr (2004), the minimum time needed to acquire total isolation between species is 1.1 million years for allopatric species pairs and only 0.1 million years for sympatric species pairs. Although we at this point cannot say much about the mode of speciation between *L. fabalis* and *L. obtusata* (if it involved an initial allopatric divergence or not), currently they are completely sympatric over the whole distribution range (from Spain to the White Sea and Iceland, except for a recent [~10,000 years] range expansion of *L. obtusata* to the NW Atlantic; Wares and Cunningham 2001) and therefore a divergence of 1.25 million years seems quite sufficient for complete isolation.

However, reviewers immediately questioned two things *i*) can we be sure that *L. fabalis* and *L. obtusata* are different species and *ii*) can we be sure that we have identified them correctly. This prompted me to complement the study with microsatellite data (four loci) and morphology from three locations; two sites where previous data had shown that many haplotypes were shared and introgression here could potentially be common (Koster, Sweden and Bergen, Norway) and one additional location in Shetland that had not been sampled before. Both *L. fabalis* and *L. obtusata* have direct develop-

Box 2. Species trees vs. gene trees

The figure shows all the possible gene genealogies after species divergence. At T_0 the population is separated by a geographic barrier (the exra bold line). At T_1, ancestral lineages (dashed lines tracks the ancestry of all gene copies present in the population at T_1) have not been given enough time to dissapper from the two species and some gene copies from species A share a closer relative with species B than species A (the species are polyphyletic). Given more time (bold lines tracks the ancestry of all gene copies present in the population at T_2) the lineage sorting proces has been completed and all gene copies share the most recent common ancestor within their own species (the species are monophyletic). This process is highly dependant on the effective population size (N_e) which determines the rate of the random genetic drift and the amont of variation in the ancestral population. It takes on avarage $4N_e$ generations to reach monophyly and since mitochondrial DNA is maternally inherited and essentially haploid (only one gene copy exist in each individual) the effective population size is up to four times smaller that of nuclear-autosomal DNA (Avise 2000). It has also been shown that one would need 16 unlinked nuclear genes to gain as much confidence in the inference of the species tree as that embodied in the single mtiochondrial DNA-haplotype tree (Moore 1995). Introgression yealds polyphyly by introducing phylogenetically divergent lineages across species boundaries (see Fig. 4).



ing larva with limited dispersal. After the retreat of the North European glaciation there was likely a rather immediate colonisation of both sides of the North Sea from a common source area and it is therefore likely that the populations on each side have been largely separated since then (i.e. about 10,000 years). If hybridisation between the two species has been substantial after the separation of these populations (or is still occurring), individuals should not group according to species but rather by geographic location. Two analyses on the microsatellite data were performed to test this; principal component analysis (PCA; using PCAGEN sofware; J. Gouldet, unpublished) and assignment test (GENECLASS 2; Piry et al. 2004). PCA is a way to reduce multidimensional data, (in this case genetic distances between populations), to lower dimensions for visualisation (starting from the dimension that explains most of the data) and assignment tests use multi-locus genotypes to assign individuals to different reference populations (in this case species). In addition, the NEWHYBRIDS software (Anderson and Thompson 2002) was used to see if any individuals were of a hybrid origin. The results were clear: all populations with individuals that were initially a priori characterised as either *L. fabalis* or *L. obtusata* (based on morphology only) clearly clustered by species and not by geographic location (PCA; Fig. 2), no hybrids could be detected and all individuals were correctly assigned to the right species (Fig. 3) with only one exception (out of a total of 320 individuals; even though we were unable to use landmark based morphometrics to separate the species, colour [all *L. obtusata* were olive] and in Shetland size [all *L. obtusata* were larger than *L. fabalis*], were diagnostic). Thus the most important question still remained unanswered: why do we not see any divergence between *L.*

fabalis and *L. obtusata* in the mitochondrial DNA despite clear differences in ecology, morphology (coloration and size) and several types of nuclear genetic markers (mikrosatellites and allozymes). The two most likely explanations are: incomplete lineage sorting (Box 2) and introgressive hybridisation (Box 1). If introgression *does* occur it will reintroduce previously lost allelic lineages into one or both gene pools and the question is therefore whether the lack of a divergence between *L. fabalis* and *L. obtusata* solely can be explained by incomplete lineage sorting or if it is necessary to involve introgression in order to explain the excessive overlap of cyt-b haplotype distribution? In **paper I**, I used a coalescent based approach (lineages are simulated back in time and for each generation these lineages will merge e.g. coalesce until only the most recent common ancestor is left in the population) to test weather introgression can be rejected, using the software MDIV (Nielsen and Wakeley 2001), which simultaneously assigns maximum likelihood estimates for population divergence time (PDT), migration (M; which in our case directly translates to introgressive hybridisation) and effective population size. If introgression can be rejected, we cannot exclude the possibility that the lack of divergence and all the shared haplotypes in the cyt-b haplotype network (Fig. 1) is simply due to a very large effective population size (this would however not make this study less interesting as the lack of mitochondrial divergence has never before been observed in so widely distributed species as *L. fabalis* and *L. obtusata,* between which contemporary hybridisation has not been observed). Indeed, the estimate of the effective population size (from MDIV) in the pooled *L. fabalis* and *L. obtusata* populations was high (0.24-1.2 million). Nevertheless, two lines of evidence indicate that incomplete lineage sorting is not the only explanation. Firstly, the existence of many geographically restricted haplotypes in both *L. obtusata* and *L. fabalis* (most importantly the divergent Spanish haplotypes, see Fig. 1) can hardly be explained exclusively by ancestral polymorphy (Funk and Omland 2003). Secondly, introgression could not be rejected by MDIV, and although this analysis makes many assumptions this, in conjunction with the geographically shared haplotypes, makes a strong case for the occurrence of introgression of mitochondrial DNA during the divergence of *L. fabalis* and *L. obtusata.* A likely scenario is therefore that hybridisation between *L. fabalis* and *L. obtusata* has been common enough to prevent differentiation in mtDNA but at the same time no more frequent than to allow for nuclear genes to diverge (at least those that are linked to incompatibility genes; Box 1).

Because there are opportunities for hybridisation in sympatry but not in allopatry, sympatric sharing of haplotypes is often taken as evidence for introgressive hybridisation (e.g. Donnelly et al. 2004, Morando et al. 2004, Peters 2007). Consequently, species differences my continue to accumulate in allopatry where the divergence is much less likely to be impeded by introgression. However *L. fabalis* and *L. obtusata* occur sympatrically throughout their entire distribution range and therefore the possibility of unconstrained divergence in allopatry is impossible and this may well be one of the most important reasons why we cannot see any general divergence in the cyt-b haplotype network for these two species.

Reticulate Evolution and the Divergent Spanish Clade

At least two basic scenarios may explain the lack of a divergence between *L. fabalis* and *L. obtusata*, given introgression of mtDNA: *i*) either gene flow has been re-occurring between *L. fabalis* and *L. obtusata* throughout their history and species differences have largely been driven by divergent natural selection in sympatry (Fig. 4a), or *ii*) there has been an allopatric separation between them but also a period of extensive introgression

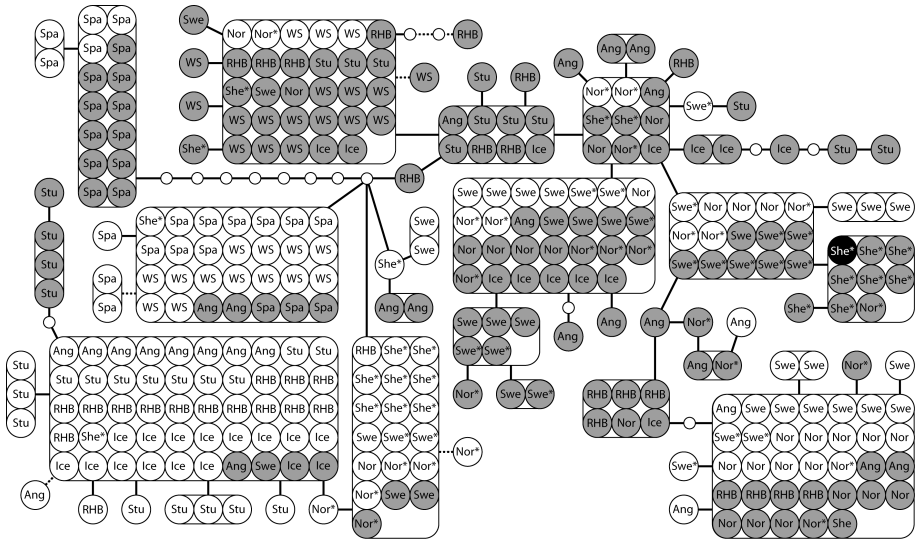


FIGURE 1. Cytochrome b haplotype network. Each large circle represents an individual sequence that has been identified (on the basis of morphology colour and size) as either *L. fabalis* (white) or *L. obutsata* (grey). Text indicates sampling location according to abbreviations in Table 3. Each line represents a mutation and small empty circles are missing, hypothetical haplotypes and dashed lines indicate replacement substitutions. Individuals from the second data set, which were only sequences with the forward primer are denoted with an asterisk. One individual out of 320 was identified incorrectly (black with white text) as *L. fabalis* when it in fact was *L. obtusata* (based on microsatellites).

where mtDNA haplotypes from one species have replaced those of the other after which some differentiation in mtDNA again was possible (Fig. 4b). Both scenarios involve reticulate evolution but differ in the extent and frequency of the introgressive events.

The population divergence time between *L. fabalis* and *L. obtusata* was estimated (by MDIV) to 0.27-0.77 million years ago which is considerably more recent than the estimate from allozymes (1.25 million years ago). However, the divergent Spanish clade is separated from the main haplotype network by 9 mutations and the minimum sequence divergence of 2.3% (K2P-distance) between them translates to a maximum divergence time of 0.72-3.65 my (based on our estimate of the minimum sequence divergence rates, see details in **paper I**), which indicates a divergence close to the base of the *L obtusata*/*L fabalis* separation. Normally a deep divergence within a single species is explained by a vicariance event where a geographic barrier has separated populations from each other (Avise 2000). The divergent Spanish clade could therefore be a result of an initial allopatric separation between *L. fabalis* and *L. obtusata*, which might have contributed to the large difference in allozymes as well. This supports the second scenario outlined above of an early divergence between *L. fabalis* and *L. obtusata*, secondary contact and replacement of haplotypes from one species to the other except for the Spanish clade. The fact that the Spanish clade includes individuals from both species does not necessarily contradict this hypothesis as this clade (like all common haplotypes) has most likely been influenced by introgression in the recent past. A major drawback of the MDIV model is however that it assumes that two populations split from a panmictic ancestral population after which gene flow may or may not have occurred whereas a more likely scenario for our cyt-b data is that there firstly have been a clear separation and then a replacement of haplotypes from one species to the other. If this is the case, all traces of prior history must have disappeared (except for the Spanish clade) and therefore the population divergence time from MDIV essentially estimates the sec-

TABLE 1. Summary of morphological and ecological character differences between *L. fabalis* and *L. obtusata*.

Character	Difference	Species status
Pigmentation	Soft parts differ	Diagnostic at most localities
Penis	Long/short tip	Always diagnostic
Spermatozoa	Size differ	Slight overlap
Pallial oviduct	Size differ	Always diagnostic
Spawn	Shape differ	Local difference reported
Radula	No of cusps differ	Some overlap
Shell	Overall size differ	Almost diagnostic (<95% of cases)
Colour	Variable	Diagnostic in some localities
Zonation	Different zones of abundance in tidal areas	Main distribution differ but some overlap
Exposure	Different preferences	Some overlap in preferred ranges of exposure

Modified from Reid (1996) (From **paper I**)

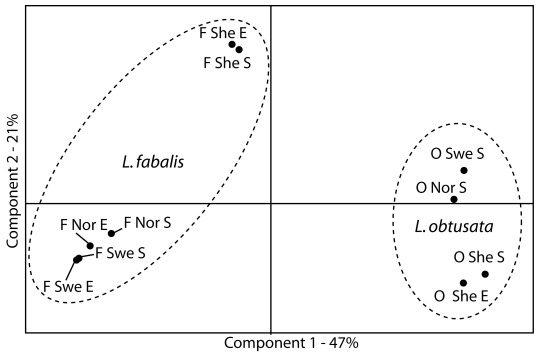


FIG. 2. Principal component analysis (PCA) based on mircosatellite allele frequencies in *L. fabalis* (F) and *L. obtusata* (O) from three different locations (Koster [Swe], Bergen [Nor] and Shetland [She]) and two different microhabitats (moderately exposed, E and sheltered, S). Percentage of explained variation per axis is indicated along axes. (From **paper I**)

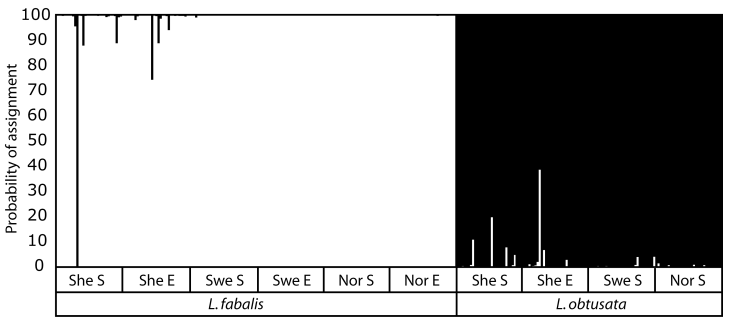


FIG. 3. Assignment of population structure with GENECLASS 2 based on four microsatellite loci. Individuals were a priori grouped into *L. fabalis* and *L. obtusata* from three locations; Koster (Swe), Bergen (Nor) and Shetland (She) sheltered (S) and moderately exposed (E) shores. Only one individual was assigned with a significant probability as a first generation migrant i.e. this individual was incorrectly identified as *L. fabalis*. (From **paper I**)

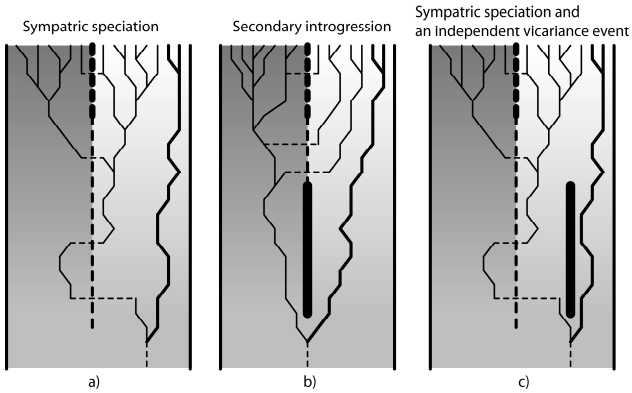


FIG. 4. Three possible hypotheses that can explain the discrepancy between nuclear and mitochondrial DNA. Background colour indicates overall genetic differences and the gene genealogy represents cytochrome b. The divergent Spanish clade is indicated in bold. a) sympatric speciation i.e. gene flow between *L. fabalis* and *L. obtusata* during the whole speciation process, b) initial allopatric speciation with secondary contact and replacement of haplotypes from one species to the other (except for the Spanish clade), c) Sympatric speciation with an independent vicariance event in one of the species. Secondary introgression in the Spanish clade is not included in this figure.

ondary divergence and not the first. From our data it is not however possible to definitely separate a scenario where the initial divergence has been between incipient *L. fabalis* and *L. obtusata* (Fig 4a) from one where a vicariance event has affected one species independently (Fig 4c). In fact, a highly diverged mitochondrial clade in *L. saxatilis* has also been detected in Spanish populations (Mäkinen et al., unpublished), suggesting that a common historical geographic separation might have similarly affected *L. fabalis* or *L. obtusata* and *L. saxatilis*.

THE ROLE OF CHROMOSOMAL REARRANGEMENTS IN SPECIATION

Major chromosomal rearrangements involving a change in the position of genes such as inversions, reciprocal translocations, Robertsonian rearrangements (centric fusions and fissions) and tandem fusions are potentially heterotic, that is unbalances gametes with duplications and deficiencies may be formed as a result of meiotic malsegregation in the heterokaryotype (King 1987). These rearrangements are common in most animal and plant taxa, and may therefore contribute substantially to the genomic incompatibilities arising during allopatric divergence (White 1978, King 1993). Since they potentially reduce hybrid fitness, it has also been suggested that chromosomal rearrangements play active roles in speciation, but in modern literature this is considered unlikely for several reasons. There is for instance some controversy on how chromosomal rearrangements initially become fixed in a population if they are strongly underdominant i.e. have heterozygote disadvantage since in the early establishment of any mutation the new allele occurs almost exclusively in heterozygotes (King 1987, Hedrick 1981, Walsh 1982, Spirito 1998). Accordingly, chromosomal mutants with very low level of selection against the heterokaryotype can become established more easily, but are on the other hand not likely to play an important role in restricting gene flow (Spirito 1998).

Some chromosomal rearrangements are effective in suppressing recombination independent of any causal effect on hybrid fitness. This effect is strongly associated with chromosomal inversions (Rieseberg 2001, Navarro and Barton 2003, Feder et al. 2003b) and to some extent the centromeres of Robertsonian heterozygotes (Panithanarak et al. 2004). Accordingly, if two populations are fixed for different chromosomal

rearrangements that restrict recombination between alleles involved in hybrid incompatibilities, gene flow will be impaired in the whole rearranged segment (Noor et al. 2001b; Box 1). Consequently the suppression of recombination by chromosomal rearrangements is likely to be more important in restricting gene flow between populations than their effects on hybrid fertility and fitness. For example, in three wild sunflower (*Helianthus*) hybrid zones, rates of introgression were shown to be 50 % lower across chromosomes carrying rearrangements than across collinear ones (Rieseberg et al. 1999). Furthermore, loci responsible for hybrid male sterility and female species preference in *Drosophila pseudoobscura* and *D. persimili*, two occasionally hybridising species, are located in rearranged chromosomes (Noor et al. 2001a) and a multi-locus analysis of the same species pair also indicated reduced gene flow in rearranged chromosomal regions (Machado et al. 2002).

In contrast to allopatric speciation, the splitting of evolutionary lineages may also occur sympatrically, solely due to natural selection despite a potential for gene flow during the whole speciation process (Johannesson 2001, Via 2001). However, a major theoretical obstacle in sympatric speciation models has been that recombination during gamete formation may mix up genes for mating preferences with genes for the selected trait, whenever occasional mating between lineages occurs. This creates individuals with preference for the opposite lineages, increasing gene flow between them and thereby opposing speciation (Tregenza & Butlin 1999, Johannesson 2001, Via 2001). In light of this, a linkage disequilibrium (a greater than chance for two alleles to be inherited together) is essential for the formation of assortative mating and thus for sympatric speciation in general. Introgression between two allopatrically diverged populations and parapatric speciation (where introgression is initially allowed in hybridzones) are likely to depend on similar circumstances - if hybrid viability/fertility is impaired by genomic incompatibilities the species boundary may be reinforced by selection favouring assortative mating (Servedio 2001). In fact, reinforcement is theoretically often treated as equivalent to a latter stage of sympatric speciation where hybrids are likely to be intermediate in phenotype compared to the parental populations and cannot therefore compete with either of them (Servedio and Noor 2003). The bottom line is that many models of speciation are dependent on a strong linkage between important speciation genes e.g. those that are necessary for assortative mating, which can be created by chromosomal rearrangements.

Inversion Polymorphism and Adaptive Divergence

Even if the new rearrangement is associated with a favourable gene complex (homozygote advantage) the problem of underdominance is still argued to be present with respect to initial establishment (Lande and Arnold 1985, Spirito 1993). Accordingly mathematical models have shown that the establishment of underdominant chromosomal rearrangements is possible only in small and inbred populations where genetic drift is strong (Hedrick 1981, Walsh 1982, Lande and Arnold 1985, Spirito 1993). Nevertheless many rearrangements, especially pericentric inversions (inversions that include the centromere) show very little or no effect on hybrid fertility (Sites and Moritz 1987, Coyne et al. 1993, see also King 1992) and can therefore be fixed in a population entirely due to random genetic drift. Furthermore, consider a heterogeneous environment with locally adapted ecotypes and that disruption of favourable allele-combinations between ecotypes is constantly taking place due to gene flow. A chromosomal rearrangement that locks closely located locally adapted alleles may be directly beneficial since it rescues ecotype lineages from the flow of maladaptive alleles from other ecotypes

(Kirkpatric and Barton 2005). This scenario applies equally to species that hybridise since, from a population genetic perspective, hybridisation and immigration are equivalent. Once inversion polymorphism has been established ecotype-specific differences may continue to accumulate despite continuous gene flow across the rest of the genome (Navarro and Barton 2003, Rieseberg and Livingstone 2003). If the rearrangement encloses both loci that are important for the adaptive differences between ecotypes and mate preference traits it could easily restrict gene flow for the rest of the genome as this generates assortative mating within each ecotype. In the course of time this may ultimately trigger sympatric speciation and drive reproductive isolation to completion (although this is depended of many other factors as well, see Coyne and Orr 2004).

If the inverted segments do not harbor isolating loci (e.g. loci for assortative mating) and do not comprise large portions of the genome (which increases the chances of genomic incompatibilities), the inversion polymorphism linked to different ecotypes does not by any means have to lead to speciation. However, the stronger the selective pressure favouring each ecotype in their respective habitats the higher the chance will be that new inversions will be fixed, as more loci are likely to have locally adapted alleles. In essence, ecological speciation is still ultimately dependent on natural selection (Schluter 1998, 2001, Via 2002), whereas chromosomal changes such as inversions may be powerful in facilitating this process rather than actively driving it (Navarron and Barton 2003, Riesenberg and Livingstone 2003). In light of this it is rather appealing to link studies of adaptive polymorphism to studies of chromosomal evolution. One such example comes from the apple maggot fly, *Rhagoletis pomonella* in North America (Feder et al. 2003a, 2003b). Traditionally the fly infests fruits of hawthorns (*Crataegus* spp.) but in the mid-1800s the fly formed a sympatric race on the introduced cultivated apple (*Malus pumila*). The derived apple-infesting race shows consistent allele frequency differences from the hawthorn host race for six allozyme loci mapping to three different chromosomes which all contain inverted segments distinguishing the races. Indeed, many of the loci coding for adaptive traits (typically the duration of overwintering pupal diapause) that discriminate between the two races are located within the inverted segments. However, nucleotide sequence data suggest that inversion polymorphism for *R. pomonella* has been segregating for ~ 1 million years, which greatly predates the ~ 150-year-old origin of the apple race (Feder et al. 2003b). Thus it seems that the host race shift in *R. pomonella* has been mediated by an already existing inversion polymorphism within the original host race infesting hawthorns. It is not clear therefore whether the initial establishments of inversions were purely neutral and later functioned only as chromosomal niches where adaptive differences could accumulate or if the inversions were adaptive from the beginning e.g. by protecting locally adapted genes from maladaptive gene flow.

Ecology Matters

As a result of divergent natural selection, species living in heterogeneous environments often show habitat-linked morphological variation in adaptive traits (Levins 1968, Hedrick et al. 1976, Endler 1977, Hedrick 1986) and this is particularly common in species with a restricted gene flow among populations (Johannesson et al. 1993, Nevo 1998, Filchak et al. 2000, Hendry et al. 2002). Shell size is most commonly negatively correlated with wave-exposure in direct developing *littorinid* gastropods (see **paper I** for a summary). This is generally explained as a result of more intense crab predation promoting large and robust shells (Boulding et al. 1999) in shores protected from heavy wave-action (Elner and Raffaelli 1980, Johannesson 1986, Boulding and Van Alstyne

1993, Boulding et al. 1999). Exposed-shore snails, on the other hand, reduce their risk of being dislodged and increase their hiding possibilities in crevices by adopting a smaller size (Underwood and McFadyen 1983, Denny et al. 1985, Trussell et al. 1993). However, as already mentioned, in *Littorina fabalis* shell size generally increases with wave-exposure, such that a large ecotype is present in moderately exposed habitats and a small ecotype in sheltered habitats. In **paper II** these have been dubbed the "Large-moderate" (LM) ecotype and the "Small-Sheltered" (SS) ecotype in order not to confuse the LM-ecotype with the exposed (E) ecotypes present in many other *littorina* species. *Littorina fabalis* dwells mostly in the canopy of fucoid algae where it mainly grazes on epiphytes (Norton et al. 1990) and the habitat is confined to the littoral algal belt of the North-East Atlantic, from Spain to the White Sea, including Iceland and Greenland (Reid 1996). Juveniles of *L. fabalis* hatch directly from egg-masses laid on plant fronds, and dispersal of juveniles and adults is restricted to a few meters per generation (Williams 1990; Tatarenkov and Johannesson 1998). Subsequently gene flow among populations is low (Tatarenkov and Johannesson 1994) and accordingly an increased size in moderately exposed habitats seems to be consistent for most of the species' geographic distribution (Reid 1996) and has been described in detail from Wales (Goodwin and Fish 1977, Reimchen 1981), British Isles, Ireland (Reimchen 1982), the west-coast of Sweden (Tatarenkov and Johannesson 1999, **paper II**), Bergen (Norway) and Shetland (**papers I** and **IV**).

The size difference between morphs is to a large extent inherited and the cessation of growth in SS-ecotype is coupled to early maturation (Tatarenkov and Johannesson 1998). Furthermore, upon reaching maturity, shell morphology changes slightly and the shell grows thicker (Goodwin and Fish 1977, Reid 1996). As a consequence small adults are more resistant to attacks from the green shore crab, *Carcinus maenas* (the main predator of *L. fabalis*), than are immature snails of similar size (Reimchen 1982). However, even small adults may be crushed by crabs of size 40-60 mm (carapace width), while large adult snails are much more resistant to these attacks (Reimchen 1982). From these observations Reimchen (1982) proposed that early maturity and therefore smaller size could be an adaptation to predation by small crabs. This hypothesis predicts that the mean size of crabs would be smaller in sheltered shores than in exposed shores. However, in **paper II** we have shown that in Swedish sites there is no difference in neither crab size nor crab density between sheltered and moderately exposed habitats and Reimchen's hypothesis is therefore not supported. We therefore proposed an alternate explanation: if the canopy of fucoid alga constitutes a refuge from predatory crabs and the risk of dislodgement from the canopy is greater in wave-exposed than in sheltered habitats, larger and less susceptible snails should be favoured by selection in exposed habitats. This model makes at least three explicit predictions: *i*) the risk of crab attacks should be lower when snails are dwelling on fucoid algae compared to the substrate below the algae; *i*) large snails (LM-ecotypes) should be more resistant to crab predation than small snails (SS-ecotypes) and *iii*) the risk of dislodgement from fucoid algae should be higher at moderately exposed habitats than sheltered habitats. In **paper II**, we found support for all three predictions. Fewer SS-morphs were attacked when tethered to algal canopy than to stones in two independent experiments (Fig. 5 a and b). In one experiment, tethered LM-ecotypes were more resistant against crab attacks than SS-morphs (both were tethered in exposed habitats) and the probability of crab attacks decreased with snail size (Fig. 6 a and b). Finally we demonstrated that overall the risk of dislodgement increased with wave exposure and thus the refuge function of algae becomes less effective in moderately exposed compared to sheltered habitats (Fig. 7 a and b). This model is based on the assumption that a small size for *L. fabalis* is favoured by selection in the absence of crab predation, which is supported by

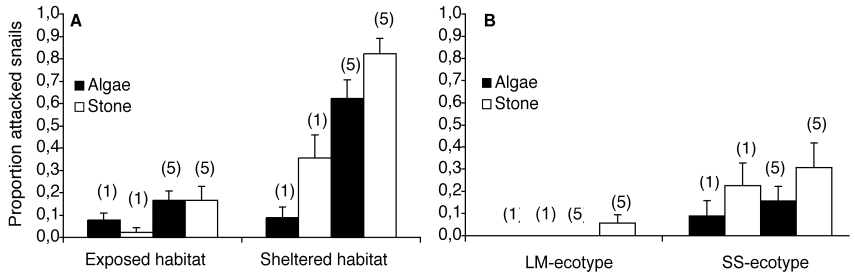


FIG. 5. (A) Attack rates when LM- and SS-ecotypes of *L. fabalis* were tethered to their own habitats. The interaction between habitat and position was significant after one day of tethering but not after five days. (B) Attack rates when LM- and SS-ecotypes of *L. fabalis* were tethered in moderately exposed habitats. LM-morphs were less attacked after both two and five days of tethering. Numbers above bars indicate how many days the snails were tethered. Bars indicate SE. (From **paper II**)

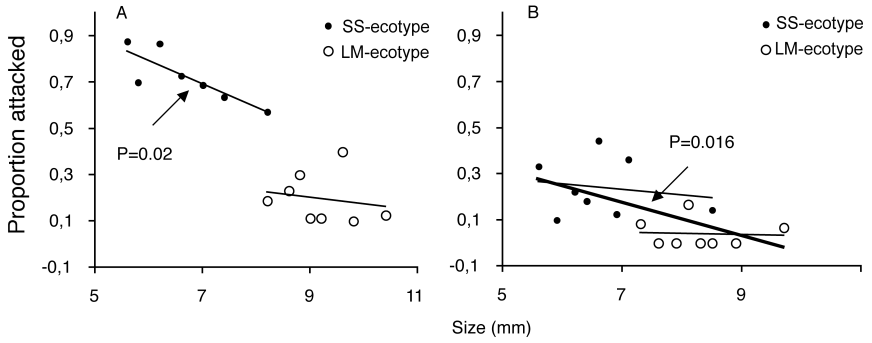


FIG. 6. Proportion attacked snails as a function of snail size in the first (A) and second (B) tethering experiment. Each point represents a size-group consisting of eight to sixteen individuals. (From **paper II**)

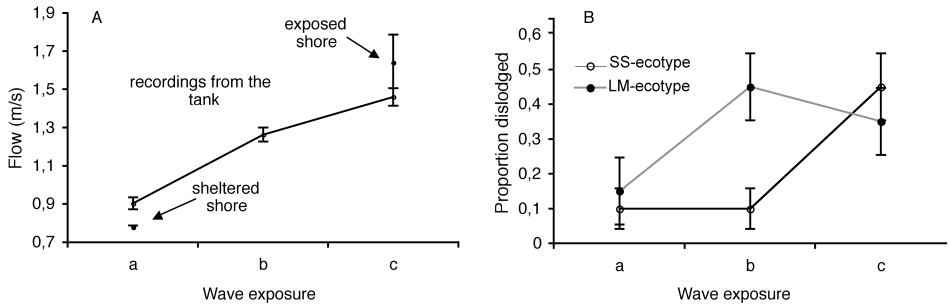


FIG. 7. (A) Water flows in a tank used in a wave simulation experiment (open circles). Level "a" is the lowest flow and "c" is the highest. Filled circles represent values that were obtained from the field during a moderately windy day of SW winds. (B) Dislodgement of LM- and SS-ecotypes of *L. fabalis* from plants of *Fucus vesiculosus* at different water flows (see A). Bars indicate SE. (From **paper II**)

the fact that in the White Sea (where crabs are absent) snails are small in both moderately exposed and sheltered habitats. In addition, a short life-span is also in general considered to select for early maturation (and small size) as this reduces the costs of prolonged growth and delayed reproduction (Blanckenhorn 2000).

Ecotype Linked Inversion Polymorphism in Littorina fabalis

Despite lacking differentiation in 29 allozyme markers, *Ark* segregates in populations of the two ecotypes of *L. fabalis*, and this difference is consistent over Welsh, French and Swedish populations (Tatarenkov and Johannesson 1999). The *Ark* allozyme substructuring was originally described from 21 locations in the Kosterfjord area in Sweden (Tatarenkov and Johannesson 1994) in which two alleles (*Ark* 100 and *Ark* 80) dominated the exposed habitats while *Ark* 120 was strictly confined to the sheltered habitats (mean frequency in sheltered habitats was 0.740 but only 0.035 in exposed habitats). Three additional alleles were also found but only in very low frequencies. In Spain and in the White Sea (Russia), the size difference between *L. fabalis* living in exposed and sheltered locations does not exist and the populations in Spain are monomorphic for the *Ark* 90 allele (Tatarenkov and Johannesson 1999), while in the White Sea (Russia) the sheltered locations are instead dominated by the *Ark* 130 allele (Fokin, pers. commun.). Individuals that are *Ark*-heterozygotes are present in zones of intermediate exposure where the two ecotypes overlap in distribution, but heterozygotes are fewer than expected by Hardy-Weinberg equilibrium (Tatarenkov and Johannesson 1999), which could be either due to selection against heterozygotes or sampling of non randomly mixing populations with different allele frequencies (i.e. Wahlund effect). In another study (in the Kosterfjord area, Sweden) almost all individuals that were homozygous for exposed *Ark* alleles (collected from habitats of intermediate exposure where both ecotypes are present) amplified an extra band (that was not present in snails that were homozygous for sheltered *Ark* alleles) for an RAPD locus, while 18 other RAPD loci showed no systematic differences (Johannesson and Mikhailova 2004). There is thus ample evidence of two molecular loci being linked to each other and to one or several loci influencing growth and adult size and furthermore, the variation in all three characters correlates with snail microhabitat (Tatarenkov & Johannesson 1999, Johannesson 2004, **paper IV**). One explanation could be that that selection on all these markers strongly coincide with wave exposure. However, this is very unlikely as the linkage between the RAPD locus and *Ark* in the Kosterfjord area (Sweden) persists over intermediate habitats where independent segregation, or recombination (if the loci are on the same chromosome), should break up the linkage. Although many chromosomal differences are capable of restricting recombination, inversions, that do not necessarily impair fertility in heterokaryotypes in any substantial way, are the most commonly occurring in natural populations (White 1978, King 1993, Miró et al. 1992) and therefore an inversion that protects alleles at the *Ark*, the RAPD an the size loci from recombining was suggested as the most probable explanation (Tatarenkov and Johannesson 1999, Johannesson and Mikhailova 2004).

While we know that size in *L. fabalis* is under differential selection in moderately exposed and sheltered habitats (see above and **paper II**) it is likely that *Ark* also is (Tatarenkov and Johannesson 1994). Arginine kinase plays a central role in both temporal and spatial ATP buffering in cells, which display high and variable rates of energy turnover (Wyss et al. 1992). For marine intertidal gastropods attachment to the substrate in order not to be dislodged by waves is crucial and for species that live in heterogeneous environments (with respect to wave exposure) *Ark* may potentially be under differential selection with different alleles favoured in different microhabitats. In addition,

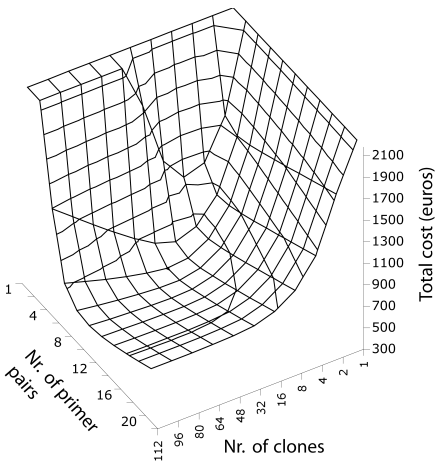
Box 3. Mark-recapture cloning

Sequencing of haploid DNA such as mitochondrial or chloroplast DNA can easily be done within a day using an automated sequencer that nowadays exist in every modern DNA lab. The sequencing of diploid DNA such as nuclear DNA is however more tedious if two gene copies in an individual commonly differ even by the smallest insertion or deletion. Without going into the details, this creates double peaks in the chromatograms (the output form the sequencer) which makes it impossible to determine the sequence. In order to sequence diploid DNA it is therefore often necessary to clone the PCR product since each bacteria only can incorporate one single DNA molecule (in this case a PCR product) and thereafter use the clones as templates for the PCR's (polymerase chain reactions). However cloning is about 10 times more expensive and takes at least one extra day of work compared to the actual sequencing and therefore sequencing nuclear DNA for population genetic studies, which requires many samples, is not practical. To circumvent this problem I have used a novel method dubbed "mark recaptured cloning" or MR-cloning, originally thought of by Nicolas Bierne (**paper III**) for the sequencing of the *Ark* intron in **paper IV**. The idea is to use combinations of primers with unique nucleotides at the 5' ends (this does not affect the annealing performance of the primers) for each individual from a population. Since sequencing is done with universal primers that anneal to the vector sequence in the bacteria, the primer combinations in each clone will be apparent and thus also the identity of the individual. This allowed me to pool PCR products from up to 49 different individuals (using seven differently tagged primer pairs) in one single cloning reaction. The actual individuals that will be "recaptured" after cloning is of course random but for population genetic studies this is sufficient (the sample need to be random anyway). The total cost for this procedure can easily be calculated by the following formula:

$$2P \times p + (S \times a + C)\frac{n}{p^2 - p^2\left(\frac{p^2 - 1}{p^2}\right)^a}$$

Where:
P=the cost of one primer
p=number of primer pairs used
S=cost sequencing one sample
a=number of sequenced clones
n=number of samples in the population that need to be sequenced

This formula excludes any costs associated with labour and it also assumes that the recapture of individuals is completely random (which is not the case, see **paper III**). The procedure is as follows: p number of primer pairs is used to amplify p^2 individuals and "a" number of clones is sequenced from each cloning reaction. This procedure is run as many times as is necessary to get all the samples (n). Let´s use some actual numbers as an example; the costs are: P=15€, C=20€, S=1.5€ (which were the approximate costs when I sequenced *Ark*) and we are interested in sequencing 100 samples. If we plot the total cost as a function of both a and p we get the graph depicted in the figure below. It is easy to see that when cloning only one individual at a time (this is, when p=1) the total cost is 2180 €, which quite a lot compared to less than 500 € using 6 primers pairs (which allows one to clone 36 individuals at a time) and sequencing 32 clones each round. It can also be seen that increasing the number of primer pairs does not increase the total costs very much and since this allows one to clone and sequence more individuals at a time it reduces also the total labour costs (more elaborate calculations have been done in **paper III**).



Box 3. Continued

However, during my work I was able to use the chemicals and competent cells from one cloning kit (price about 20 €) to six different cloning reactions and accordingly the price per cloning was now only 3.3 €. Using similar calculations as above (with C=3.3 instad of 20) the difference between cloning one individual at a time becomes 510 € (for 100 individuals) and the money saved by using the MR procedure is now only about 100 €. Considering that there were also some significant downsides of using the MR-cloning procedure (see text and **paper IV**) I would recommend using the cloning kits for more than one individual rather than using the MR-cloning procedure.

Ark allozyme studies in *L. obtusata* have suggested non-neutral variation along wave-exposure gradients in three geographical regions in the Gulf of Maine, USA (no such studies are available from the NE Atlantic; Schmidt et al. 2007). The difference between wave exposed and sheltered sites was significant but very modest compared to the compressed (<20m) gradients reported from *L. fabalis* in the Kosterfjord area (Tatarenkov and Johannesson 1994). There is thus a substantial chance that if an inversion (or other type of neutral chormosomal rearrangement that restricts recombination) has occurred in a part of a chromosome where *Ark*, and gene(s) for size are located, it will be favoured by selection. This is because a copy of a locally favoured allele at one locus never suffers the disadvantage of being found on the same chromosome with maladaptive immigrant alleles at other loci. This gene has therefore higher fitness than competing copies of the allele at the same locus on chromosomes that do recombine. Consequently, the allele on the inversion spreads, and the inversion spreads with it. This forms the basis of a model by Kirkpatric an Barton (2005) and it specifically predicts that selection on the locally adapted alleles is present when the inversion appears and it is supported by data from several studies including *Drosophila* and the *Anopheles* flies (the main malaria carrying vector; Manoukis et al. 2007).

If an inversion that restricts recombination in heterokaryotypes segregates within a population it essentially becomes divided into two groups, original and inverted, but only with respect to sequences within or close to the inverted segment (Machado et al. 2007). When an inversion appears, it only affects one chromosome in a single individual and initially its dynamics is largely determined by drift as for any other type of new mutation. If it happens to be positively selected, for instance by involving locally adapted alleles, it may rapidly drive to high frequency. Since the inverted chromosome cannot recombine with the standard one, nucleotide variability inside the inversion is expected to be highly reduced (Navarro et al. 2003) until new variation is produced by mutations.

Because the key elements that are required by the model by Kirkpatric and Barton (2005) are likely to exist in *L. fabalis* namely a strong linkage between two locally adapted loci (*Ark* and size) I was very keen on studying this. *Littorina fabalis* has 17 chromosome pairs (Janson 1983), but they are small and thus inversion loops (formed during meiosis) or reversed banding patterns that are indicative of inversions (Ayala and Kieger 1980), cannot be observed directly. Instead I sequenced ~300 bp's of an intron of *Ark* (using mark-recapture cloning see Box 3/**paper III**) for LM-and SS-ecotypes from seven different locations throughout their distribution range (**paper IV**). The prediction was clear; if *Ark* indeed is linked to a habitat/ecotype specific inversion polymorphism (or other kind of segregating rearrangement that restricts recombination, but for convenience I will from here on simply call this an inversion) we expect the variation in one of the ecotypes (irrespective of geographic location) to be much lower (if the inversion is new) and/or any new mutations in the inverted segment to be unique to the inverted

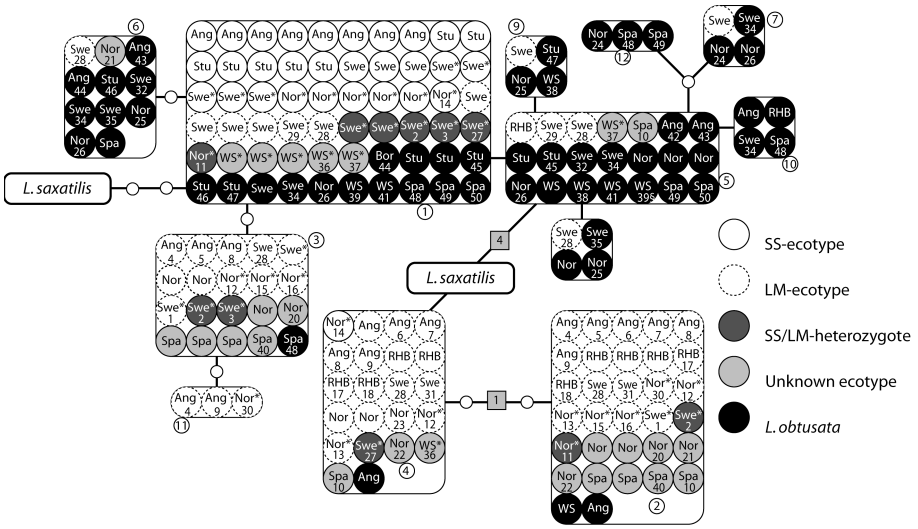


FIG. 8. Haplotype network of an intron of the arginine kinase gene. Large boxes are haplotypes, small circles represent individual sequences for which the sample location is given by the two or three letter abbreviation (Swe=Koster, Sweden; Nor=Bergen, Norway; RHB=Robinhoods Bay, east UK; Ang=Anglesey, Wales; Spa=Vigo, Spain; WS=White Sea, Russia). Lines are mutations, small empty circles represent missing hypothetical haplotypes and boxes are indels with the number of mutation involved indicated by a number. Numbers were given for all samples for which more than one haplotype was found. (Modified from **paper IV**)

chromosome (if the inversion is old). And the results were as clear as the prediction since 27 out of 28 SS-ecotypes from four different sheltered locations (from UK, Sweden and Norway) were monomorphic for one haplotype of the *Ark* intron while the LM-ecotype from the same locations was segregating for ten different haplotypes (Fig. 8). Indeed, the lack of variation in the SS-ecotype certainly shows the inversion is new (at least relative to the time it takes for new mutation to build up new variation). However, it is not useful to estimate the divergence rate for the *Ark* intron sequence as this gene clearly is under selection (and possibly involved in a gene duplication, see below) and we cannot therefore, at this time, approximate the minimum age of the inversion. Nevertheless, as the linkage between size and *Ark* exists on both sides of the North Sea we can only assume two possible scenarios for the spread of the inverted chromosome: *i*) either the inversion was already present in an ancestral population, possibly during the last glacial (a minimum of 10,000 years ago) or *ii*) it was established after the separation of the populations on both sides of the North Sea and then spread secondarily from one geographic location to all other populations through positive selection. In light of the restricted gene flow in *L. fabalis* (Tatarenkov and Johannesson 1994) and the large microsatellite and cyt-b differences between populations on both sides of the North Sea (**paper I**) the first scenario seems more plausible.

According to both microsatellites and cyt-b, the Kosterfjord (Sweden) and the Bergen (Norway) populations are very closely related (**paper I**) and overall the habitats in these geographic locations are superficially very similar. Accordingly, in Bergen a clear size difference between *L. fabalis* in moderately exposed and sheltered habitats has also been described (**paper I** and **IV**). In many sheltered Kosterfjord locations the *Ark*120 allele is completely dominating (Tatarenkov and Johannesson 1994). However in Bergen the frequency of the *Ark*120 allele was only 40% and the exposed *Ark*100

allele, which was fixed for the exposed Bergen location, dominated the sheltered location as well. At first this did not seem odd as this could simply be due to differences in local selection regimes between Kosterfjord and Bergen. When I compared the size of the *Ark*120 homozygotes and the *Ark*100 homozygotes in the sheltered location in Bergen I expected nevertheless that the size of the former would be considerably smaller than the latter as the linkage between size and *Ark* genotype was very clear in the Tatarenkov and Johannesson (1998) study from Kosterfjorden. Very surprisingly this was not the case but nevertheless I did find a significant size difference between *Ark*100 homozygotes from the sheltered locations compared to the *Ark*100 homozygotes from the exposed location. This shows that in Bergen there is no clear linkage between size and *Ark* genotype and size is instead entirely determined by wave exposure, in contrast to snails from Kosterfjord where the *Ark* genotype is a good predictor of snail size (Tatarenkov and Johannesson 1998). Still all six individuals that were homozygous for the *Ark*120 allele from Bergen indeed amplified the same haplotype as all other SS-ecotypes and this haplotype was completely absent among the five *Ark*100 homozygotes. There are at least two alternative explanations for this, either there is no inversion that locks *Ark* and the gene(s) for size in this location and selection operates independently on both characters, or (perhaps less likely) the gene(s) for size have mutated such that we no longer can observe a linkage. This matter may only be resolved once unambiguous markers for the inversion that are independent of both size and *Ark* are available.

A Gene Duplication?

When sequencing with the MR-cloning procedure (Box 3) we only cloned one population at a time since only a minute contamination of primer with the wrong 5′-tag may yield incorrect identification of an individual. This is normally not a problem for population-level analyses, as it does not matter from exact which individual a haplotype comes from as long as we know that it is from the correct population. Nevertheless, in some instances, we repeatedly recovered more than two haplotypes for some individuals. This was highly unexpected as *L. fabalis* and *L. obtusata* are diploid and the *Ark* gene in invertebrates, specifically for arthropods and molluscs, is known to be coded by a single gene (e.g. Suzuki at al. 2000, Munneke 1988). Some of the individuals from which more than two haplotypes were found were sequenced a second time to ensure that no contamination of the primers had occurred in the MR-cloning procedure and we did not expect to recover the exact same haplotypes if contamination indeed had occurred (Table II). Further, in order to exclude that the template itself would contain DNA from more than one individual, I re-extracted DNA from the eight individuals from which more than two haplotypes were obtained and these were in addition cloned individually instead of using the MR-cloning protocol (to exclude the possibility of contamination of tagged primers).

Indeed contamination or mix up of samples during the MR-cloning procedure most likely has occurred as completely different haplotypes were obtained from two independent cloning for individual nr 38 in Fig. 8 (Table 2). Nevertheless from five out of eight samples that were cloned individually (i.e. not using the MR-cloning procedure) more than two haplotypes were also found and here contamination of tagged primers was impossible. In addition from four of these eight individuals, the exact same haplotypes were obtained from at least two independent clonings, even after re-extracting the DNA. This shows that the recovery of more than two haplotypes from one individual is not entirely due to mistakes during the MR-cloning procedure, PCR-

TABLE 2. Summary of sequencing of individuals with more than two haplotypes. Individual and haplotype numbers refer to the numbers given in Fig. 8. Individuals for which at least three haplotypes could be recovered from a minimum of two independent clonings are indicated in bold. Fab=*L. fabalis*, LM=Large-moderate ecotype, Obt=*L. obtusata*.

		First MR-cloning		Second MR-cloning		After re-extraction and individual cloning	
Individual Nr. in FIG. 8	Species	Nr. of sequences	Haplotypes	Nr. of sequences	Haplotypes	Nr. of sequences	Haplotypes
4	**Fab LM**	**5**	**2,3,11**	**14**	**2,3,11**	**6**	**2,3**
8	**Fab LM**	**9**	**2,3,4**	**3**	**2,4**	**9**	**2,3,4**
9	Fab LM	4	2,4,11	0	na	4	2,11
25	**Obt**	**4**	**6,8,9**	**0**	**na**	**9**	**6,8,9**
26	Obt	4	5,6,7	0	na	5	1,5,6
28	Fab LM	6	1,6,8	0	na	5	2,3,4
34	**Obt**	**7**	**1,5,6,10**	**1**	**7**	**11**	**1,5,6,10**
48	Obt	15	1,10,12	8	1,3,12	5	1,12

artefacts or contamination of the DNA, but indeed these are real haplotypes, perhaps due to a gene duplication. Contamination of sperm in females is also very unlikely since a small piece of foot tissue was consistently used for extraction of DNA. If two gene copies evolve independently following a duplication, the longer time since the gene duplication event the more diverged they are expected to be (due to accumulation of mutations). Because of the relatively small nucleotide diversity in the total data set, there is reason to believe that this putative gene duplication is relatively young perhaps even unique to this taxon. Based on the variation in the LM-ecotype and *L. obtusata* in contrast to the SS-ecotype, the gene duplication is nevertheless very likely to predate the inversion and it would therefore be natural to assume that two gene copies are present in the inverted chromosome as well. In light of this it is even more interesting why the SS-ecotype was almost completely monomorphic for one haplotype. Perhaps the linkage between *Ark* and size and the possible gene duplication are not independent of each other but in what way they could be related will have to await further studies.

CONCLUSIONS AND FUTURE PROSPECTS

In **paper I** we have shown that despite *L. fabalis* and *L. obtusta* being different in morphology, ecology, several different nuclear DNA markers and that no contemporary hybridisation between them is likely to occur, they are not diverged in the mitochondrial cyt-b gene. This could be either due to incomplete lineage sorting but a more likely scenario is however that hybridisation between them has occurred which has allowed mitochondrial DNA to introgress between the species boundary. One explanation could therefore be that there has been a replacement of haplotypes from one species to the other and the only evidence of an early divergence between *L. fabalis* and *L. obtusata* (from the cyt-b haplotype network) is the split between the Spanish clade and the main haplotype network, but several alternative explanations exist that cannot be ruled out. Although this study cannot give any definitive answers to the lack of divergence in cyt-b between *L. fabalis* and *L. obtusata*, it complements the already existing literature of mtDNA introgressions such that we now know that this phenomenon is not only restricted to local populations (or otherwise geographically restricted species) and to species where there is evidence of contemporary hybridisation also from nuclear markers and/or morphology.

I only sequenced 350 base pairs of the cyt-b gene since the focus of **paper I** was to first of all demonstrate that the lack of divergence in this gene is universal for the whole pan-European distribution range. Of course, the more nucleotides that would be

sequenced the fewer haplotypes would in general be shared between individuals including between *L. fabalis* and *L. obtusata*. But the patterns of polyphyly would still be indicative of hybridisation and a larger sequence would enable me to more accurately date the suspected introgressions. In the future I would therefore like to focus on one or a few locations where introgression is most likely to have occurred (e.g. Bergen, Norway) and sequence fewer individuals but much larger blocks of the mitochondrial DNA molecule.

The overall genetic variation in a gene is primarily related to the effective population size (larger populations have more variation) but also the mutation rate (higher mutation rate gives more variation). Given some assumptions, the effective population size of nuclear genes is on average four times higher than in mitochondrial DNA and since the nucleotide diversity in cyt-b is about two times higher than in the *Ark* intron sequence it is clear that the mutation rate in the *Ark* intron is much lower than in the cyt-b (if the mutation rates would be the same we would expect the *Ark* intron to show four times more variation as the cyt-b gene). In addition, as already mentioned the higher effective population size of nuclear genes will slow down the lineage sorting process which makes monophyly of the nuclear genotypes less likely in comparison. A lack of divergence in the *Ark* intron is therefore not at all as surprising as in the cyt-b gene. In fact, *L. saxatilis,* which is the closest relative to the flat periwinkles is not noticeably separated from *L. fabalis* and *L. obutsata* either (Fig. 8; in mitochondrial genes they are; Reid 1996). When introgressive hybridisation occurs, only genes that are unlinked to Dobzhansky-Muller incompatibilities are likely to introgress. If introgression has indeed been as common as indicated by the cyt-b data I would therefore expect to find at least some nuclear genes that also show the same characteristics as cyt-b. In light of the data from *Ark* (and some preliminary data from an intron of calmodulin) I do not however hold much promise to such an approach as it is difficult to find nuclear genes that evolve fast enough such that a difference between *L. fabalis* and *L. obtusata* can be found. Instead in future studies I would like to focus on AFLP's (Amplified Fragment Length Polymorphism) from which a high number of bi-allelic markers that represent the whole genome will be available and from this it would be interesting to study how much of the genome has been restricted from gene flow and how many regions show signs of introgression.

In **paper II** we have shown that the size difference between the LM-ecotype and the SS-ecotype of *L. fabalis* is maintained by an interaction of four factors: *i*) life history optimisation selecting for early maturation and small size, *ii*) crab predation, selecting for large size *iii*) fucoid algae functioning as refuges from crab predation and *iv*) wave-induced dislodgement, reducing the effects of the refuge function in moderately exposed habitats. Size in sheltered habitats are instead expected to be small due to life history reasons, which is supported by the fact that that in the White Sea, where crabs are absent, snails are small in both habitats. The two ecotypes of *L. fabalis* differ also in *Ark* and the sequencing of an intron of *Ark* revealed that the SS-ecotype is nearly fixed for one haplotype while the diversity among LM-ecotypes is much higher, which supports a scenario where a recently derived chromosomal rearrangement, possibly an inversion, restricts recombination between *Ark* and one or several loci that influence size. I would like to (once again) stress that we cannot be sure that linkage between *Ark* and size indeed is due to an inversion or even a chromosomal rearrangement but the patterns we see are difficult explain otherwise (but see Butlin 1995 for and additional view). Numerous studies support the idea that chromosomal rearrangements play causal roles in speciation but they are predominantly conducted in closely related species (e.g. Noor 2001a, Brown et al 2004) where it is difficult to know what role the rearrangements played during the initial divergence. As pointed out by Coyne and Orr (2004, p.60), "the

isolation barriers currently most important in restricting gene flow are not necessarily those historically most important during speciation". There is no doubt that the divergence between populations may be facilitated by occurrence of already fixed chromosomal rearrangements (acting as partial gene flow barriers) but not until the model by Kirkpatrick and Barton (2005) was the initial dynamics of chromosomal rearrangements rigourously addressed and they suggested that an inversion may be positively selected if it protects locally adapted alleles at two or more loci (as indeed could to be the case in *L. fabalis*). What makes the ecotype linked inversion polymorphism in *L. fabalis* particularly appealing is that we have no reason to suspect that any genic or chromosomal differences other than the putative inversion exist and we can specifically study the effects of this phenomenon without the confounding effects of other genomic differences. So far the *Ark* substructuring has been shown to be most pronounced in areas where there is also a clear differences in snail size between exposed and sheltered locations (UK, Sweden, South Norway, North France) but some *Ark* substructuring occurs also in the White Sea (although this is different from the other locations) where such a size difference does not exist. In addition, there are some intriguing details around the LM- and SS-ecotypes in Bergen (Norway) as both size and *Ark* strongly correlate with wave exposure but still appear not to be genetically linked. The correlation with the inversion and the specific selective regimes in respective geographic location will hopefully not only allow us to test explicit predictions about the role of this particular inversion for the evolution of local adaptation in *L. fabalis*, but also to draw conclusions about the role of inversions at the initial stages of speciation in general. To study this further we need however a definitive way to identify the linkage block where *Ark* and the gene(s) for size are situated and since cytological staining of the chromosomes is very difficult, a multi locus approach such as AFLPs' would probably be the most efficient approach.

AFLP is based on a relatively simple methodology that can generate large amounts of data. As AFLP's would be the most straightforward way to study the divergence between *L. fabalis* and *L. obtusata* as well as the inversion hypothesis I hope this can be my primary focus in the near future.

References

Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217-1229.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Harward University Press, Cambridge, Massachusetts.

Ayala, F. J., and J. A. Kiger. 1980. Modern genetics. Benjamin/Cummings publishing company, California.

Bachtrog, D., K. Thornton, A. Clark, and P. Andolfatto. 2006. Extensive introgression of mitochondrial DNA relative to nuclear genes in the *Drosophila yakuba* species group. Evolution 60:292-302.

Barton, N., and O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. Heredity 56:357-376.

Bazin, E., S. Glemin, and N. Galtier. 2006a. Mitochondrial DNA and population size - Response. Science 314:1390-1390.

Bazin, E., S. Glemin, and N. Galtier. 2006b. Population size does not influence mitochondrial genetic diversity in animals. Science 312:570-572.

Berry, O. F. 2006. Mitochondrial DNA and population size. Science 314:1388-1388.

Berthier, P., L. Excoffier, and M. Ruedi. 2006. Recurrent replacement of mtDNA and cryptic hybridization between two sibling bat species *Myotis myotis* and *Myotis blythii*. Proc. R. Soc. London, Ser. B 273:3101-3109.

Blanckenhorn, W. U. 2000. The evolution of body size: What keeps organisms small? Q. Rev. Biol.75 (4): 385-407

Boulding, E. G., M. Holst, and V. Pilon. 1999. Changes in selection in gastropod shell size and thickness with wave-exposure in Northeastern Pacific shores. J. exp. mar. Biol. Ecol. 232:Rochette.

Boulding, E. G., and K. L. Van Alstyne. 1993. Mechanisms of differential survival and growth of two species of *Littorina* on wave-exposed and protected shores. J. exp. mar. Biol. Ecol. 169:139-166.

Brown, K. M., L. M. Burk, L. M. Henagan, and M. A. F. Noor. 2004. A test of the chromosomal rearrangement model of speciation in Drosophila pseudoobscura. Evolution 58:1856-1860.

Butlin, R. K. 2005. Recombination and speciation. Mol. Ecol. 14:2621-2635.

Carson, E. W., and T. E. Dowling. 2006. Influence of hydrogeographic history and hybridization on the distribution of genetic variation in the pupfishes *Cyprinodon atrorus* and *C. bifasciatus*. Mol. Ecol. 15:667-679.

Chan, K. M. A., and S. A. Levin. 2005. Leaky prezygotic isolation and porous genomes: Rapid introgression of maternally inherited DNA. Evolution 59:720-729.

Coyne, J. A., Meyers, J. A., Crittenden, A. P., Sniegowski. P. 1993. The fertility effects of pericentric inversions in *Drosophila-Melanogaster*. Genetics 134 (2): 487-496

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc, Massachusetts.

Darwin, C. 1859. The origin of species. Science:1449-52.

de Queiros, K. 1998. The generak lineage concept of species, species criteria, and the process of speciation *in* J. D. Howard and S. H. Berlocher, eds. endless forms. Oxford univrsity press, New York.

Denny, M. W., T. L. Daniel, and M. A. R. Koehl. 1985. Mechanical limits to size in wave-swept organisms. Ecol. Mon. 55:69-102.

Donnelly, M. J., J. Pinto, R. Girod, N. J. Besansky, and T. Lehmann. 2004. Revisiting the role of introgression vs shared ancestral polymorphisms as key processes shaping genetic diversity in the recently separated sibling species of the Anopheles gambiae complex. Heredity 92:61-68.

Elner, R. W., and D. G. Raffaelli. 1980. Interactions between two marine snails *Littorina rudis* Maton and *Littorina nigrolineata* Grey, a predator *Carcinus maenas* (L.) and a parasite, *Microphallus similis* Jägerskiold. J. Exp. Mar. Biol. Ecol. 43:151-160.

Endler, J. A. 1977. Geographic variation, speciation and clines. Princeton Univ. Press., Princeton.

Feder, J. L., S. H. Berlocher, J. B. Roethele, H. Dambroski, J. J. Smith, W. L. Perry, V. Gavrilovic, K. E. Filchak, J. Rull, and M. Aluja. 2003a. Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. Proc. Natl. Acad. Sci. U. S. A. 100:10314-10319.

Feder, J. L., F. B. Roethele, K. Filchak, J. Niedbalski, and J. Romero-Severson. 2003b. Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. Genetics 163:939-953.

Filchak, K. E., J. B. Roethele, and J. L. Feder. 2000. Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. Nature 407:739-742.

Funk, D. J., and K. E. Omland. 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34:397-423.
Goodwin, B. J., and J. D. Fish. 1977. Inter- and intraspecific variation in *Littorina obtusata* and *L. mariae* (Gastropoda: Prosobranchia). J. Moll. Stud. 42:241-254.
Hartl, D. L., and A. G. Clark. 1997. Principles of population genetics. Sinauer Associates, Inc, Sunderland, Massachusetts.
Hedrick, P. W. 1981. The establishment of chromosmal variants. Evolution 35:322-332.
Hedrick, P. W. 1986. Genetic polymorphism in heterogeneous environments: a decade later. Ann. Rev. Ecol. Syst. 17:535-566.
Hedrick, P. W., M. E. Ginevan, and E. P. Ewing. 1976. Genetic polymorphsm in heterogeneous environments. Ann. Rev. Ecol. Syst. 7:1-32.
Hendry, A. P., E. B. Taylor, and J. D. McPhail. 2002. Adaptive divergence and the balance between selection and gene flow: Lake and stream stickleback in the misty system. Evolution 56:1199-1216.
Janson, K. 1983. Chromsome number in two pheotypically distinct populations of *Littorina saxatilis* Olivi, and in specimens of the *Littorina obtusata* (L.) species-complex. J. Moll. Stud. 49:224-227.
Johannesson, B. 1986. Shell morphology of Littorina saxatilis Olivi: the relative importance of physical factors and predation. J. Exo. Mar. Biol. Ecol. 102:183-195.
Johannesson, K. 2001. Parallel speciation: a key to sympatric divergence. Trends Ecol. Evol. 16:148-153.
Johannesson, K. 2003. Evolution in *Littorina*: ecology matters. J. Sea Res. 49:107-117.
Johannesson, K., Johannesson, B., Rolan Alvarez, E. 1993. Morphological differentiation and genetic cohesiveness over a microenvironmental gradient in the marine snail *Littorina saxatilis.* Evolution 47 (6): 1770-1787
Johannesson, K., and N. Mikhailova. 2004. Habitat-related genetic substructuring in a marine snail (*Littorina fabalis*) involving a tight link between an allozyme and a DNA locus. Biol. J. Linnean Soc. 81:301-306.
King, M. 1987. Chromosomal rearrangements, speciation and the theoretical approach. Heredity 59:1-6.
King, M. 1992. A dual level model for speciation by multiple pericentric inversions. 68:437-440.
King, M. 1993. Species evolution - the role of chormosomal change. Cambridge University Press, New York.
Kirkpatrick, M., and N. Barton. 2006. Chromosome inversions, local adaptation and speciation. Genetics 173:419-434.
Lande, R., and J. Arnold. 1985. Evolution of mating preference and sexual dimorphism. J. Theor. Biol. 117:651-664.
Levins, R. 1968. Evolution in changing environments. Princeton Univ. Press., Princeton, NJ.
Machado, C. A., T. S. Haselkorn, and M. A. F. Noor. 2007. Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis*. Genetics 175:1289-1306.
Machado, C. A., R. M. Kliman, J. A. Markert, and J. Hey. 2002. Inferring the history of speciation from multilocus DNA sequence data: The case of *Drosophila pseudoobscura* and close relatives. Mol. Biol. Evol. 19:472-488.
Marincovich, L., and A. Y. Gladenkov. 1999. Evidence for an early opening of the Bering Strait. Nature 397:149-151.

Martinsen, G. D., T. G. Whitham, R. J. Turek, and P. Keim. 2001. Hybrid populations selectively filter gene introgression between species. Evolution 55:1325-1335.
Melo-Ferreira, J., P. Boursot, F. Suchentrunk, N. Ferrand, and P. C. Alves. 2005. Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. Mol. Ecol. 14:2459-2464.
Miro, R., C. Fuster, I. C. Clemente, M. R. Caballin, and J. Egozcue. 1992. Chromosome inversions involved in the chromosome evolution of the hominidae and in human constitutional chromosome-abnormalities. J. Hum. Evol. 22:19-22.
Moore, W. S. 1995. Inferring Phylogenies from Mitochondrial DNA Variation - mitochondrial-gene trees versus nuclear-gene Trees. Evolution 49:718-726.
Morando, M., L. J. Avila, J. Baker, and J. W. Sites. 2004. Phylogeny and phylogeography of the *Liolaemus darwinii* complex (Squamata: Liolaemidae): Evidence for introgression and incomplete lineage sorting. Evolution 58:842-861.
Morjan, C. L., and L. H. Rieseberg. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Mol. Ecol. 13:1341-1356.
Munneke, L. R., and G. E. Collier. 1988. Cytoplasmic and mitochondrial arginine kinases in *Drosophila* - Evidence for a single gene. Biochem. Genet. 26:131-141.
Navarro, A., and N. H. Barton. 2003. Chromosomal speciation and molecular divergence - Accelerated evolution in rearranged chromosomes. Science 300:321-324.
Nevo, E. 1998. Molecular evolution and ecological stress at global, regional and local scales: The Israeli perspective. J. Exp. Zool. 282:95-119.
Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: A Markov chain Monte Carlo approach. Genetics 158:885-896.
Noor, M. A. F., K. L. Grams, L. A. Bertucci, Y. Almendarez, J. Reiland, and K. R. Smith. 2001a. The genetics of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. Evolution 55:512-521.
Noor, M. A. F., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001b. Chromosomal inversions and the reproductive isolation of species. Proc. Natl. Acad. Sci. U. S. A. 98:12084-12088.
Norton, T. A., S. J. Hawkins, N. L. Manley, G. A. Williams, and D. C. Watson. 1990. Scraping a Living - a Review of Littorinid Grazing. Hydrobiologia 193:117-138.
Panithanarak, T., H. C. Hauffe, J. F. Dallas, A. Glover, R. G. Ward, and J. B. Searle. 2004. Linkage-dependent gene flow in a house mouse chromosomal hybrid zone. Evolution 58:184-192.
Peters, J. L., Y. Zhuravlev, I. Fefelov, A. Logie, and K. E. Omland. 2007. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (Anas spp.). 61:1992-2006.
Piry, S., A. Alapetite, J. M. Cornuet, D. Paetkau, L. Baudouin, and A. Estoup. 2004. GENECLASS2: A software for genetic assignment and first-generation migrant detection. 95:536-539.
Reid, D. G. 1990. Trans-Arctic Migration and Speciation Induced by Climatic-Change - the Biogeography of *Littorina* (*Mollusca*, *Gastropoda*). Bull. Mar. Sci. 47:35-49.
Reid, D. G. 1996. Systematics and evolution of *Littorina*. Ray Society, London.

Reimchen, T. E. 1981. microgeographical variation in *Littorina mariae* Sacci & Rastelli and a taxonimic consideration. J. Conch 30:341-350.
Reimchen, T. E. 1982. Shell size divergence in *Littorina mariae* and *L. obtusata* and predation by crabs. Can. J. Zool. 60:687-695.
Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. Trends Ecol. Evol. 16:351-358.
Rieseberg, L. H., and K. Livingstone. 2003. Evolution - Chromosomal speciation in primates. Science 300:267-268.
Rieseberg, L. H., J. Whitton, and K. Gardner. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics 152:713-727.
Roca, A. L., N. Georgiadis, and S. J. O'Brien. 2005. Cytonuclear genomic dissociation in African elephant species. Nature Gen 37:96-100.
Ruber, L., A. Meyer, C. Sturmbauer, and E. Verheyen. 2001. Population structure in two sympatric species of the Lake Tanganyika cichlid tribe *Eretmodini*: evidence for introgression. Mol. Ecol. 10:1207-1225.
Schluter, D. 1998. Ecological causes of speciation *in* J. D. Howard and S. H. Berlocher, eds. Endless forms. Oxford University press, New York.
Schluter, D. 2001. Ecology and the origin of species. Trends Ecol. Evol. 16:372-380.
Schmidt, P. S., M. Phifer-Rixey, G. M. Taylor, and J. Christner. 2007. Genetic heterogeneity among intertidal habitats in the flat periwinkle, *Littorina obtusata*. Mol. Ecol. 16:2393-2404.
Servedio, M. R. 2001. Beyond reinforcement: The evolution of premating isolation by direct selection on preferences and postmating, prezygotic incompatibilities. Evolution 55:1909-1920.
Servedio, M. R., and M. A. F. Noor. 2003. The role of reinforcement in speciation: Theory and data. Annu. Rev. Ecol. Evol. Syst. 34:339-364.
Sites, J. W., and C. Moritz. 1987. Chromosomal evolution and speciation revisited. Syst Zool 36:153-174.
Spirito, F. 1993. The establishment of underdominant chromosomal rearrangements in multi-deme systems with local extinction and colonization. Theor. Pop. Biol. 44:80-94.
Spirito, F. 1998. The role of chromosomal rearrangements in speciation *in* J. D. Howard and S. H. Berlocher, eds. Endless Forms.
Suzuki, T., N. Inoue, T. Higashi, R. Mizobuchi, N. Sugimura, K. Yokouchi, and T. Furukohri. 2000. Gastropod arginine kinases from *Cellana grata* and *Aplysia kurodai*. Isolation and cDNA-derived amino acid sequences. Comp. Biochem. Physiol. B-Biochem. Mol. Biol. 127:505-512.
Tatarenkov, A., and K. Johannesson. 1994. Habitat related allozyme variation in a microgeographical scale in the marine snail *Littorina mariae* (Prosobranchia: Littorinacea). Biol. J. Linn. Soc. 53:105-125.
Tatarenkov, A., and K. Johannesson. 1998. Evidence of a reproductive barrier between two forms of marine periwinkle *Littorina fabalis* (Gastropoda). Biol. J. Linn. Soc. 63:349-365.
Tatarenkov, A., and K. Johannesson. 1999. Micro- and macrogeographical allozyme variation in *Littorina fabalis*; do sheltered and exposed forms hybridize? Biol. J. Linn. Soc. 67:199-212.
Tatarenkov, A. N. 1995. Genetic-Divergence, between Sibling Species *Littorina mariae* Sacchi and Rastelli and *Littorina obtusata* (L) (Mollusca, Gastropoda) from the White-Sea. Ophelia 40:207-218.

Templeton, A. R. 1998. Species and speciation *in* J. D. Howard and S. H. Berlocher, eds. Endless forms. Oxford university press, New York.
Tregenza, T., and R. K. Butlin. 1999. Speciation without isolation. 400:513-513.
Trussell, G. C., A. S. Johnson, S. G. Rudolph, and E. S. Gilfillan. 1993. Resistance to dislodgment - Habitat and size-specific differences in morphology and tenacity in an intertidal snail. Mar. Ecol.-Prog. Ser. 100:135-144.
Underwood, A. J., and K. E. McFadyen. 1983. Ecology of the intertidal snail *Littorina acutispera* Smith. J. Exp. Mar. Biol. Ecol. 66:169-197.
Walsh, J. B. 1982. Rate of accumulation of reproductive isolation by chromosome rearrangements. Am. Nat. 120:510-532.
Wares, J. P., and C. W. Cunningham. 2001. Phylogeography and historical ecology of the North Atlantic intertidal. Evolution 55:2455-2469.
White, M. J. D. 1978. Modes of speciation. W.H. Freeman and Company, San Fransisco.
Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. Trends Ecol. Evol. 16:381-390.
Via, S. 2002. The ecological genetics of speciation. Am. Nat. 159:1-7
Williams, G. A. 1990. *Littorina mariae* - a factor structuring low shore communities? Hydrobiol 193:139-146.
Wilson, C. C., and L. Bernatchez. 1998. The ghost of hybrids past: fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). Mol. Ecol. 7:127-132.
Wright, S. 1931. Evolution in mendelian populations. Genetics 16:97-159.
Wright, S. 1938. Size of population and breeding structure in relation to evolution. Science 87:430-431.
Wyss, M., J. Smeitink, R. A. Wevers, and T. Wallimann. 1992. Mitochondrial creatine-kinase - a key enzyme of aerobic energy-metabolism. Biochim Biophys Acta 1102:119-166.
Zaslavskaya, N. I., S. O. Sergievsky, and A. N. Tatarenkov. 1992. Allozyme similarity of Atlantic and Pacific Species of *Littorina* (Gastropoda, Littorinidae). J. Moll. Stud. 58:377-384.

ACKNOWLEDGEMENTS

Kerstin Johannesson, my supervisor, is truly one of the most positive and encouraging people I have ever met. I first met her as a teacher during my undergraduate studies and I passed on an essay to her in case she would be interested in reading it and perhaps giving some comments on it. I did not expect much and I was highly surprised when she took contact and gave me positive feedback of a kind I had never received from anyone before. I think this was crucial for my decision to do my degree project for her on the ecology of *L. fabalis* despite the fact that periwinkles did not initially strike me as any interesting study organisms (I was at that time mostly interested in salmonid fish). After my degree project I realised however that my interest in biology was not tied to any specific organisms but rather to the questions that could be studied with them and I was very fortunate when Kerstin finally found financing for my PhD project. I have always enjoyed my discussions with Kerstin and every time I pop my head through the door to her office she seems to have time for me. Her belief that everything, no matter how complicated, can be explained in the simplest of terms is something that I really have tried adopting in my own scientific work.

Although Andrei Tatareknov (who was supposed to be my assistant supervisor) left for California almost immediately after I started my PhD studies, he has been a great help for me during times I really needed it. He was the one who, together with Kerstin and Natalia Mikhailova did all the pioneering work with the *Ark* substructuring in *L. fabalis,* without which my work could not have been possible. In addition, the idea that the linkage between size and *Ark* could be due to an inversion was initially Andreis.

I would also like to thank John Graham at Leeds University and his undergraduate student Celina Bio~Gatta who initially sequenced the *Ark* intron, which I could use in my further studies. Without this it would have taken me much longer to figure out how to sequence anything from the *Ark* gene. Thank you John also for inviting me to Leeds and use your BAC-library. Although none of that work is presented in this Thesis, it gave me lots of experience and some ideas that were very valuable. Above all I really enjoyed your hospitality.

During my PhD studies I have been unfortunate since I have not had a consistent assistant supervisor but there are however many people that still deserve lots of credit: Thomas Dahlgren, Pieternella Luttikhuizen, Carl André, Fredrik Pleijel and Carles Vilá. I would especially like to thank Thomas whom I have always been able to call whenever things have got really desperate and Pieternella who I learned a lot from. Me and Carl share the same interest for outdoor activities and he deserves a special thanks for introducing me to Nordic ice-skating.

I would like to thank Marina Panova for her positive attitude and all the help during my last years of PhD studies. Without her help with microsatellites it would have been a struggle. Thank you also Mikhail Foikin for giving me access to the allozyme data you have on the Swedish and White sea *L. fabalis* and Johan Hollander for helping me with the morphometrics in **paper I**.

During my work I have been involved in some teaching which I have enjoyed a lot. I would therefore like thank all my students especially my graduate students; Tobias Lindskog, who did a substantial amount of the work in **paper IV**; Karl Larsson, for his sparkling personality and Lisa Sundqvist, for impressing me with her mathematical skills.

At least when it comes to science I have lots of ideas and many people I have encountered during my career have been (un)fortunate in having no choice but to listen to them. I would therefore like to thank the following people for their patience: Erik Boström (my flat mate who probably has suffered the most), Mikael Dahl (my office partner), Emman Vodoti, Anders Eriksson, Serik Sagitov and not the least Ricardo Pereyra. A special thanks to goes also to Per Nilsson, Per Jonsson and Mats Lindegarth - the statistics Mafia at Tjärnö. And thank you Benno for putting out the fires after me in the Lab.

Last of all I would like to thank my mother, father and sister for all the support and all my friends for giving me lots of positive energy.

This thesis was funded by the Swedish Research council, Kungliga Fysiografiska sällskapet i Lund, Helge Axelsson Johnssons stiftelse, Wilhelm och Martina Lundgrens Vetenskapsfond, Rådman och Fru Ernst Colliander stiftelse and Filosofiska fakulteternas gemensamma donations-nämnd.