Biodiversity and Genetic Patterns in Marine Invertebrates

Emma Thuróczy Vodoti

UNIVERSITY OF GOTHENBURG
Department of Zoology / Systematics and Biodiversity
Faculty of Science
Sweden, 2010

ISBN 978-91-628-7913-6
Printed by Intellecta Infolog, Göteborg, Sweden 2010
Cover photo: Emma Thuróczy Vodoti

Thuróczy Vodoti E. 2010.
Biodiversity and genetic patterns in marine invertebrates.

Department of Zoology, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden

Abstract:

Systematics based on genetic data has both confirmed and contradicted earlier, morphologically defined species and their relatedness. Morphology does not always correspond to genetic lineages, and this will inevitably affect both traditional systematics as well as biodiversity assessments. My thesis aims to investigate genetic biodiversity in some marine invertebrates, dealing with both species and population (intraspecific) relationships. I discuss gene lineage relatedness in correlation to species morphs and geographical species distribution, which has bearing upon the species recognition problem, the barcoding approach and the meaning of phylogeographic patterns.

Results show that morphology does not generally reflect gene lineage relationships in the considered species and that cryptic species are common, mainly in the nemertean groups studied (*Oerstedia dorsalis*; *Cerebratulus* spp). Despite having a relatively established status as a species, the horse mussel (*Modiolus modiolus*) is also found to hold cryptic lineages. Further, the difficulties in species delimitation become apparent, as an almost continuous range of genetic divergence prevails between most of the found clades. The Antarctic krill (*Euphausia superba*) is the only study species lacking cryptisism, at least in the samples and geographical region studied. The intraspecific genealogies were in this case used to estimate population parameters by means of the coalescent. Despite being an abundant species of modest size, results suggest a low effective population size and a population with common ancestry, most likely capable of migrating over relatively great distances. This stresses that factors such as demography and genetics can provide a preliminary, general base for management issues.

I conclude that cryptic species are common, and therefore taxonomic work cannot be isolated from the barcoding quest of summoning genetic data for species identification. Since genetic relationships and morphological traits do not always go hand in hand, a barcoding approach could in some cases be misleading. Further, it is probably impossible to find a universal way of defining, identifying and delimiting species. This thesis illustrates some practical examples of these problems, and suggests that a case-to-case evaluation is likely needed in future taxonomic and phylogenetic efforts.

Keywords: *biodiversity, molecular genetics, intraspecific patterns, morpho-species, genealogy, marine invertebrates, phylogeography, haplotype network*

ISBN 978-91-628-7913-6

Svensk sammanfattning

Systematik baserat på genetiska data har både bekräftat och stridit mot tidigare, morfologiskt definierade arter och deras släktskap. Morfologi motsvarar inte alltid genetiska släktlinjer, vilket påverkar både traditionell systematik och biodiversitetsskattningar. Min avhandling berör genetisk biodvierstiet i några marina ryggradslösa djur och tar upp släktskapet mellan både arter och populationer (inomartssläktskap). Jag diskuterar genetiska släktlinjer i relation till artmorfer och geografisk artutbredning. Detta är kopplat till problemet med att känna igen en art Barcoding metoden samt betydelsen av fylogeografiska mönster.

Resultaten visar att morfologi generellt inte avspeglar genetiska släktlinjer i de studerade arterna och att kryptiska arter är vanliga, särskilt i de undersökta nemertingrupperna (*Oerstedia dorsalis*; *Cerebratulus* spp). Trots sin sedan länge etablerade status som ”god” biologisk art, så visade sig även den vanliga hästmusslan (*Modiolus modiolus*) bestå av kryptiska artgrupper. Dessutom blir problematiken med artavgränsning uppenbar, eftersom de flesta av de funna genetiska grupperna, eller kladerna, visade en mer eller mindre kontinuerlig grad av genetisk diversitet utan tydliga gränser. Den Antarktiska krillen (*Euphausia superba*) visade sig vara den enda av de undersökta djurgrupperna som inte bestod av kryptiska arter, åtminstone inte i de individer och geografiska områden som studien omfattade. De genetiska släktskapslinjerna inom arten Antarktisk krill användes för att uppskatta populationsparametrar med hjälp av coalescentmetoder, d v s släktskapslinjers samband och deras koppling till vad som skett i populationen historiskt sett. Trots att krill förekommer i stort antal och är en liten art, så visar resultaten på en låg effektiv populationsstorlek och ett gemensamt ursprung för populationen med förmåga att aktivt förflytta sig över stora områden. Resultaten betonar vikten av att ta hänsyn till demografiska och genetiska faktorer, inte bara biomassa, när man bedömer nivån av utnyttjandet av biologiska resurser.

Sammanfattningsvis bedömer jag att kryptiska arter är vanliga, och att taxonomiskt arbete inte kan isoleras från Barcoding-metoden, vars uppgift är att samla data för artidentifiering. Eftersom genetiska släktlinjer inte alltid går hand i hand med arters morfologi så kan Barcoding-metoden vara vilseledande. Dessutom är det troligtvis omöjligt att hitta ett universellt sätt att definiera, identifiera och avgränsa arter. Denna avhandling illustrerar några praktiska exempel på dessa problem, och föreslår att en enskild bedömning från fall till fall troligtvis kommer att behövas i framtida taxonomi och fylogenetiska studier.

LIST OF PAPERS[1]

This thesis is based in the following papers, referred to in the text by their Roman numerals:

I. SUNDBERG, P., VODOTI THURÓCZY, E., ZHOU, H., STRAND, M. (2009) Species limits in ribbon worms (phylum Nemertea) - genetic and morphological polymorphism in the hoplonemertean *Oerstedia dorsalis*. *Biological Journal of the Linnean Society* 98: 556-567.

II. SUNDBERG, P., VODOTI THURÓCZY, E., STRAND, M. (2009) DNA barcoding needs taxonomy – the case of *Cerebratulus* spp (Nemertea). *Molecular Ecology Resources* 10: 274-281.

III. VODOTI THURÓCZY, E., HALANYCH, K.M., SUNDBERG, P., DAHLGREN, T. (manuscript) Phylogeography of the horse mussel *Modiolus modiolus*.

IV. VODOTI THURÓCZY, E., BERGSTRÖM, B. (manuscript) Antarctic krill genealogy reveals a small effective population size and a history of high mobility in the Atlantic sector of the Southern Ocean.

[1]A doctoral thesis at a Swedish university is often presented as a collection of papers, with a summarising introductory part that comprises the formal thesis. The papers have either already been published or are manuscripts at various stages (in press, submitted or ms).

CONTENTS

AIMS OF THE THESIS

Population genetic processes are the evolution of gene lineages within species lineages, and they can provide crucial lines of evidence for understanding intrinsic genetic patterns of what we call species, as well as means to delimit these species. This thesis aims to investigate the genetic patterns and divergence of gene lineages in species hitherto only recognised on diagnosable morphological grounds. The correlation between morphology and genetic relatedness is explored in some taxonomically well known marine ribbon worms (**Paper I, II**) with the intention of adding genetic data to the process of species identification, and propel the revision of taxa where needed. I further investigate the genetic biodiversity and phylogeographic patterns in two relatively well-established taxonomic species, the horse mussel (**Paper III**) and the Antarctic krill (**Paper IV**). In the exploited Antarctic krill, I study the historical population structure and the effective population size in the Southern Ocean. In the horse mussel, I assess whether patterns of genetic population differentiation exist in the North Atlantic Ocean, and if so, use these to make inferences about how paleoclimate and oceanographic or demographic factors might have shaped populations. This in turn can be used as indicators of climate change effects.

1. INTRODUCTION AND BACKGROUND

1.1 Morphology, systematics and biodiversity

Systematics is one of the oldest disciplines of biological sciences, today classifying organisms into groups based on common shared-derived traits (Hennig 1965). This attempt to define and group all living organisms into hierarchical, taxonomically defined entities has been complemented with modern genetic techniques, analyses and increased computational power (see e.g. Felsenstein 1981; Swofford 1998; Lewis 2001; Nylander *et al.* 2004). However, it turns out that molecular relationships are quite often at odds with the classical, morphology based systematics, especially so at the "lower" levels (e.g. species and genus) in a wide range of organisms (Hundsdoerfer & Wink 2006; Lee *et al.* 2007; Flagel *et al.* 2008; Plattner *et al.* 2009). Considering the fact that all (species) phylogenies have been based solely on morphological characters until just roughly 20 years ago, this might not come as a great surprise. Consequently, our understanding of the biological world around us has increased as genetic information reveals earlier hidden patterns.

As data have compiled, the range of genetic divergence within and between organism groups has been found to overlap (Knowlton 1993; Wiemers & Fiedler 2007; Lou & Golding 2010). To some extent, the different results depend on genetic markers and methodology, but also illustrate the limited knowledge about species in the sea. An ironic twist of the scientific development within the field is consequently that a species is still difficult to universally define and delimit. This is of course an impediment to species identification, since at some point one has to decide what to do with non-

identical DNA sequences – at what level of genetic divergence should the border between different species be drawn? And will this border be universally applicable to all, or at least closely related organism groups? Biodiversity for instance, is traditionally associated with the actual number of species (species count) or taxonomic surrogacy, i.e. letting the number of higher ranked taxonomic units representing the species) (Bertrand *et al.* 2006). Hence, the problem of defining and delimit species is clearly also affecting our progress in conservation biology issues. The effects might be under- or overestimations of biodiversity, with subsequently inappropriate actions taken to conserve it. Needless to say, the way we define, identify and delimit species is central here (Maurer 2000; Mace 2004; Leonard *et al.* 2006).

The well over 20 currently used species concepts have not made the picture any clearer (Coyne & Orr 2004). One obstruction is the range of definitions, and that the criteria presented in the different concepts are seen as i) necessary properties of species and ii) they are based on different biological characteristics *depending on the interests of the scientist* (de Queiroz 2005, de Queiroz 2007). The major evolutionary forces that shape organisms, or groups of organisms, are recognised as genetic drift, natural selection, migration and mutation (Futuyma 1998). Naturally, they will sustain an endless amount of concerting expressions in an organism at all levels, from genetics to physiology and behaviour. It seems clear to me that, logically, neither of these processes or affected properties can be valued as more or less "true" or "speciating" than another, and they cannot be confined to happen in any particular order or during a certain period of time. This might seem an overly obvious statement, but nevertheless, it is these confinements that advocates of different species concepts are trying to demarcate. Recently, de Quieroz (2005, 2007) proposed a merging outlook on how species concepts can be used in a less conflicting way, proposing a unified concept where species are recognised as "separately evolving metapopulation lineages" (*the general metapopulation lineage concept of species*). As a base, it uses *the general species concept* and *the general metapopulation concept of species* developed by Ernst Mayr and Sewall Wright, among many others, during the mid and late 1900's in The Modern Evolutionary Synthesis (see de Quieroz 2005 for a more detailed overview on the development of the concepts). In the unified concept, a population can be viewed as an instantaneous cross-section through the metapopulation lineage. The properties of the population changes as time goes by, and at every sampling event, we can find different acquired properties.

All existing species concepts of today provide facts concerning organism group properties, and subsequently lines of evidence concerning the species status of the group. They differ only in the theoretical concept used, the empirical evidence, and the characteristics needed for the population to be considered a new species (for an in-depth discussion on how to delimit *populations*, see Waples & Gaggiotti 2006). A population that at one point has acquired reproductive isolation from other populations (the biological species concept) might not have acquired diagnosable differences yet (the phylogenetic species concept requires a group to be diagnosable with fixed differences in e.g. morphology or genes). However, when sampled at a later stage, the population

might have evolved such diagnosable differences. Ironically, species concepts are sometimes seen as conflicting, but it should not come as a surprise to a biologist that speciation is a gradual and often slow process, not occurring in a particular ranked order set by scientific species concepts.

1.2 Populations, gene trees and species trees

In phylogenetics, the time of speciation (species tree bifurcation) might not reflect the tokogeny, i.e. the intraspecific or population genealogy (for an overview, see Posada & Crandall 2001) (Figure 1). When a *single* specimen represents a morphologically (or genetically) identified species in a phylogram, the intra- and interspecific relationships are virtually hidden. Gene lineage reticulations or divergences, for instance, might exist in unsampled specimens not presented in the phylogram. Tokogenies, then, show these "hidden" non-hierarchical networks of genetic relatedness between both ancestral and recent gene lineages existing simultaneously in natural populations, within the erected species tree. Gene lineages that exist in different (morphological) species but do not differ by any mutations will show as identical haplotypes with a given frequency in a haplotype network. This would not show in a species phylogeny, where prerequisites are fitted to illustrate interspecific evolution (e.g. no recombination, terminal positioning of taxa on a bifurcating tree, no ancestral sequences present) - indeed, the phylogenetic tree is a special case of phylogenetic networks restricted to show bifurcating evolutionary relationships (Huson & Bryant 2006). Turning to Figure 1, we consider the gene lineages marked with asterisks, and how different intraspecific sampling schemes would affect the species tree topology. If a sampled specimen happens to belong to one of the marked gene lineages, incongruence between the gene tree and the morphological species tree will arise. For example, a specimen/individual of species A may be considered a closer relative to species B than to its conspecifics within the A-clade. In contrast, if the marked cryptic lineages are not sampled at all, biodiversity will be underestimated. The need for multispecimen sampling is clearly warranted when inferring species trees, especially when it comes to recent divergences (Knowles & Chan 2008).

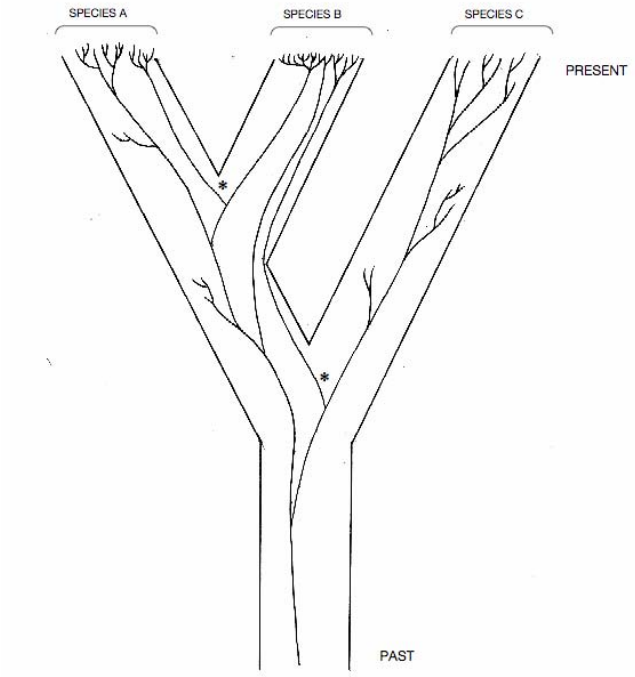

Figure 1. A species tree (thick branches), here based on morphological characters, showing the relationship between species A, B and C. Thin branches show the tokogeny, or intraspecific population gene lineages, that branch into present time specimens (top of figure). Gene lineages that are extinct end within the species tree without reaching present time. * denotes incongruence between the gene tree and the morphology species tree.

Tokogeny approaches show all sampled gene lineages by both frequency and relatedness, independent of the perhaps pre-defined morphological species tree, and thus provides an excellent base for further hypothesis building to infer resolved species trees. In this thesis, we apply the statistical parsimony network method (software TCS; Clement *et al.* 2000), proven to identify a high rate of biological species, as well as cryptic species (Hart & Sunday 2007). Barcoding approaches (e.g. Barcodes of Life Initiative) indirectly depend on how and where to set the limit of genetic discontinuity that can universally identify species (Hebert *et al.* 2003). The 95% connection limit of the statistical parsimony network method has proven useful for this purpose, especially with mtDNA data (Hebert *et al.* 2004) (although see DISCUSSION for critics of the approach). Mainly, it provides an initial overview of the genealogy in a group of organisms identifying general patterns that can propel further studies (i.e. it does not state a *universal* species-delimiting tool). Simply

put, the statistical parsimony method identifies the 95% statistical confidence limit for connecting haplotypes with the maximum number of single substitution differences. From thereon, haplotypes are connected starting with those that differ by one mutational change, then two and so on, until the network contain all haplotypes within the connecting limit. Remaining haplotypes appear as singletons (not connected to the network) or may together form additional networks. For illustration and terms, see e.g. figure 2 in **paper II**.

So, is there any light in the tunnel of grasping what a species is? In this thesis, I do not try to provide a final, general answer to this immense question burdened with philosophical, practical and semantic issues. However, I acknowledge that extended specimen sampling on the intraspecific (population) level, with the intention of depicting gene lineages "ignoring" earlier taxonomic groupings, can indeed reveal disregarded specimen relationships. These can then be used to further resolve issues of species status. In two of the papers, we study the association of morphology with genetic relatedness in the nemertean taxa *Oerstedia dorsalis* and *Cerebratulus* spp. Here, the arbitrariness of nemertean morphological characters as species delimitators/identifyers is examined, and we employ a multiple-specimen sampling scheme to test the reliability of morphology as a morph or species identifier. Apart from casting light on the species definition issues, genetic diversity studies can also show how variations in population dynamics can leave different genetic imprints, and that these can be used to estimate the effective population size and the historical genetic patterns in a population. In the last papers then, we aim to implement this in the horse mussel (*Modiolus modiolus*) and the Antarctic krill (*Euphausia superba*). Due to the subtidal habitat of the horse mussel, refuge areas and recolonization routes might differ from intertidal species. Genetics of subtidal populations might therefore trace different geological and climate events than the genetics of intertidal species. In summary, paleontological studies in general can provide information about climate effects on sampled areas, which might be difficult or impossible to deduce from geological data only (Hewitt 2000). In the case of the Antarctic krill, my approach concerns historical genetic patterns and the effective population size, which can both be used to discuss the population history, mobility between geographical regions, and to some extent provide indications of the population resilience.

3. METHODS

3.1 Study species – an overview

The Antarctic krill - *Euphausia superba*

Antarctic krill is one of the most abundant metazoan species in Antarctica with a biomass of 500 million tons, and it occupies a key position in the food web as a link between primary production and higher trophic levels such as baleen whales, sea birds, fish and squid (Hewitt & Linen Low 2000). The 50 mm long krill has a circum-polar distribution in Antarctica only, where it

readily migrates between ice-covered, near shore over-wintering areas, and open-water spawning areas during summer (Siegel 2000). Antarctic krill is strongly gregarious, which is facilitated by their bioluminescent ability providing a visual que for fellow krill in dark waters, and they have the peculiar habit of arranging themselves into clusters of different size classes (Siegel 2000). Unfortunately, this has also enabled harvesting of big, reproductively mature adults since the late 1970's, when the fishery industries discovered and began to utilise the krill resource. Today, this constitutes one of the biggest fisheries in the world (Hewitt & Linen Low 2000), but thorough genetic studies on the population are scarce. The fishery industry mainly supports its outtake on biomass estimations, but also to some extent on other ecological factors (Hewitt & Linen Low 2000; Hewitt *et al.* 2004).

Species	*Euphausia superba*
Systematics	Phylum Arthropoda; Class Malacostraca; Order Euphausiacea
Range	circum-polar in Antarctica
Habitat	nektonic, 0-3000 m depth
Reproduction	gonochoristic with internal fertilization through deposited sperm plugs, brooding, nauplius larvae and eleven succeeding sub-adult stages

The ribbon worms - *Cerebratulus* spp and *Oerstedia dorsalis*

Phylum Nemertea is a diverse and abundant group, but yet relatively poorly investigated. In spite of their global distribution and great abundance in any marine ecosystem, morphological and histological descriptions are still the main base for modern taxonomists (Gibson 1998), although molecular phylogenetics has entered the field (Thollesson & Norenburg 2003). External characters are, in the best case, few. The first species descriptions from the late 19th and early 20th century are based on histology and morphological characters (e.g. Bürger 1892). Unfortunately, classification based on histology has not provided an ultimate solution to the identification of morphological species, since it is severely inconsistent due to the extreme elasticity of the nemertean body wall, and the subsequent contraction during chemical fixation treatment prior to histological sectioning (Sundberg 1979).

In this thesis we study two genera that are morphologically variable, and which show two opposite aspects of the problem with nemertean taxonomy. The first case is represented by *Oerstedia dorsalis*, a nominal species with great variability in coloration and patterns, displaying a continuum from pale-white specimens to dark purple-brownish ones with iridescent silver spotting. Morph-status has probably been more commonly assigned to specimens with fewer clear-cut differences (e.g. specimens with numerous blurred spots of different size and colour shades). As new morphs are discovered, they can either be recognised as new species, or as new intraspecific morphs, which in many cases is a subjective task (P. Sundberg, pers. comm.). The second group, *Cerebratulus* spp, is morphologically less variable than *O. dorsalis* when it

comes to colour pattern, but these species can have rather distinct morphological differences. These have been used together with histology to establish some species descriptions, but relatively few modern revisions have followed. In summary, traditional nemertean taxonomy is likely to represent several over- and underestimations of the number of existing species.

Species	*Oerstedia dorsalis*
Systematics	Phylum Nemertea; Class Enopla; Order Hoplonemertea
Range	northern hemisphere; intertidal or shallow sublittoral
Habitat	in *Corallina* sp and other algal beds, shells and gravel, algal holdfasts
Reproduction	gonochoristic with direct, or unknown, development

Species	*Cerebratulus* spp
Systematics	Phylum Nemertea; Class Anopla; Order Heteronemertea
Range	global distribution; intertidal and down to depths of 150 m or more
Habitat	clay, sand, shells and gravel, reefs, holdfasts, limestone burrows
Reproduction	gonochoristic with external fertilization (eggs attached to substrate) and planktotrophic pilidium larvae; or reproduction unknown

The horse mussel - *Modiolus modiolus*

This brownish-black bivalve, with resemblance to the common blue mussel, occurs in the Northern hemisphere. The horse mussel is strictly subtidal and does not tolerate desiccation or raised temperatures (Davenport & Kjørsvik 1982). It can form either beds or biogenic reefs, which have been found to host hundreds of other invertebrate species, in addition to its own juveniles (Brown & Seed 1977). The species is thus sedentary, but capable of loosening itself from the substratum. Its ability of dispersal is most likely pronounced in the planktotrophic larval stage, which can last for several months before settling. After reaching a size of about 40 mm, the young individuals become less vulnerable to their main predators, starfish and crabs; adult specimens can reach a length of around 200 mm (Tyler-Walters 2007).

Species	*Modiolus modiolus*
Systematics	Phylum Mollusca; Class Bivalvia; Order Mytiloida
Range	northern hemisphere; subtidal down to 280 m depth

Habitat	all substrates, can form biogenic reefs
Reproduction	gonochoristic with broadcast spawning and planktotrophic larvae

3.2 Specimen sampling and study areas

The specimens used in this thesis were collected either by trawling (krill: Siegel *et al.* 2002; Watkins *et al.* 2004), dredging and snorkelling (horse mussels and ribbon worms) or by collecting algae tufts and turning over boulders in the intertidal zone (ribbon worms). To make sampling of minute ribbon worms possible (e.g. *O. dorsalis*), dredged or manually collected material was left to deoxygenate in a container with some seawater, forcing the animals to escape the bottom layer and aggregate at the water rim, where they were collected with a pipette. Following collection, specimens were identified, and in some cases photographed for documentation of morph, before preserving them in 70-80% ethanol. DNA extraction, PCR amplification and sequencing were conducted with standard methods and procedures (see papers in this thesis for details), and DNA extraction was concentrated to somatic tissue to minimise the risk of sampling gut content, germ cells or paternally inherited mitochondrial lineages. Specimens and extracted DNA were stored at -20°C or -80°C for reference. Voucher specimens are deposited at the Natural History Museum, Gothenburg (Antarctic krill), and the National Natural History Museum, Stockholm (nemerteans, horse mussels).

When sampling Antarctic krill, we focused on the Atlantic sector of the Southern Ocean, and samples spanned over-wintering areas, spawning areas and oceanic current systems from the Antarctic Peninsula in the west, the Lazarev Sea (east Weddell Sea) in the southeast, and South Georgia islands in the north (at the rim of the maximum extent of the circum-polar drift ice). Furthermore, sampling was distributed as evenly as possible among the fishing sectors erected by the Commission of Conservation of Antarctic Marine Living Resources (CCAMLR 2009).

Horse mussels were collected throughout the species' known distributional range (except for Asia and Japan), from the White Sea and Iceland in the northeast to the Atlantic and Pacific coasts of North America in the west, and France in the south. By collecting from both the Atlantic and the Pacific, we aimed to deduce if populations from the two oceans were genetically divergent, apart from studying paleontological patterns.

The ribbon worm specimens were solely from Europe; *Oerstedia dorsalis* was sampled over a relatively large range (considering it is a brooding species) from the Azores in the south to Norway in the north, spanning the western Mediterranean Sea and the British Isles. The samples represent a wide variety of pre-defined, presumably intraspecific morphs (Sundberg 1984; Envall & Sundberg 1993). This sampling strategy made it possible to investigate whether geographic and/or morphological clusters represent genetically differentiated groups. The *Cerebratulus* spp samples represented a relatively

small-scale range (considering their relatively large body size, abundance and planktonic larval phase) of about 100 km off the northwest coast of Sweden. Based on observations that nemertean species often are misidentified due to a highly variable morphology (Strand & Sundberg 2005), we compared taxonomic morphospecies on a local scale to investigate whether this misconception, and subsequent erroneous species identification, might be an extensive problem in this ribbon worm group.

3.3 Genetic markers and analyses

Mitochondrial DNA (mtDNA) is a circular, organelle genome that classifies as a single locus containing roughly 37 genes in metazoans (averaging 16 kb in vertebrates), and has an elevated mutation rate (Hartl & Clark 1997; Nei & Kumar 2000). Since many of the mtDNA gene products are involved in the cell respiration cycle, it came as a great surprise when it was revealed that the mutation rate in these genes are 5-10 times higher than in the nucleus. It is most likely thought to depend on the lack of a proofreading system in the mitochondrion, which however is compensated by the fact that most mutations are synonymous, i.e. they do not change the protein codon product. The mtDNA is commonly maternally inherited, but some findings show that paternal leakage of mitochondrial DNA can occur at fertilisation, and in the common blue mussel, *Mytilus edulis*, purely paternal lineages of mtDNA has been discovered (Skibinski *et al.* 1994). Due to the high mutation rate, and therefore an increased saturation rate of substitutions, the mitochondrial genes are only suitable for divergence studies spanning relatively recent evolutionary times (Avise 2001). The mtDNA genes have subsequently been widely used in population genetics and intraspecific phylogeography as a complement to allele frequency studies (Sunnucks 2000).

The cytochrome oxidase subunit I gene (CO1 or cox1) is a mitochondrial gene widely used for resolution at the species or genus level, perhaps most famously as a hypothesised species identifier in the Barcode of Life Project summarised in Hebert *et al.* (2003). In this thesis, we use the CO1 gene to investigate intra- and interspecific genetic variation and geographic patterns of gene lineages. Nuclear DNA, on the contrary, typically exhibits all the properties absent in the mitochondrion, i.e. recombination, biparental inheritance and a considerably slower mutation rate (Hartl & Clark 1997). This has caused most nuclear genes to be used in higher level phylogenies, but relatively fast evolving loci such as the internal transcribed spacers (ITS) segments between the rRNA genes (5.8S, 18S and 28S) have increased in molecular systematics and population genetic studies, where it proves to be a valuable tool (Boyer *et al.* 2001). In this thesis, we likewise use and recognize the ability of the ITS regions (ITS1-5.8S-ITS), as well as the mtDNA 16S and CO1 genes, to identify both intra- and interspecific genetic clusters. When using the ITS, however, one must be cautious due to the different gene copies existing within the same individual. In **paper I**, we note this phenomenon in the ribbon worm *O. dorsalis*, and chose the option of excluding ambiguous sites

(heterozygous double peaks), since excluding specimens would not eliminate the problem - included specimens could still have un-sampled, multiple allele copies varying at the same sites.

In all our analyses, suitable evolutionary models for the included genes were determined under the Akaike Information Criterion as implemented in MrModeltest2 (Nylander 2004). When constructing phylogenies, we incorporated different evolutionary models for separate codon positions and genes, if more than one gene was part of the analysis. We used both maximum likelihood analysis in PAUP* 4.0 (Swofford 1998), as well as Bayesian inference with MrBayes v3.1.2 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Genetic statistics and parameters of genetic differentiation at the population level were estimated in DNAsp v.4.00.6 (Rozas & Rozas 1999), Arlequin v.3.1 (Excoffier *et al.* 2005) and LAMARC v2.1.2b (Kuhner 2006). The latter provides a genealogy-based estimation of the composite parameter theta (θ) from which the effective population size can be calculated. To infer the genealogy, we implement the coalescent (Kingman 1982a, b; Hudson 1990), i.e. the random merging of gene lineages backwards in time from descendants to parents. The inferred genealogy can be used to estimate the historical effective population size (Kuhner *et al.* 1995; Kuhner 2006), since the time for all lineages to coalesce and reach the most recent common ancestor logically depends on the number of lineages present.

4. RESULTS AND DISCUSSION

4.1 Applications of intraspecific genealogies

In **paper III and IV**, the aim was to use intraspecific genealogies to extract population parameters such as effective population size and genetic diversity, but also to investigate phylogeography and paleontology.

Paleontological perspectives

The majority of marine, paleontological studies have so far focused on, and drawn conclusions from intertidal species, as their habitat is inevitably affected by climate changes (e.g. glaciations or changes in sea level) (e.g. Wares & Cunningham 2001). Not surprisingly, it has lately become evident that subtidal organisms can also display genetic patterns related to climate changes, but as different versions of that of their intertidal counterparts (see e.g. Ball *et al.* 2006; Carr & Marshall 2008). In **paper III**, we corroborate that subtidal species provide additional information on the geographical range of climate changes, indicating that widely distributed, oceanic organisms have been affected as well.

Our haplotype network reveals a preliminary picture of dispersal routes and probable refugia from the last climate optimum affecting the horse mussel in the Atlantic Ocean (Figure 2; paper III). First, the mixed geographical origin of the deeply nested (i.e. ancient) haplotypes strongly suggests shared history and/or a rapid range expansion possibly with following high migration rates, although this needs to be confirmed by further studies on e.g. fast evolving

genetic markers. Second, high genetic diversity in specimens from the White Sea suggests a possible refugium there. Since the White Sea is temperate and located just south of the Arctic Circle, it further indicates a cold refugium from a climatic optimum, rather than a warm refugium from glaciations. This is consistent with the subtidal horse mussel being sensitive to high temperatures, but capable of coping with an extended period of below-zero temperatures (Howland *et al.* 1999).

The Icelandic population deviates in several aspects. It only had half of the haplotype diversity found at all other sites in the Atlantic, probably indicating either bottleneck and/or founder effects (Table 3; paper III). Deeply nested haplotypes in the network are dominated by Icelandic haplotypes, which means that they are relatively ancient, and thus rather points toward a bottleneck event in an old population. However, the relatively shallow sites where mussels were sampled for the study were likely affected more severely by climate changes, and probably host a young founder population with low genetic diversity. In summary, we see the utility of using a study species from another habitat than the intertidal in paleontological studies, with increased information gained on climate history – previous results on intertidal species concerning glaciations are here complemented with different genetic patterns providing information on climate optima ranges as well.

Genealogies are population chronicles

Paper IV extends the use of the intraspecific genealogy to infer genetic population parameters. We discuss whether the historical genetic patterns of the Antarctic krill show connections to the geographic distribution of sampling sites, i.e. if the inferred populations have been isolated in different regions. Further, we estimate the effective population size, and relate it to the population history and the biology of large populations. Current results from acoustic density surveys as well as allozyme and allele frequency studies suggest both possible panmixia and subdivision of krill populations (see e.g. Fevolden & Schneppenheim 1989; Zane *et al.* 1998; Siegel *et al.* 2004).

Genetic approaches can provide an important dimension to the present knowledge of the krill population by indicating population origin and possible long-term isolation between geographical sites. Widely distributed Antarctic krill samples from the Atlantic sector of the Southern Ocean showed no genetic clustering correlated to geographical sampling sites in our study (Figure 3, Table 2; **paper IV**). This suggests that the population share a common origin (possibly the Antarctic Peninsula region) and/or panmixia seems to prevail on a long-term basis. A high mobility of krill aggregations has also been indicted by acoustic density surveys (Siegel 2000). Although geographically distant from each other, the western Antarctic Peninsula and the south-eastern Lazarev Sea (East Weddell Sea) host a blend of deeply nested (ancient) haplotypes in the network, suggesting a south-eastern range expansion route. Our results further show that the female effective population size is surprisingly low as compared to the total population size – a mere fraction of the population seems to be contributing to the gene pool of the next generation ($Ne_{female} \approx 4x10^{-5}$ % of the

population census size). Both circumpolar sampling and more regional studies would likely provide a clearer picture of krill population origins and dynamics.

The crucial effects of demography

Demographic factors related to fluctuating population size, family size (lifetime production of offspring per individual), internal fertilization, and female size at fecundity has strong effects on the effective population size (Frankham 1995). An extreme variation in reproductive success (and thereby family size) is displayed in Antarctic krill, with density and biomass varying as much as 21-23 times between years (Siegel 2000; Siegel *et al.* 2002). This is probably mainly due to the internal fertilization of females (as opposed to free-spawning), high juvenile mortality, and older or larger size classes (> 45 mm) having about three to ten times higher egg production than younger and smaller size classes (Cuzin-Roudy 2000). Additionally, small size classes represent the largest clusters in the Southern Ocean (Siegel *et al.* 2004). Taken together, these features most likely affects the effective population size of the population and illustrates how large populations do not necessarily represent stable and resilient gene pools. Together with the discovery that census and effective population size are not positively correlated, N_e estimates are now regarded as some of the most important conservational concerns since the relative size and genetic composition of N_e depict future population viability (Frankham *et al.* 2002; Schwartz *et al.* 2008).

In concert with the panmictic behaviour of this species, conservation and fishery management of Antarctic krill need to focus on size selection and ecological factors, not on even catch distribution between fishing nations and geographical areas. In summary, we have gained insight into important factors of a high abundance exploited species using relatively cheap and easily accessible DNA data, as compared to e.g. acoustic surveys. It nicely illustrates how a broad preliminary picture about a species history and general genetic condition can be delivered through intraspecific genealogy studies, and used for subsequent hypothesis building and population management.

4.2 Variable genetic patterns

In **paper I and II**, we consider the existence of some taxonomic species, as well as discuss species identification and delimitation.

Cryptic species and genealogy patterns

Cryptic species are increasingly discovered and acknowledged in the marine environment (Knowlton 1993; Muths *et al.* 2006; Bickford *et al.* 2007; Harper & Hart 2007; Boissin *et al.* 2008; Chen & Hare 2008). The results in this thesis provide additional examples of cryptic species within well known organism groups. Our first study concerns the nemertean *O. dorsalis*, and we find that genetic relatedness generally do not correspond to either morphology or geographical sampling sites (Figure 2 & 3; **Paper I**). Further, some of the revealed clades show signs of shared ancestry and/or gene flow, especially

pronounced in the D-clade, where haplotypes from northwest Spain (El Ferrol) are found together with haplotypes from Great Britain (Wales). Individuals from both sites occur in deeply nested (ancient) as well as in tip (recent) haplotypes. Moreover, whenever a specimen from El Ferrol is present in a network, it is always given the highest outgroup probability in the statistical parsimony analysis, i.e. representing the most ancient haplotype presumably a "founder" (specimen EF36, EF38 and EF44 in network D, F and I, respectively). This suggests a northward dispersal of *O. dorsalis* from Spain, possibly originating from the southernmost sample site in the Azores. Whether this pattern reflects current migration levels remains to be studied. In contrast, we never see specimens from eastern Spain (Blanes) or southern Spain (Cadiz) in association with non-Spanish specimens. Additionally, they form separate (clade B, C) or divergent (clade I) groupings. Uneven sampling of geographic sites could cause this pattern, but it might also indicate limited gene flow between the Mediterranean Sea and the other sites. On the whole, we see a seemingly minor role of colour patterns as camoflauge in *O. dorsalis*, such that colour patterns do not group according to genetic clades or geographic occurrence. However, we cannot substantiate this observation without investigating e.g. genes directly coupled to morphology. In summary, our results are multifaceted, and obviously need further studies to clarify the possible correlation between some morphs, their genetic relatedness, speciation processes and what clades to call species in the cryptic *O. dorsalis* group.

In **paper II** then, the same issue of morphological species identification is subsequently taken to the taxonomic *inter*specific level in the nemertean *Cerebratulus* spp group, represented by specimens from the Swedish west coast. The nemertean identification literature present *Cerebratulus* spp as seemingly straightforward to identify based on external characters (see e.g. McIntosh 1873-74; Gibson 1994). However, considering that descriptions are relatively few, it was not surprising to find a haphazard blend of the species' haplotypes (Figure 2, paper II). Finally, **paper III** represents a common, well known and morphologically easy to identify species, the horse mussel bivalve (*M. modiolus*). It unexpectedly turned out to be two genetically distinguished groups. A common misconception about cryptic species, according to Bickford *et al.* (2007), is that they are recent speciation events, thus supposedly explaining why morphological traits have not developed yet. Paper III shows that cryptic species can be morphologically identical (at least superficially) for an extended period of time, even though the genetic divergence between the two is pronounced (22% in our case between the Pacific and Atlantic haplotype clades in Figure 2, paper III).

Species identification

Taking a closer look at our results, we find that identifying and announcing a new species is not straightforward. It seems clear that just a fixed percentage value of genetic divergence is not going to do the trick - the papers in this thesis demonstrate a genetic (mainly for the CO1 gene) divergence span of roughly 0.5-2% within clades and 8-29% between clades. Furthermore, the

genetic distances span both intra- and interspecific ranges found for marine (cryptic) taxa (Knowlton 1993; Dawson & Jacobs 2001; Govindarajan *et al.* 2005; Boissin *et al.* 2008; Luttikhuisen & Dekker 2010), which make it difficult to straightforwardly evaluate how many species to assign in our case. Some *Cerebratulus* spp haplotypes were the most divergent ones at 29% (Figure 2, Table 2; **paper II**) followed by the Pacific and Atlantic horse mussel clades as well as *O. dorsalis* specimens from the Azores and northeastern Spain at around 22% divergence (**paper III** and Table 3; **paper I**, respectively). There seems to be a "break" at around 2% genetic divergence, suggesting a possible species delimitation border for the investigated groups. The “mid-ranges” of around 8-11 % divergence in *O. dorsalis* illustrates the complications that might arise after posting a fixed species limit of genetic divergence. In this case, one can only hypothesise what this “intermediate” range might mean; perhaps an incipient species with a different evolutionary history or an inadequate sampling range for the group (Moritz & Cicero 2004, see also *Species delimitation in theory and practice* below). For further studies on taxa and their identification, this uncertainty is of course not a problem, just a challenge.

However, in the light of DNA taxonomy and identification of species using barcodes, we have an issue (DeSalle *et al.* 2005). Our findings have implications for species identification in general, and the barcoding approach in particular for two reasons. First, some of the species descriptions studied in this thesis are solely based on morphological and/or histological characters and could therefore be misleading. Second, major additional (intraspecific) studies seems to be needed to extend the taxonomic base upon which the barcoding species identification naturally must rest. DNA barcoding can be compared to a library card identifying a book by a certain code. The card is useless if it leads us to a book we did not anticipate – the code is wrong. Biologically speaking then, books are taxonomically defined species and the library card is a genetic sequence. Needless to say, without taxonomy (*known* books) barcoding is shooting blanks (leaving aside the fact that species are not invariable entities, like books). In spite of this dilemma, there is still a sprouting progression linking the two – barcoding provides DNA taxonomy with a continuous influx of DNA data, which can extend current (DNA) taxonomy by e.g. revealing cryptic species among organisms with few morphological characters (Schander & Willassen 2005). To decide at what level of genetic divergence to draw the taxon line, one has to compare the genetic variation within a population to the variation between populations, or taxa (Blaxter 2004), and the organism should ultimately be a "known, accepted taxon".

Species delimitation in theory and practice

If we consider a couple of the numerous species concepts (summarized and discussed in Coyne & Orr 2004), we could conclude that any of our distinct genetic clades in this thesis correspond to a species, following the evolutionary species concept and/or the phylogenetic species concepts. In summary, the proponents of these concepts define a species by an exclusive group of genetic lineages ("monophyletic" in phylogenetics) that share a

common ancestor and whose gene lineages therefore coalesce more recently than other lineages outside the group. However, strictly speaking both concepts encounter difficulties. First, organisms can display different evolutionary histories at different genes, as discussed in INTRODUCTION AND BACKGROUND (Figure 1). Second, the concepts do not specify how many genes that must differ (and by how much) for two lineages to be considered different species – if *all* gene lineages of the genome need to be monophyletic, we have a practical and time consuming issue at hand. Considering that most genealogical studies so far have only dealt with a fraction of the genome, the implications remain to be discovered. In **paper I**, we find evolutionary congruence among nuclear and mitochondrial loci, although it might simply be due to the low resolution of the ITS and 16S genes (results not shown, but see Figure 3; paper I). Third, genealogical speciation can "reverse", as compared to reproductive incompatibility, which is considered a permanent species defining condition in the biological species concept. For example, if individuals from different genealogical lineages (e.g. a number of small, isolated, monophyletic populations) become sympatric, they could still be able to hybridize if reproductive isolation had not yet evolved. Gene lineages would from thereon share their evolutionary history, and then perhaps be considered a single species. This can easily be imagined for the closely distributed populations of nemerteans in **paper I and II**. Although the sampled specimens seem to be part of genetically distinct groups, many individuals can instantaneously be deposited into another population through a rafting event (Gutow *et al.* 2005). In contrast, reproductive isolation would retain the species status even under such sympatric conditions. Fourth, the amount of genetic sequence divergence (or variability in a morphological trait) needed for a group to be considered a species is subjective, as we have seen for coloration patterns in the nemertean studies of this thesis. Still, considering the difficulties in keeping nemerteans in the lab (M. Strand, pers. comm.) to perform e.g. cross-fertilization tests, genetic studies can at least provide a preliminary insight on the relationships in the nemertean groups, and probably in many others as well.

To add to the problem of different concepts, species *delimitation* is confused with the issue of species *conceptualization*, which leads to disagreements about the species category and methods for inferring the number of species and their boundaries (de Queiroz 2007). For example, if reproductive isolation is found between specimens *within* a clade, it may put the biological and the phylogenetic species concepts at odds (Knowlton & Weigt 1997). A flexible solution is presented by the same author, where a separately evolving metapopulation lineage can be seen as the only *property* needed for a species to exist, i.e. *the theoretical concept* (see more in INTRODUCTION AND BACKGROUND). The existing species concept criteria instead represent *lines of evidence (operational criteria)* relevant for assessing lineage separation to some degree. More lines of evidence are consequently associated with a higher degree of corroboration between the hypothesis and speciation event. Species concepts of today will thus not be used to conceptualize species, only to delimit them. This approach thus successfully separates methodology from conceptual issues, giving them room to function as complementary lines of evidence rather

than conflicting theories. Logically, the *absence* of an operational criterion is not evidence against a hypothesis for lineage separation (no more than absence of any kind of data would prove that it does not exist). The investigated feature (e.g. reproductive isolation) might not have evolved yet! The farther lineages have reached in the process of divergence, the larger is the number of evidence traits expected to be found.

Most species concepts only succeed in defining a coherent, intrinsic individual as a species if external factors are considered in parallel (Horvath 1997). Reproductive isolation, postzygotic isolation and natural selection are all examples of factors that cannot define the individuality of species intrinsically, since their definitions require interaction with another species or ecological factors. Population structure, mating system and mechanisms that can limit gene flow are, on the contrary, considered to be factors that can strengthen the cohesiveness of the individuality of a species. In this thesis, we have focused on a genealogical approach to investigate intra- and interspecific genetic patterns. By doing so, we wished to circumvent the pre-existing traditional taxonomic definitions and descriptions of species, and investigate whether natural groups exist on the level of genetic relatedness. **Papers I, II and III** shows indications of several intrinsic lines of evidence for new species as genetic (population) structure is discovered as distinct clades. Yet, these clades rarely correspond to the geographical sampling site of the included specimens, the taxonomic species, or morphological traits. Rather, they contain a blend of these. In this manner, the great majority of our inferred clades are only distinguished by genetic characters. It is easy to see how a small subset (e.g. a single individual in traditional systematics) of the specimens in our gene genealogies could be misleading when erecting species trees based solely on morphology. In contrast to our results then, and reconnecting to the barcoding approach, Pons *et al.* (2006) found that mitochondrial DNA delimitation of tiger beetle species (*Rivacindela* sp) corresponded to earlier traditional definitions (e.g. coherence of geographical range and congruence with morphologically identified species). The authors propose a likelihood method that "determines the point of transition from species level (speciation and extinction) to population-level (coalescence) evolutionary processes", by detecting an increased branching rate in ultrametric trees. This approach has been judged as too simple, ignoring the fact that sampling scheme and migration will heavily affect branching events within demes (Lohse 2009). Further, the "barcoding gaps" widely used to delimit intraspecific diversity from interspecific diversity may simply reflect this (inadequate) spatial sampling scheme (Moritz & Cicero 2004). Furthermore, inferring species trees from gene trees, which presently are based on concatenated DNA sequence data, has been criticized (Edwards 2009). The latter authors argue that focus on population genetic processes, such as coalescence, and inferring trees from a range of heterogeneous gene trees (instead of a range of variable DNA sequences) will be the future of molecular phylogenetics at the species level. In addition, some attempts have been made to develop statistical tests for taxonomic distinctness (e.g. O´Meara 2010), and Rosenberg (2007) presents quite promising results when it comes to rejecting random branching events – a sample of ten lineages from the studied group, and three "outgroup"

lineages, is sufficient to confirm non-random monophyly at high probability levels. This provides some hope to the gene lineage oriented species concepts, and some confirmation of the results in this thesis indicating new species within existing taxonomic groups. We can only conclude that there is a wide range of outcomes in these kinds of studies, and the only thing that seems obvious is the fact that nothing is obvious. This stresses the need of case-to-case evaluations of species identification, definition, and of integrating species concepts.

5. CONCLUSIONS AND PROSPECTS

This thesis explores the genetic divergence and patterns of organisms without forcing results onto *a priori* species descriptions. The empirical data and conclusions from the included studies should hopefully stimulate and develop the discussion about what a species is, how we can delimit it and work with it in different contexts. Methods based on statistical parsimony is useful when screening a relatively large number of samples from previously un-investigated taxa, and in combination with species distribution data it can reveal additional information about a species' dispersal routes and lineage sorting. Sampling should ultimately span the organelle and nuclear genomes, as well as investigating the effects of multiple specimen sampling on lower-level phylogenies. Perhaps cytological methods (e.g. organelle compatibility) can add up to the quest, providing a practical *in vivo* composite aspect, as coadapted complexes between mtDNA and the nuclear genome can develop in reproductively isolated populations) (Willett & Burton 2001).

It obviously seems like we need taxonomy and to know what a *taxon* is; it is after all hard to refer to "genetically exclusive groups" at field trips. However, what can traditional taxonomy do in the case of cryptic or morphologically very similar organisms? Spontaneously I say it stands beaten in many cases. On the other hand, taxonomic methodology might just need to become more developed and refined, although it might require more sophisticated protocols, such as the mesoglea species test for hydroids (Govindarajan *et al.* 2005). However, in scientific taxonomic revisions and development of species identification methods, taxonomic methodology should perhaps be the prime focus of future taxonomic and barcoding cooperation.

In practical and conservational terms though, it all boils down to the simple fact that we want to conserve all kinds of biodiversity; genetic, ecological, morphological or functional, perhaps depending on the circumstances, and not necessarily confine all living organisms to the same theoretical framework at any price. It seems to be a case-to-case science. What really make the matter so complicated are scientists that seek a universally applicable definition of species, which are in constant change. Since the degree of distinctness in factors such as geographical subdivision and genetic divergence varies substantially among organisms, it is not surprising that each species concept gives an incomplete view of the multitude of existing speciation events. I conclude that studies on the gene lineage level is an excellent place to

start when addressing any kind of species or population aspect, as it is after all the most "unmasked" representation of the evolutionary process, not influenced by the subjectivity of human beings. The identification and delimitation of species lineages is still imposing a problem, as genetic divergence and patterns will differ between groups, but gene lineage relationships will *always* provide us with some objective information, no matter the questions asked.

6. REFERENCES

Avise JC (2001) *Phylogeography The history and formation of species*. Harvard University Press, London.

Ball AO, Beal MG, Chapman RW, Sedberry GR (2006) Population structure of red porgy, *Pagrus pagrus* in the Atlantic Ocean. *Marine Biology* 150: 1321-1332.
Bertrand Y, Pleijel F, Rouse GW (2006) Taxonomic surrogacy in biodiversity assessments, and the meaning of Linnean ranks. *Systematics and Biodiversity* 4: 149-159.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* 22: 149-155.

Blaxter ML (2004) The promise of DNA taxonomy. *Proceedings of the Royal Society B: Biological Sciences* 359: 669-679.

Boissin E, Féral JP, Chenuil A (2008) Defining reproductively isolated units in a cryptic and synoptic species complex using mitochondrial and nuclear markers: the brooding brittle star, *Amphipolis squamata* (Ophiuroidea). *Molecular Ecology* 17: 1732-1744.

Boyer SL, Flechtner VR, Johansen JR (2001) Is the 16S-23S rRNA internal transcribed spacer region a good tool for use in molecular systematics and population genetics? A case study in Cyanobacteria. *Society for Molecular Biology and Evolution* 18: 1057-1069.

Brown RA & Seed R (1977) *Modiolus modiolus* (L.). An autoecological study. In: *Biology of Benthic Organisms* (eds) Keegan BK, Cleidigh PO, Boddin PJS. Pergamon press, Oxford.

Bürger O (1892) Zur Systematik der Nemertinefauna des Golfs von Neapel. *Nachrichten von der Königlichen Gesellshaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen* 5: 137-178.

Carr SM, Marshall HD (2008). Intraspecific phylogeographic genomics from multiple complete mtDNA genomes in Atlantic Cod (*Gadus morhua*): Origins of the "Codmother," trans-Atlantic vicariance and midglacial population expansion. *Genetics* 180: 381-389.

CCAMLR (2009) Map of the CCAMLR Convention Area. Boundaries of the statistical reporting areas in the Southern Ocean. Retrieved on March 9, 2009 at http://www.ccamlr.org/pu/E/conv/map.htm

Chen G, Hare MP (2008) Cryptic ecological diversification of a planktonic estuarine copepod, *Acartia tonsa*. *Molecular Ecology* 17: 1451-1468.

Clement M, Posada D and Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657-1660.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer, Sunderland, MA.

Cuzin-Roudy J (2000) Seasonal reproduction, multiple spawning, and fecundity in northern krill, Meganyctiphanes norvegica, and Antarctic krill, *Euphausia superba*. *Canadian Journal of Fisheries and Aquatic Sciences* 57: 6-15.

Davenport J, Kjørsvik E (1982) Observations on a Norwegian intertidal population of the horse mussel *Modiolus modiolus* (L.). *Journal of Molluscan Studies* 48: 370-371.

Dawson MN, Jacobs DK (2001) Molecular evidence for cryptic species of Aurelia aurita (Cnidaria, Scyphozoa). *Biological Bulletin* 200: 92-96.

DeSalle R, Egan MG, Siddall M (2005) The unholy trinity, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society of London* 360: 1905-1916.

de Queiroz K (2005) Ernst Mayr and the modern concept of species. *Proceedings of the Natural Academy of Sciences of the United States of America* 102: 6600-6607.

de Quieroz K (2007) Species concepts and species delimitation. *Systematic Biology* 56: 879-886.

Edwards SV (2009) Is a new and general theory of molecular systematics emerging? *Evolution* 63: 1-19.

Envall M, Sundberg P (1993) Intraspecific variation in nemerteans (Nemertea): synonymization of the genera *Paroerstedia* and *Oerstediella* with *Oerstedia. Journal of Zoology, London* 230: 293-318.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetic analysis. *Evolutionary Bioinformatics Online* 1: 47-50.

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368-376.

Fevolden SE, Schneppenheim R (1989) Genetic homogeneity of krill (Euphausia superba Dana) in the Southern Ocean. *Polar Biology* 9: 533-539.

Flagel LE, Rapp RA, Grover CE, Widrlechner MP, Hawkins J, Grafenberg JL, Alvarés I, Chung GY, Wendel JF (2008) Phylogenetic, morphologcial, and chemotaxonomic incongruence in the North American endemic genus *Echinacea*. American Journal of Botany 95: 756-765.

Frankham R (1995) Effective population size/adult population size ratios in wildlife: A review. *Genetical Research* 66: 95-107.

Frankham R, Ballou JD, Briscoe DA (2002) *Introdution to conservation genetics*. Cambridge University Press, Cambridge.

Futuyma DJ (1998) *Evolutionary biology*. Sinauer Associates, Sunderland, MA.

Gibson R (1994) Nemerteans. In: Barnes RSK, Crothers JH, eds. Synopses of the British Fauna (New Series), Vol. 24, 2nd edn. Shrewsbury: Field Studies Council (FSC).

Gibson R (1998) Epilogue – one hundred years of nemertean research: Bürger (1985) to the present. Hydrobiologia 368: 301-310.

Govindarajan AF, Halanych KM, Cunningham CW (2005) Mitochondrial evolution and phylogeography in the hydrozoan Obelia geniculata (Cnidaria). Marine Biology 146: 213-222.

Gutow L, Thiel M, Franke H-D (2005) Epi-pelagic invasions - the drift vector, MarBEF and LargeNet Workshop on Long Term Data Set Networking, 3-4 March 2005, Helgoland, Germany.

http://hdl.handle.net/10013/epic.22717

Hart MW, Sunday J (2007) Things fall apart: biological species form unconnected parsimony networks. *Biology Letters* 3: 509-512.

Hartl DL, Clark AG (1997) *Principles of population genetics* (3rd ed) Sinauer Associates, Inc., Sunderland, Massachusetts.

Harper FM, Hart MW (2007) Morphological and phylogenetic evidence for hybridization and introgression in a sea star secondary contact zone. *Invertebrate Biology* 126: 373-384.

Hebert PDN, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences* (Suppl.) 270: 96-99.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2004) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* 270: 313–322.

Hennig W (1965) Phylogenetic Systematics. *Annual Review of Entomology* 10: 97-116.

Hewitt G (2000) The genetic legacy of the Quarternary ice ages. *Nature* 405: 907-913.

Hewitt PR, Linen Low HE (2000) The fishery on Antarctic krill: Defining an ecosystem approach to management. *Reviews in Fisheries Science* 8: 235-298.

Hewitt RP, Watkins J, Naganobu M, Viacheslav S, Brierley AS, Demer D, Kasatkina S, takao Y, Goss C, Malyshko A, Brandon M, Kawaguchi S, Siegel V, Trathan P, Emery J, Everson I, Miller D (2004) Biomass of Antarctic krill in the Scotia Sea in January/February 2000 and its use in revising an estimate of precautionary yield. *Deep Sea Research II: Topical Studies in Oceanography* 51: 1215-1236.

Horvath CD (1997) Some questions about identifying individulas: failed intuitions about organisms and species. *Philosophy of Science* 64: 654-668.

Howland RJM, Pantiulin AN, Millward GE, Prego R. 1999. The hydrography of the Chupa estuary, White Sea, Russia. *Estuarine, coastal and shelf science* 48: 1-12.

Hudson RR (1990) Gene genealogies and the coalescent process. *Oxford Survey of Evolutionary Biology* 7: 1-44.

Huelsenbeck JP, Ronquist F (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17: 754-755.

Hundsdoerfer AK, Wink M (2006) Incongruence of morphology and genetic markers in *Hyles tithymali* (Lepidoptera: Sphingidae) from the Canary Islands. *Journal of Zoological Systematics and Evolutionary Research* 44: 316-322.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254-267.

Kingman JFC (1982a) The coalescent. *Stochastic Processes and Their Applications* 13: 235-248.

Kingman JFC (1982b) On the genealogy of large populations. *Journal of Applied Probability* 19A: 27-43.

Knowles LL, Y-H Chan (2008) Resolving species phylogenies of recent evolutionary radiations. *Annals of the Missouri Botanical Garden* 95:224-231.

Knowlton N (1993) Sibling species in the sea. *Annual review of Ecology and Systematics* 24: 198-216.

Knowlton N, Weigt LA (1997) Species of marine invertebrates: a comparison of the biological and phylogenetic species concepts. *In:* Species: the units of biodiversity (eds Claridge MF, Dawah HA, Wilson MR), Chapman & Hall, London, UK, pp. 199-219.

Kuhner MK (2006). LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics* 22: 768-770.

Kuhner M, Yamato J, Felsenstein J (1995) Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. *Genetics* 140: 4121-1430.

Lee T, Hong CH, Kim JJ, Foighil D (2007) Phylogenetic and taxonomic incongruence involving nuclear and mitochondrial markers in Korean populations of the freshwater snail genus *Semisulcospira* (Cerithioidea: Pleuroceridae). *Molecular Phylogenetics and Evolution* 43: 386-397.

Leonard DRP, Clarke KR, Somerfield PJ, Warwick RM (2006) The application of an indicator based on taxonomic distinctness for UK marine biodiversity assessments. *Journal of Environmental Management* 78: 52-62.

Lewis PO (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* 50: 913-925.

Lohse K (2009) Can mtDNA barcodes be used to delimit species? A response to Pons et al.. (2006). *Systematic Biology* 58: 439-442.

Lou M, Golding GB. Assigning sequences to species in the absence of large interspecific differences. *Molecular Phylogenetics and Evolution* (2010), doi: 10.1016/j.ympev.2010.01.002.

Luttikhuisen PC, Dekker R (2010) Pseudo-cryptic species *Arenicola defodiens* and *Arenicola marina* (Polychaeta: Arenicolidae) in Wadden Sea, North Sea and Skagerrak: Morphological and molecular variation. *Journal of Sea Research* 63: 17-23.

Mace GM (2004) The role of taxonomy in species conservation. *Proceedings of the Royal Society B: Biological Sciences* 359: 711-719.
Maurer D (2000) The dark side of taxonomic sufficiency (TS). *Marine Pollution Bulletin* 40: 98-101.

McIntosh WC (1873-74) A monograph of the British annelids. Part 1. The nemerteans. London: Ray Society, 1-213.

Moritz C, Cicero C (2004) DNA barcoding: promise and pitfalls. *Public Library of Sceince Biology* 2:1529–1531.

Muths D, Davoult D, Gentil F, Jollivet D (2006) Incomplete cryptic speciation between intertidal and subtidal morphs of *Acroconida brachiata* (Echinodermata: Ophiuroidea) in the Northeast Atlantic. *Molecular Ecology* 15: 3303-3318.

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, Inc., New York.

Nylander JAA (2004) MrModeltestv2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey J (2004) Bayesian phylogenetic analysis of combined data. *Systematic Biology* 53: 47-67.

O´Meara BC (2010) New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology* 59: 59-73.

Plattner A, Kim JJ, Reid J, Lim YW, Yamaoka Y, Breuil C (2009) Resolving taxonomic and phylogenetic incongruence within species *Ceratocystiopsis minuta*. *Mycologia* 101: 878-887.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution* 16: 37-45.

Pons J, BarracloughTG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* 55: 595-609.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.

Rosenberg NA (2007) Statistical tests for taxonomic distinctiveness from observations of monophyly. *Evolution* 61: 317-323.

Rozas J, Rozas R (1999) DNAsp version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174-175.

Schwartz MK, Luikart G, Waples RS (2006) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution* 22: 25-33.

Schander C, Willassen E (2005) What can biological barcoding do for Marine Biology? *Marine Biology Research* 1: 79-83.

Siegel V (2000) Krill (Euphausiacea) demography and variability in abundance and distribution. *Canadian Journal of Fisheries and Aquatic Sciences* 57: 151-167.

Siegel V, Bergström B, Mühlenhardt-Siegel U, Thomasson M (2002) Demography of krill in the Elephant Island area during summer 2001 and its significance for stock recruitment. *Antarctic Science* 14: 162-170.

Siegel V, Kawaguchi S, Ward P, Litvinov F, Sushin V, Loeb V, Watkins J (2004) Krill demography and large-scale distribution in the southwest Atlantic during January/February 2000. *Deep-Sea Research II* 51: 1253-1273.

Skibinski DOF, Gallagher C, Beynon CM (1994) Sex-limited mitochondrial DNA transmission in the marine mussel *Mytilus edulis*. *Genetics* 138: 801-809.

Strand M, Sundberg P (2005) Delimiting species in the hoplonemertean genus Tetrastemma (phylum Nemertea): morphology is not concordant with phylogeny as evidenced from mtDNA sequences. *Biological Journal of Linnean Society* 86: 201-212.

Sundberg P (1979) Statistical analysis of variation in characters in *Tetrastemma laminariae* (Nemertini), with a redescription of the species. *Journal of Zoology, London* 189: 39-56.

Sundberg P (1984) Multivariate analysis of polymorphism in the hoplonemertean *Oerstedia dorsalis* (Abilgaard, 1806). *Journal of Experimental Marine Biology and Ecology* 78: 1-22.

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology and Evolution* 15: 199-203.

Swofford DL (1998) PAUP* 4.0 - Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, MA.

Thollesson M, Norenburg JL (2003) Ribbon worm relationships: a phylogeny of the phylum Nemertea. *Proceedings of the Royal Society B: Biological Sciences* 270: 407-415.

Tyler-Walters H (2007) *Modiolus modiolus*. Horse mussel. Marine Life Information Network: Biology and Sensitivity Key Information Sub-programme [on-line]. Plymouth: Marine Biological Association of the United Kingdom. [cited 25/03/2008]. Available from: <http://www.marlin.ac.uk/species/Modiolusmodiolus.htm>

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* 15: 1419-1439.

Wares JP, Cunningham CW (2001) Phylogeography and historical ecology of the north Atlantic intertidal. *Evolution* 55: 2455-2469.

Watkins JL, Hewitt R, Naganobu M, Sushin V (2004) The CCAMLR 2000 survey: A multination, multi-ship biological oceanography survey of the Atlantic sector of the Southern Ocean. *Deep-Sea Research II* 51: 1205-1213.

Wiemers M, Fiedler K (2007) Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae). *Proceedings of the Royal Society B: Biological Sciences* 274: 2881-2889.

Willet CS, Burton RS (2001) Viability of cytochrome C genotypes depends on cytoplasmic backgrounds in *Tigriopus californicus*. *Evolution* 55: 1592-1599.

Zane L, Ostellari L, Maccatrozzo L, Bargelloni B, Battaglia B, Patarnello T (1998) Molecular evidence for genetic subdivision of Antarctic krill (*Euphausia superba* Dana) populations. *Proceedings of the Royal Society Series B: Biological Sciences* 265: 2387-2391.

ACKNOWLEDGEMENTS

There were many rivers to cross, but here I am, soaking wet and freezing, but I'm here! Or as Leonard Cohen put it: "…by the rivers dark I wandered on". I thank you for unknowingly being the greatest source of inspiration in my life, side by side with the ocean. You have dragged me through the darkest times, made me laugh in the middle of nowhere, receiving enlightment when I least expected it, just like the ocean and Nature. I celebrate your wisdom and creativity by citing parts of your work before each paper of this thesis. I also thank all the other great sources of powerful voices and music, which has carried me to work through any horizontalness of Gothenburg rain, spinning on my beloved bike SODA - thank you for the nine years! Without all of you, I would definitely have been a victim of stress hormone-related death.

First in line of academics to be thanked is of course my supervisor Per Sundberg, who has fought his way through many (thousands?) of text drafts. I hope you will be able to read again! I also thank my assistant supervisor Thomas Dahlgren for always being a positive chap, and for providing clear guidelines for my work. Last, but not least, of seniors I thank my examiner Christer Erséus for being a hard-core rock in hard-core times. I say no more, because silence is golden. Another of greatest colleagues is Urban Olsson, who I was always happy to see and soak positive vibes from. Thank you for that, and also for memorable questions on early mornings at work, such as: "Do you know what a hairy sea squirt looks like?" And Daniel-san, what can I possibly say?! THE Room mate of all Room mates. Sharing an office with you is never boring, and comes close to the feeling of being in a time machine (set to the Medieval), a kaleidoscope, a Russian babushka doll, outer space and inner circle at the same time. Thank you for always believing in me. Finally, my colleagues and fellow students throughout my studies and PhD, you have been invaluable. Thanks (in no taxonomical order) Ann-Sofie and Lena for ALWAYS smiling at me, zoophysiologists for great lunch talks, Bernth for making me happy, Emelie for making me feel young, Anna for looong talks, Lisa for loud laughs, Pierre for walks on the wild side, Helena for bare feet, Judith for having balls (she is a woman), Jenny for coming back, Haixia for the fish, Arne for dry comments and eye-blowingly strong coffee, Niklas for "Terror på traversen", Johan Eklöf for creativity, Susanne Viker for enthusiasm, Monica for dissecting gutted squid with students, Anna Zeffer (and Bitte) for openness, Mare for great spirit, Petri for psychological support, Kalle for horses to ride on, Benno for great attitude, all the other GREAT people at Tjärnö, Bo Bergström for believing I could do this, Serik for mathematical support (is there?!), Julia for being such a piece of chocolate, Kristina for being such a piece of vanilla, Malin for apple cheeks – I love them, Linda for being strong and Mats for being so bloody tall. My long-time friends from the age of three and seven, Christina and Anna, we are still in touch and hope we always will be! I'm happy to have you. Some of my latest friends are Johanna – you are a ray of sun, I feel as welcome in your home with your family as I do in my own. Thank you for the support!! And Mirushka, thank you for being so strong and having philosophical discussions with me. The AMP people, I thank you especially for making me feel so competent, valuable and

alive! The same goes for all the people at the East Indiaman ship Götheborg for a mind-blowing experience.

A big thank you to colleagues at EBM –especially Haamid, Mia, Susanne, Gabriella, Emilia and Petra, for making me look forward to work in the morning and welcoming me as an old friend to my new work place. And lastly, the people who has made my everyday life function and the thesis to be written – Lena, Anette, Lotta, Ulrika and Sofia at Danska vägen children daycare center. Without your positive attitude and flexibility I would have been in even greater stress! Thank you for understanding!

Lastly, my realtives. Scattered all over the country and the world, I thank you for the support you have been giving, unconsciously or openly. My Mom, always there, always happy about a magpie or a small flower – I learnt to love the little things by watching you, and they make up most of Life. Thank you for showing me that! You could have looked a little bit less worried about me finishing my PhD, though. My Dad, my thunderstorm. You thought me to act today, and never wait for tomorrow, and that I can do anything I set my mind to. With your crazy energy and will of steel you have showed me what it it possible to accomplish, and plowed the way for my own stubborness to grow strong. Without that, this thesis would not be here. Kösszi szépen, nagyon szeretlek Papuska! And then there is my sister, who thought me not to take things for granted, and also saved me by babysitting a million times - tack Gambi! I thank my grandparents for a welcoming home and the ever-filled trays with homebaked cookies. I thank my cousins and their families for showing their interest in my work. My Hungarian grandparents rest in peace, but I will never forget great words of wisdom, my Grandmothers food and my Grandfathers harmonica under the grape vine.

The latest (and the first, as we systematists know) branch of my family tree is in Kenya, holding some of the purest, strongest and greatest people on Earth. Thank you all for accepting the Mzungu into your family! I'm so grateful. Ephraim, thank you for welcoming me into your home as your own daughter - I love my house! Brigit, thank you for being my sister with shining loving eyes. John, thank you for sleeping on the floor for weeks so that I could have the bed. Lee, thank you for many loud talks and laughs. Shadrack, thank you for always trying to teach me Swahili. Shawn, Michael and Eugene, thank you for treating me like a real auntie, hugging me every time. Izo, Robert, Newton, Alo, Judah, Ali, Daudi, Frankie, Soldier, Baba Gong - thank you for letting me be "one of the guys". Shee, Mariam, Faraj, Mohamad, William, Jackson - thank you for showing me some of the most beautiful places in the World - Lamu Island and Maasai Mara (and thanks for the goat William!). In fact, thank you all the people of Riruta Satellite for giving me a second home always scented with grilled corn cobs and sounding of laughter and music. Thank you Atur and Mama for loving me so much.

And then there is Morris, my Husband, Mando, my little lightning of a Son, and NoName, our baby still in unborn orbit -thank you for kickin' my ass from the *inside* to finish this thesis! Where did you all come from? Suddenly I had you all. I thank you Morris for being my greates master in teaching me paitience and silence. I love you for being as crazy as I am, wanting to marry someone after a

weeks aquaintance - we did the right thing and many wonderful things has come out of it, we will make it together. I admire all the efforts you have gone through since you moved to Sweden. You got a job, learnt Swedish and recieved a Swedish drivers license in no time, and you are still working hard. Thank you for trying to make it for us! Nakupenda sana! Mando - thank you for keeping my mind open and to think outside the box (of course the kitchen-pot lid should go on the potty, and the birds should also ride with the bus). I hope you will forgive me for the times when I've been tired and less patient. I love you endlessly, and I would sacrifice to finish this thesis for you if I had to - the most important thing is that we are happy together. *Däää va den*!