LIFESTYLE TRANSITIONS AND RATES OF ANIMAL MOLECULAR EVOLUTION

A Thesis

Presented to

The Faculty of Graduate Studies

of

The University of Guelph

by

JEREMY RYAN DEWAARD

In partial fulfilment of requirements

for the degree of

Master of Science

February 2004

Your file Votre référence
ISBN: 0-612-92945-0
Our file Notre référence
ISBN: 0-612-92945-0

ABSTRACT

LIFESTYLE TRANSITIONS AND RATES OF ANIMAL MOLECULAR EVOLUTION

Jeremy Ryan deWaard
University of Guelph, 2004

Advisor:
Professor P. D. N. Hebert

This thesis investigates how lifestyle transitions influence the rates of molecular evolution in animals. This is accomplished by employing a phylogenetic comparative method to contrast relative rates of molecular evolution with lifestyle transitions in a diverse range of taxa. The results suggest that adopting a planktonic lifestyle is associated with an enhanced rate of substitution, and that this rate acceleration is likely caused by their increased exposure to ultraviolet radiation. This study also reveals that parasitic lineages have an increased rate of sequence evolution relative to their free-living allies, likely a consequence of a genetic arms race between parasite and host. These results have broad implications for the application of molecular clocks, as well as for understanding both diversification patterns, and the modulation of animal molecular evolution.

Acknowledgements

This research was supported by an Ontario Graduate Scholarship to myself and a Natural Sciences and Engineering and Research Council of Canada grant and Canada Research Chairs appointment to Paul Hebert.

Chapter 1 was written in collaboration with Veronika Sacherova, Melania Critescu, Elpidio Remigio, Theresa Crease, and Paul Hebert. It has been modified and submitted to *Molecular Phylogenetics and Evolution* for publication. It was a great pleasure and fantastic learning experience to work with more 'senior' scientists and I am thankful for the great opportunity.

I would like to acknowledge and thank several people who made the completion of this thesis possible. Firstly, I would like to thank those that I worked with in the field: Dave Hardie, Jennie Knopp, Monica Pasquali, Pia Salathe, Kareen Schnabel, Ainslie Little, and Adam Phipps. Whipping around in helicopters and zodiacs, diving in Watson's Bay to Creswell Bay, reckless driving in northern Ontario and northern Queensland – all would have been miserable without you... well, less enjoyable anyways...

Secondly, I am indebted to those individuals who provided specimens and/or sequence data: Melania Critescu, Rowan Barrett, Ryan Gregory, Suzanne Gray, Russ Hopcroft, Simon Jarman, Stephanie Kirk, Ed Remigio, Veronika Sacherova, Kareen Schnabel, Glen Wagner, and Jonathan Witt. In some cases, this involved smuggling

across foreign borders and illegally shipping hazardous goods – thanks for taking the risk.

In addition, I wish to thank my advisory committee (Teri Crease, Jinzhong Fu, and Paul Hebert) for their guidance over the past two years. In addition, I would like to acknowledge my defence committee (Roy Danzmann, Jinzhong Fu, Paul Hebert, and Pat Wright) for providing useful comments and revisions for the finished thesis.

As well, I have been extremely fortunate to work in a large and ever-changing, but always enjoyable, lab setting. I'd like to thank each and every past and present lab-mate: Sarah Adamowicz, Shelley Ball, Rowan Barrett, Tanya Brown, Heather Cole, Andrea Cox, Melania Cristescu, Alina Cywinska, Rob Dooh, Ryan Gregory, Dave Hardie, Sujeevan Ratnasingham, Ed Remigio, James Rhydderch, Veronika Sacherova, Lisa Schieman, Kareen Schnabel, Jonathan Witt, and Tyler Zemlak.

And last but not least, I wish to thank my family – Tollina, Frank, Carolyn and Andrew deWaard – and especially Stephanie Kirk, for all being extremely supportive during the past couple of years. There were times, particularly when I was writing up, when I was undoubtedly 'crabby' and 'no fun', and you all managed not only to tolerate me, but to help out and encourage me to finish – I am tremendously thankful for that.

TABLE OF CONTENTS

LIST OF TABLES.....vi

LIST OF FIGURES.....viii

LIST OF APPENDICES.....ix

GENERAL INTRODUCTION: Molecular rate variation.....1

The molecular clock and rate variation.....2

Partitions of rate variation.....3

Determinants of rate variation.....4

Comparative method in molecular evolution.....7

The present study.....8

CHAPTER 1: Probing the origins of the branchiopod crustaceans.....10

Abstract.....11

Introduction.....12

Materials and methods.....14

Taxonomic sample & gene selection.....14

Molecular techniques and sequence alignments.....16

Phylogenetic analysis.....17

Results.....18

More taxa approach.....18

More genes approach.....19

Comparison of approaches.....20

Discussion.....21

Branchiopod interordinal affinities.....21

Relationships within the Cladocera.....22

Strategies of taxon and gene sampling.....23

Conclusions.....24

CHAPTER 2: UV exposure, the planktonic lifestyle and accelerated rates of molecular evolution.....35

Abstract.....36

Introduction.....37

Materials and methods.....39

Taxonomic sample and independent comparisons.....39

Sequence data.....40

Relative rates analysis.....41

Nucleotide and dinucleotide frequency analysis.....42

Tests of association.....42

Results.....43

Discussion.....43

Rate acceleration and UV radiation.....43

Alternative reasons for rate acceleration in plankton.....44

Future directions.....46

CHAPTER 3: Accelerated genomic evolution in metazoan parasites.....51

Abstract.....52

Introduction.....53

Materials and methods.....55

Taxonomic sample.....55

Sequence data.....57

Relative rates analysis.....59

Base composition.....60

Genome size variation.....60

Tests of association.....60

Results.....61

Discussion.....62

Parasitism and rate acceleration.....62

Parasitism and genome reduction.....67

Parasitism and other distinctive genetic features.....68

GENERAL DISCUSSION AND CONCLUSIONS.....76

Summary of results.....77

The application of molecular clocks.....78

Adaptive zones and diversification.....79

Factors influencing animal molecular evolution.....80

REFERENCES.....82

APPENDIX 1.....100

APPENDIX 2.....103

LIST OF TABLES

CHAPTER 1

Table 1. Taxonomic sample analyzed in this study, with GenBank accession numbers....25

Table 2. Support for hypotheses of branchiopod relationships....28

Table 3. Sequence statistics for the two datasets....29

Table 4. Base compositions of the 6 genes used in this study and results of chi-square tests for base homogeneity....30

CHAPTER 2

Table 1. Phylogenetically independent comparisons used in this study....48

Table 2. Relative rate calculations for 8 independent comparisons using nucleotide data....49

Table 3. Nucleotide and dinucleotide frequency data for the COI and 18S comparisons....50

CHAPTER 3

Table 1. Phylogenetically independent comparisons of parasitic and free-living taxa used in the relative rate analysis....70

Table 2. Relative rate calculations for 21 independent comparisons using nucleotide data....72

Table 3. Relative rate calculations for 14 independent comparisons using amino acid and codon sequence data....73

Table 4. Base composition for 21 independent comparisons of parasitic versus free-living metazoans....74

Table 5. Genome size variation for 10 phylogenetically independent comparisons.....75

LIST OF FIGURES

CHAPTER 1

Figure 1. Currently accepted phylogeny of the class Branchiopoda.....31

Figure 2. Phylogenetic relationships of the Branchiopoda as determined by maximum likelihood analysis of the MT dataset.....32

Figure 3. Phylogenetic relationships of the Branchiopoda as determined by maximum parsimony analysis on the MT dataset.....33

Figure 4. Phylogenetic relationships of the Branchiopoda as determined by Bayesian inference of the MG dataset.....34

LIST OF APPENDICES

Appendix 1. Taxonomic sample and sources of sequence data for Chapter 2.....100

Appendix 2. Taxonomic sample and sources of sequence data for Chapter 3.....103

General introduction: molecular rate variation

General introduction: molecular rate variation

The molecular clock and rate variation

The molecular clock hypothesis was first postulated in the 1960s when Zuckerlandl and Pauling (1962, 1965) first observed an approximately constant rate of amino acid substitution in hemoglobin and other proteins. This observed linear rate of molecular evolution across species for a given protein was later explained by Kimura and Ohta (1971), who suggested that most amino acid substitutions were neutral with respect to fitness. Under this explanation, the rate of molecular evolution is governed by the mutation rate and will be constant unless the neutral mutation rate changes. Dickerson (1971) expanded on this explanation by proposing why proteins have different substitution rates – proteins possess different proportions of neutral amino acid sites. Hence, the rate of molecular evolution is determined by both the mutation rate and by the proportion of sites not constrained by selection. These novel hypotheses conflicted with two dominant views of the nature of evolution – that there was significant variation in rates of evolution across species and time, and that most amino acid changes would be positively selected or deleterious and selected against. The implications of this surprising discovery were immediately recognized; proteins and DNA sequence data can be employed to date evolutionary events, much like the use of radioactive isotopes for dating geological events. Studies estimating times of divergence between taxa soon proliferated (e.g. Dickerson 1971; Kimura and Ohta 1973), and have continued to the present day (e.g. Bromham et al. 1998; Korber et al. 2000; Smith and Peterson 2002; Mercer and Roth 2003).

Unlike earlier studies estimating divergence times, recent authors generally acknowledge the existence of substantial variation in substitution rates. By the mid 1980s, it became widely recognized (e.g. Wu and Li 1985; Britten 1986; Vawter and Brown 1986; Li and Tanimura 1987; reviewed in Mindell and Thacker 1996; Li 1997; Bromham and Penny 2003), that rates of molecular evolution vary widely across taxonomic lineages. The disparities in rates of molecular evolution among taxa have been linked to influences at the level of mutation and selection, as early molecular clock work alluded to, which in turn, can be further broken down into several variables that determine the molecular evolutionary rate of a given lineage.

Partitions of rate variation

To acknowledge the two major processes in evolution, mutation and selection, it is convenient to divide the factors that influence variation in molecular rates among taxonomic lineages at these two levels. These factors may produce rate variation in both the overall substitution rate and in the rate of particular types of change (Mindell and Thacker 1996). Influences on mutation rate include the efficiency of DNA repair, tolerance to DNA damage, efficiency of DNA replication, frequency of DNA replication, exposure to mutagens, and the extent of base composition or codon biases. Influences at the level of selection include degeneracy of the genetic code, conservation of secondary and tertiary structure, conservation of gene function, rate of diversification and effective population size. These factors vary between species and over time, providing a basis for lineage specific rate heterogeneity (Gillespie 1991; Bromham and Penny 2003).

Determinants of rate variation

A few factors have been correlated with rate variation in several taxonomic groups and each has an established mechanistic basis. The interplay between these factors may explain differences in molecular evolutionary rates between organisms. The four most widely proposed variables include generation time, DNA repair and replication efficiency, metabolic rate, and population size. Exposure to ultraviolet (UV) radiation and to genetic arms races are two factors that are less widely considered, but they also have a solid mechanistic basis.

The generation time hypothesis (Kohne 1970; Wu and Li 1985; Ohta 1993; Li et al. 1996) proposes that organisms with short generation times will experience an increased evolutionary rate relative to those with longer generation times. This conclusion supposes that an organism which copies its DNA more frequently per unit time has a greater opportunity for replication errors to occur. This assumes that the number of germ-line divisions is approximately constant in most organisms and that most mutations are a result of replication errors. There is support for this hypothesis in a number of vertebrate groups including birds (Mooers and Harvey 1994), mammals (Bromham et al. 1996), and reptiles (Bromham 2002).

Variation in mutation rate can also be explained by differences in the enzymes responsible for DNA replication and repair (Britten 1986). The disparity in the efficiency can be attributed to the size and biochemical properties of the repair and replication machinery. Some organisms possess a complex and intricate system of repair enzymes, such as those possessed by mammals, while others, such as RNA

viruses, employ low fidelity polymerases and lack proofreading functions (Bromham and Penny 2003). This repair machinery is likely shaped by selection, balancing the high cost of replication and repair (Kunkel 1992) against the costs associated with deleterious mutations. Support for this hypothesis has been gained from comparisons of rodents and humans, because rodents possess a less efficient repair system (Hart and Setlow 1974) and exhibit a faster substitution rate (Li et al. 1996).

The metabolic rate hypothesis (Martin and Palumbi 1993; Rand 1994) posits that higher metabolic rates result in an increased level of oxygen radicals, which can enhance mutation rates. This linkage is based on early studies that revealed that concentrations of mutagenic free radicals are positively correlated with metabolic rates (Gross et al. 1969; Shigenaga et al. 1989). There is some empirical support for this hypothesis (e.g. Martin et al. 1992; Martin and Palumbi 1993), but most studies are hampered by analyses that fail to account for phylogenetic bias or multiple substitutions (Slowinski and Arbogast 1999). Several recent studies that have critically tested this hypothesis have failed to provide support for its operation (e.g. Mooers and Harvey 1994; Bromham et al. 1996; Rowe and Honeycutt 2002).

The population size hypothesis (Ohta 1972, 1992) predicts that organisms with smaller effective population sizes will experience increased fixation rates of weakly deleterious mutations, leading to an overall increase in substitution rates. This hypothesis is supported by a wealth of studies where small effective populations can be inferred, as for example, in asexuals (Moran 1996), endosymbionts (Woolfit and Bromham 2003), eusocial insects (Schmitz and Moritz 1998) and island birds (Johnson and Seger 1998).

UV radiation has a direct mutagenic effect on genomic DNA, most notably through the formation of photoproducts such as cyclobutane pyrimidine dimers (Friedberg et al. 1995). The UV exposure hypothesis predicts an association between an elevated rate of mutation in organisms with a greater exposure to solar radiation. This has been demonstrated best in 'UV-exposed' lineages of fungi (Lutzoni and Pagel 1997), but is also consistent with observations on several other groups of organisms (Smith et al. 1992; Pawlowski et al. 1997; Kooistra and Medlin 1996; Hebert et al. 2002).

Another factor that may influence rates of molecular evolution is a genetic arms race between host and parasite (Dowton and Austin 1995; Page et al. 1998). The antagonistic relationship between parasite and host may produce both an acceleration of adaptive evolution, resulting from the Red Queen process (Van Valen 1973), or direct selection for an increased rate of mutation (Haraguchi and Sasaki 1996). Comparative studies for parasites are limited, but an arms race may explain the increased substitution rate in parasitic wasps (Dowton and Austin 1995), lice (Page et al. 1998) and in at least one parasitic fly lineage (Castro et al. 2002).

Unfortunately, the comparative data necessary to thoroughly evaluate these hypotheses over a broad range of taxa are lacking for all but a few well-studied groups (i.e. vertebrates) (Bromham et al. 1996). As an alternative, it can be useful to look at convergent patterns across diverse taxa that may impose similar qualities in these or in other factors. For example, groups of organisms that share common biological attributes, such as a parasitic or a planktonic lifestyle, may share a common exposure to agents that would influence rates of evolution. For example, because of the fact that UV light is transmitted through water and that their bodies are usually transparent, one could

expect that planktonic organisms would be exposed to higher UV levels. To test whether transitions to these lifestyles are associated with shifts in rates of molecular evolution, it is necessary to use a proper phylogenetic comparative method.

Comparative method in molecular evolution

When testing for an association between species traits, it is imperative to consider their phylogenetic affinities (Harvey and Pagel 1991). Species cannot be treated as independent data points since a hierarchy is created by their phylogenetic relationships and traits may vary according to the relatedness of species. A failure to accommodate for this phylogenetic bias may lead to type I and II errors and provide misleading results. However, this problem can be overcome by the method of phylogenetically independent contrasts (Felsenstein 1985; Burt 1989; Harvey and Pagel 1991). This approach involves the selection of pairs of species or clades from a phylogeny that do not overlap on the phylogeny and that differ in the trait(s) of interest. These selected contrasts can then be used as single data points for statistical analysis to test for association between the species traits.

For example, to apply the independent comparisons method to test for an association between rates of molecular evolution and a transition in lifestyle, it is first necessary to select contrasts from a phylogeny. This phylogeny may be published previously (e.g. Mooers and Harvey 1994; Bromham 2002) or reconstructed with the sequence data used to calculate molecular rates (e.g Nunn and Stanley 1998; Woolfit and Bromham 2003). It is not required that the phylogeny is fully resolved before one selects monophyletic pairs for comparison (see Bromham et al. 1996). At this point, rates of molecular evolution must be measured and these are most commonly calculated

with relative rate tests (Sarich and Wilson 1973). This test compares the sequences (or multiple sequences) from each species or clade (ingroups) to a known outgroup, resulting in the relative amount of divergence between the two ingroups. This relative rate value is then paired with the corresponding lifestyle variable (e.g. parasitic or free-living), providing a single data point for statistical analysis. Although simple in design, this method has been demonstrated to be effective for comparing rate differences and other variables between *a priori*-defined lineages (Mooers and Harvey 1994; Bromham et al. 1996; Bromham 2002; Bromham and Cardillo 2003).

The present study

This thesis explores the effect of lifestyle transitions on the rates of molecular evolution in animals. Firstly, the phylogeny of the branchiopod crustaceans is reconstructed to elucidate the origin of a transition from the benthos to the plankton in the Cladocera. Secondly, I analyze this comparison and others to test for a general association between a planktonic lifestyle and enhanced substitution rate. Finally, I perform a second comparative analysis, testing for an association between transitions to parasitism and increased rates of molecular evolution.

In Chapter 1, the phylogenetic relationships of the Branchiopoda are explored through a comprehensive multi-gene analysis. I employ two approaches to taxon and gene sampling, generate sequence data for three genes, incorporate previously published sequence data, and use modern techniques of phylogeny reconstruction. The results corroborate previously established affinities, but suggest a unique arrangement of the cladoceran orders, and identify areas that require further work.

The lifestyle transition from a benthic to planktonic existence, and the subsequent increase in UV exposure, is investigated for an association with rates of molecular evolution in Chapter 2. This comparative analysis examines eight independent plankton-benthic contrasts, including the one determined in Chapter 1, to test for an association. In addition, I compare the base composition for each lineage to detect signatures of UV-induced mutagenesis.

Chapter 3 examines the transition to parasitism and its association with rates of sequence evolution. A comparative analysis of 21 phylogenetically independent contrasts is performed to test whether parasitic lineages have accelerated rates of molecular evolution in comparison with their free-living relatives. Also, variation in base composition and genome size is compared to further elucidate the factors affecting rate variation.

Finally, the results of the thesis are briefly discussed in the General Discussion with reference to their implications for the application of molecular clocks to calculate divergence times, and for our understanding of both the patterns of diversification, and the factors influencing the tempo of animal molecular evolution.

CHAPTER 1

Probing the origins of the branchiopod crustaceans

Abstract

The Branchiopoda are unrivalled in their variation in body forms, even within the morphologically diverse crustaceans. To fully understand the origin and evolution of these morphological reconfigurations, a robust phylogeny of the group is essential. To infer the affinities among branchiopods, I employed two approaches to taxon and gene sampling, presented new sequence data from 3 genes, incorporated previously published sequence data from 3 additional genes, and utilized modern techniques of phylogeny reconstruction. The results provided support for a number of longstanding hypotheses for the relationships among the orders, such as the Cladoceramorpha and Gymnomera concepts, as well as favouring a unique arrangement of the cladoceran orders. A few affinities remain to be resolved, particularly at the base of the Phyllopoda and within the Anomopoda, but the results suggest that increased gene sampling is recommended for future investigations of branchiopod systematics.

Introduction

In contrast to other arthropod lineages, which show limited *bauplan* diversity, the crustaceans show striking variation in body form. However, among the eight commonly recognized crustacean classes, the class Branchiopoda shows exceptional diversity, especially given the fact that it includes just 800 described species. In contrast to the relatively static body plans of the other classes, these crustaceans display marked variation in their body segmentation patterns, and in the morphology, number, and function of their limbs. While this provides a unique opportunity to investigate the exploration of phenotypic space, our understanding of the origin and evolution of these morphological reconfigurations is inhibited by our lack of a robust phylogeny for the group.

A great deal of effort has been expended to define affinities between the 8 orders and 24 families of extant branchiopods (reviewed in Spears and Abele 2000; Martin and Davis 2001). Past studies have included examinations of fossil taxa (e.g. Walossek 1993, 1995), investigations of the embryology, ontogeny, and morphology of key species (e.g. Olesen et al. 1997, 2003), and phylogenetic analyses of morphological (e.g. Olesen 1998, 2000; Negrea et al. 1999) and molecular characters (Hanner and Fugate 1997; Schwenk et al. 1998; Taylor et al. 1999; Spears and Abele 2000; Richter et al. 2001; Braband et al. 2002; Swain and Taylor 2003). In addition, a few recent studies have established the utility of several complex genetic characters, including rRNA structural characters (Swain and Taylor 2003), the distribution patterns of introns (Braband et al. 2002) and rRNA expansion regions (Crease and Taylor 1998). These

studies, a diverse assemblage in themselves, have failed to achieve the holy grail: a consensus on branchiopod relationships.

Although the phylogeny of the branchiopods has been redrawn on numerous occasions, many details remain incomplete (Fig. 1). The class is generally divided into 8 extant orders: the Anostraca (fairy shrimps); the Notostraca (tadpole shrimps); the Laevicaudata and Spinicaudata (collectively known as clam shrimps and previously classified together as the Conchostraca); and the final four orders (Anomopoda, Ctenopoda, Haplopoda, and Onychopoda) which collectively comprise the Cladocera (water fleas). The affinities between these orders remain unclear, although the placement of Anostraca as sister group to the remaining branchiopods (Phyllopoda) is well supported (Negrea et al. 1999; Olesen 1998; Spears and Abele 2000), as is the monophyly of the Cladocera (Taylor et al. 1999; Spears and Abele 2000; Braband et al. 2002). Furthermore, *Cyclestheria hislopi*, formerly placed within the Spinicaudata, is unquestionably the sister clade to the Cladocera (= Cladoceromorpha) (e.g. Crease and Taylor 1998; Ax 1999; Spears and Abele 2000). Within the Cladocera, several studies have corroborated the monophyly of the predatory water fleas (Haplopoda + Onychopoda = Gymnomera) (Richter et al. 2001; Swain and Taylor 2003), but their relationship to the Anomopoda and Ctenopoda, as well as the affinities of varied cladoceran families, remains under debate.

This study, which expands on previous molecular studies, seeks to obtain a well-supported phylogeny for the branchiopod orders and families, particularly within the speciose Cladocera. I follow two strategies of data collection, provide new sequence data for three genes, incorporate previously published sequence data, and employ

modern techniques of tree reconstruction. These investigations are used to corroborate previous relationships, to introduce novel ones, and to identify areas of ambiguity that require further study.

Materials and Methods

Taxonomic sample & gene selection

This study employed two approaches to data collection. The first, a 'more taxa' approach (e.g. Pollock et al. 2002), involved the collection of sequence data from a broad diversity of branchiopods for three gene fragments. The second approach, a 'more genes' approach (e.g. Rosenberg and Kumar 2001), added sequence data for three additional genes for a subset of the original taxa.

In the 'more taxa' (MT) approach, I examined 56 taxa that included representatives from 22 of the 24 recognized branchiopod families (Martin and Davis 2001, Table 1). In addition to maximizing taxonomic breadth, an effort was made to include multiple species of key genera (e.g. *Lynceus*) and speciose taxa (e.g. Chydoridae). Specimens were collected from numerous locations in North America, South America, Europe, and Australia. The complete collection and locality details are available from the Cladoceran website (http://www.cladocera.uoguelph.ca).

The malacostracan *Anaspides tasmaniae* was employed as the outgroup in the analysis of these data. While prior studies have strongly supported the monophyly of the Branchiopoda (e.g. Sanders 1963, Wingstrand 1978, Walossek 1993, Spears and Abele 2000), their relation to other Crustacea remains uncertain (reviewed in Martin and Davis 2001, and references therein). *Anaspides tasmaniae* certainly lies outside the

branchiopods, but it may not be their closest relative. However, my preliminary analyses indicated that a) when multiple outgroups were included, malacostracans were the sister taxa to a monophyletic branchiopod clade and b) that topology was unaffected by the choice of single or multiple outgroups from other crustacean classes (trees not shown).

The MT approach involved the analysis of sequences from two mitochondrial genes; *cytochrome c oxidase subunit 1* (COI) and the *large subunit 16S rRNA* (16S), and a nuclear gene; the *small subunit 18S rRNA* (18S). These three genes were chosen because of their wide use in past studies of arthropod phylogenetics (e.g. Giribet et al. 2001), including studies of branchiopod relationships (Remigio and Hebert 2000; Spears and Abele 2000; Cristescu and Hebert 2002; Sacherova and Hebert 2003). Furthermore, these three loci provide phylogenetic signal over varying time scales, aiding the resolution of both shallow and deep nodes. 135 sequences were generated for this dataset, referred to as the 'MT' dataset.

In the 'more genes' (MG) approach, I gathered additional sequence information for a subset of 17 taxa, representing all 8 orders of branchiopods. These genes included the mitochondrial *small subunit 12S rRNA* (12S), the nuclear *large subunit 28S rRNA* (28S), and the nuclear protein-coding gene *elongation factor 1 alpha* (EF-1α). These gene regions have proven informative in the studies from which the sequences were extracted (Hanner and Fugate 1997; Taylor et al. 1999; Braband et al. 2002; Cristescu and Hebert 2002; Swain and Taylor 2003; Sacherova and Hebert 2003). Due to the lack of a suitable outgroup species represented by all 6 genes, I rooted all trees with the anostracan *Artemia franciscana*. I term this the 'MG' dataset.

Molecular techniques and sequence alignments

Genomic DNA was extracted from whole animals using 25-50 μL aliquots of proteinase K extraction buffer and the method described in Palumbi (1996). The primer pairs LCO1490/HCO2198 (Folmer et al., 1994) and 16Sar/16Sbr (Palumbi, 1996) were used to PCR amplify a 680 base pair (bp) fragment of COI and a 570 bp fragment of 16S, respectively. An approximately 1995 bp fragment of 18S was amplified with the primers 9F (5'-TGG GGA TCA TTG CAG TTC CCA ATC-3'; T. Crease, pers. comm.) and 2004R (Crease and Colbourne, 1998) with about 800 bp near the 5' terminus targeted for sequencing. The 50μl PCR reactions contained 0.5-2.0 μl (out of 25-50 μL) of DNA template, 5.0 μl 10x PCR buffer (Boehringer-Mannheim), 0.2 μM of each primer, 2.2 mM $MgCl_2$, 0.2 mM of each dNTP, and 1 unit of *Taq* DNA polymerase. The PCR conditions for COI and 16S consisted of 1.5 min at 94°C, followed by 35 cycles of 45 s at 93°C, 1 min at 50°C, and 1 min at 72°C, followed by one cycle of 5 min at 72°C. The PCR conditions for 18S consisted of 1 cycle at 94°C, 35 cycles of 30 s at 93°C, 30 s at 50°C, and 3 min at 72°C, followed by one cycle of 5 min at 72°C. PCR products were excised from agarose gels and purified using the Qiaex II gel extraction kit (Qiagen Inc.) and sequenced using an ABI 377 automated sequencer and ABI prism BigDye terminator 3 sequencing kit (Applied Biosystems). Gene products were sequenced in both directions or twice in the same direction whenever ambiguous sites were encountered. Some sequences used in my analysis were obtained from previously published studies (Table 1).

DNA sequences were initially aligned in Sequence Navigator (Applied Biosystems). The alignments for the ribosomal genes required adjustments with

reference to proposed secondary structure models (De Rijk et al. 2000, Van de Peer et al. 2000; Crease and Colbourne, 1998; Taylor, Finston and Hebert, 1998). Sites within the ribosomal genes that were not easily aligned were excluded from subsequent analyses. The sequence alignments are available from the Cladoceran website (http://www.cladocera.uoguelph.ca).

Phylogenetic analysis

In order to reconstruct the phylogenetic relationships of the ingroup taxa, I concatenated the nucleotide sequence alignments for the three and six genes included in the MT and MG datasets respectively. Tree-building was performed by maximum likelihood (ML), Bayesian inference (BI) and maximum parsimony (MP). These three techniques were used because concordance among different analytical approaches strengthens support for the tree (Cunningham 1997) and because there are varied opinions on how to best reconstruct phylogenies (Crandall et al. 2000). The best-fit model of sequence evolution was selected by analyzing distance-based topologies with hierarchical likelihood ratio tests using ModelTest 3.06 (Posada and Crandall 1998) following the procedure outlined by Huelsenbeck and Crandall (1997). The ML analysis was performed in PAUP* v4.0b10 (Swofford 2001) using the model and parameters estimated in ModelTest and the heuristic search option (10 replicates, 1 tree held per replicate, sequences added at random, branch swapping by nearest neighbour interchanges, starting tree obtained by neighbour-joining).

The BI analysis was computed in the program MrBayes 2.01 and 3.0b4 (Huelsenbeck and Ronquist 2001). Again, the model and parameters estimated by ModelTest were used for the analysis. Three independent runs, each consisting of four

Markov chains, were run for 1 050 000 generations, with the first 50 000 generations discarded as the burn-in. Each run was inspected to ensure that likelihood stationarity was reached during burn-in, and that parameters and posterior probabilities were consistent between runs. MP trees were estimated with PAUP* using the unweighted heuristic search option (1000 replicates, 100 trees held per replicate, sequences added at random, and tree bisection-reconnection branch swapping) on the parsimony-informative sites.

I performed a goodness-of-fit (X^2) test, as implemented in PAUP*, to test for stationarity in base composition. I evaluated the strength of the phylogenetic signal in the dataset by calculating the g_1 skewness statistic (Hillis and Huelsenbeck 1992) and performing a relative apparent synapomorphy analysis (RASA) (Lyons-Weiler et al. 1996). To assess confidence in the phylogenies, I performed nonparametric bootstrapping (100 pseudoreplicates for ML, 1000 for MP). Finally, incongruence length difference (ILD) tests (Farris et al. 1994) were employed to determine the congruence of phylogenetic signal from the different genes. These tests had no bearing on my decision to combine the genes for a total evidence approach (Kluge 1989), but instead, to investigate the nature of the phylogenetic signal (Remsen and DeSalle 1998).

Results

More taxa approach

The final alignment for the MT dataset was 1546 bp in length (689 variable sites), comprised of 639, 353, and 554 bp long fragments of COI, 16S, and 18S, respectively. The hierarchical likelihood ratio tests indicated that the best-fit model for

subsequent analysis was the general time reversible model with invariable sites and gamma shape parameter (GTR + I + G) with the following parameters selected: unequal base frequencies: A = 0.34, C = 0.14, G = 0.13, T = 0.39; six substitution categories: A→C = 0.46; A→G = 4.85, A→T = 0.72, C→G = 1.05, C→T = 5.97, G→T = 1.00; proportion of invariant sites = 0.48; and gamma distribution shape parameter = 0.40.

The Bayesian and maximum likelihood analyses produced nearly the same topology (Fig. 2). The node support, assessed with posterior probabilities, was generally high (> 80) for nodes at the family level and above. The parsimony analysis resulted in 26 equally parsimonious trees with a length of 6536 steps (consistency index = 0.18; retention index = 0.40). In contrast to the BI and ML tree, the parsimony analysis failed to recover several nodes (bootstrap percentages < 50) such as the Cladocera and Anomopoda (Fig. 3).

More genes approach

The length of the final alignment for the MG dataset, including the 1546 bp from the MT dataset, was 4096 bp (1543 variable sites). The sequence data, taken from GenBank and previous studies (Table 1), consisted of 1300 bp of 28S, 250 bp of 12S, and 1000 bp of EF-1α. The best-fit model selected for this dataset was also the GTR + I + G model, with the following parameters: unequal base frequencies: A = 0.25, C = 0.23, G = 0.26, T = 0.26; six substitution categories: A→C = 1.09; A→G = 4.73, A→T = 4.22, C→G = 1.42, C→T = 9.45, G→T = 1.00; proportion of invariant sites = 0.45; and gamma distribution shape parameter = 0.69.

All three tree-building approaches used on the MG dataset produced an identical topology (Fig 4). There was a single most parsimonious tree that is 4950 steps in length

(consistency index = 0.48; retention index = 0.35) with only two nodes not recovered with bootstrap support > 50 (Spinicaudata + Cladoceromorpha and Chydoridae + Daphniidae). In general, all three trees had modest to strong node support.

Comparison of approaches

The two approaches to taxon sampling had a significant impact on the specific hypotheses supported by the results, as well as the strength of these nodes (Table 2). To a lesser extent, and mostly limited to the MT dataset, the tree building approach also impacted the outcome (Table 2). In general, the MT dataset provided good resolution of relationships at the family level and below, but only moderate or poor resolution of deeper divergences. The MG dataset on the other hand, provided good resolution of interordinal divergences, particularly within the Cladoceromorpha. However, there was disagreement between the two approaches on several important nodes, for example, the position/monophyly of Cladocera and Gymnomera (Table 2).

It is likely that several factors contribute to the discrepancies in topology obtained with the two datasets. The MG dataset is roughly twice as large as the MT dataset, both in variable and parsimony informative characters (Table 3). My calculation of the g_1 skewness statistics, and the RASA test statistics (Table 3) indicate that there is significant phylogenetic signal in the two datasets, likely ruling this out as a factor in their incongruence. In contrast, chi-square tests (Table 4) provide evidence for heterogenous nucleotide composition across the taxa in the MG dataset, which appears to derive from the two protein-coding genes. In addition, the partition homogeneity tests indicated heterogeneity of phylogenetic signal from the various genes. I detected significant heterogeneity ($p \leq 0.01$) in all comparisons performed: all 3 genes (MT

dataset), all six genes (MG dataset), mitochondrial genes only, nuclear genes only, protein-coding genes only, and ribosomal genes only.

Discussion

Branchiopod interordinal affinities

This study provides support for a number of longstanding hypotheses concerning higher level branchiopod relationships. The monophyletic status of Anostraca, Notostraca, Laevicaudata, and Spinicaudata (excluding *Cyclestheria)* are supported in all analyses and using both datasets, consistent with recent studies (e.g. Braband et al., 2002; Spears and Abele, 2000). The Cladocera is also found to be monophyletic with the MG dataset and in the maximum parsimony analysis of the MT dataset. Conversely, the clam shrimp *Cyclestheria* groups among the Cladocera in the maximum likelihood and Bayesian trees constructed from the MT dataset, rendering the latter paraphyletic. The low node support of this placement, and the deep divergences in the MT trees in general, cause us to favour cladoceran monophyly and instead interpret this as support for the Cladoceromorpha concept (Ax, 1999). All other trees support a *Cyclestheria* + Cladocera sister group relationship, which now seems uncontroversial (e.g. Crease and Taylor, 1998; Spears and Abele, 2000; Swain and Taylor, 2003).

As in Braband et al. (2002) and Spears et al. (2000), I am unable to determine the exact relationships among the large branchiopod orders. The MG dataset suggests that the Laevicaudata may be the sister clade to the remaining groups of the Phyllopoda, whereas maximum parsimony analysis of the MT dataset places it as the sister group to the Notostraca. Both of these hypotheses are congruent with the analysis of Braband et al. (2002). Also consistent with Braband et al. (2002), as well as Spears and Abele

(2000), is the close affinity between Spinicaudata and Cladoceromorpha suggested by analysis of the MG dataset. This suggestion challenges the traditional "Conchostraca" taxon (Negrea et al., 1999; Schram, 1986; Walossek, 1993), but requires further confirmation.

Relationships within the Cladocera

My analysis also provides support for previous hypotheses concerning relationships within the Cladocera. Inferences employing both of the datasets and all three tree-building approaches support the monophyly of the orders Onychopoda, Ctenopoda, and Anomopoda. The MG analysis further supports two other hypotheses; the Calyptomera, comprised of the Ctenopoda and the Anomopoda (Negrea et al., 1999; Spears and Abele, 2000), and the Gymnomera (e.g. Richter et al., 2001; Swain and Taylor, 2003). In addition, the sister grouping of Calyptomera and Gymnomera is very well-supported, providing a new hypothesis of cladoceran relationships.

The taxon sampling within the Anomopoda in my MT dataset allows inferences about the affinities within this large order. Firstly, it would appear that the Moinidae, recently demoted to subfamily status within the Daphniidae (Fryer, 1995; Olesen, 1998), actually warrant their traditional family status, since they appear most closely related to the Bosminidae (ML and BI), and perhaps some macrothricid lineages (MP). Secondly, the two speciose families, Chydoridae and Macrothricidae, are paraphyletic under all methods, suggesting that a revision of these two families is needed before anomopod affinities can be clarified.

Strategies of taxon and gene sampling

How best to approach data collection for phylogenetic estimation remains a contentious issue and incomplete taxon sampling is often cited as a major source of error in phylogenetic studies (reviewed in Graybeal, 1998; Poe, 1998; Pollock et al., 2002; Rosenberg and Kumar, 2001). Increasing taxon sampling (e.g. Pollock et al., 2002) or increasing sequence length (e.g. Rosenberg and Kumar, 2001) are two opposing strategies, each possessing merit and empirical support. In the present study, I naturally do not have the 'true phylogeny' with which to determine with certainty the superior strategy, but I can evaluate them with three metrics that are possessed: node support, congruence across phylogenetic techniques, and ability to reconstruct generally accepted and well-supported relationships. For addressing phylogenetic affinities above the family level, it is apparent that the 'more genes' approach provides stronger support on the nodes shared in both datasets, shows greater congruence across tree-building methods (evident by the inference of the same topology with all three techniques), and displays a better ability to reconstruct well-supported relationships (e.g. Cladocera and Gymnomera). Previous work suggests that for any given phylogenetic problem, there is a threshold amount of sequence data below which an increase in taxonomic sampling does not improve, or may even decrease, phylogenetic accuracy (Cummings et al., 1995; Mindell et al., 1997). Following this suggestion, it is unclear if the threshold was reached in the present study with 6 genes, but it is almost certain that it was not reached with 3 genes. For this reason, future work will likely benefit most from increased sequence sampling for the currently used exemplar taxa, as opposed to more extensive taxon sampling.

Conclusions

In summary, I have provided the most comprehensive molecular study of branchiopod relationships to date, both in terms of taxonomic representation and the amount and diversity of sequence data. My study highlights the importance of sampling strategies for future investigations of branchiopod systematics, as well as phylogenetic analyses in general. A few details of the branchiopod phylogeny remain incomplete, particularly near the root of the Phyllopoda and among families within the Anomopoda, but the lineage relationships have been clarified. In the near future, we may attain a complete and robust phylogeny, finally providing the vantage point needed to interpret the striking morphological reconfigurations of the branchiopods that hampered the creation of a phylogeny in the first place.

Table 1. Taxonomic sample analysed in this study, with GenBank accession numbers. Asterisks (*) denote sequences novel to this study and dashes (-) indicate missing sequences. The genes sampled were *cytochrome c oxidase subunit 1* (COI), the *large subunit 16S rRNA* (16S), the *small subunit 18S rRNA* (18S), *elongation factor 1 alpha* (EF-1α), the *small subunit 12S rRNA* (12S), and the *large subunit 28S rRNA* (28S). For the EF-1α, 12S, and 28S genes, the same species or a congeneric taxon was acquired from GenBank and/or previous studies.

Taxonomy	Species	COI	16S	18S	EF-1α	12S	28S
Class Branchiopoda							
Subclass Sarcostraca							
Order Anostraca							
Family Artemiidae	*Artemia franciscana*	NC_001620	AF209051	X01723	X03349	X69067	a ,AY137143
Family Thamnocephalidae	*Thamnocephalus platyurus*	AF209066	AF209057	AF144217	-	-	-
	Branchinella pinnata	AF308940	*	*	-	-	-
Family Branchionectidae	*Branchinecta paludosa*	AF209064	AF209055	AF144206	-	-	-
Family Streptocephalidae	*Streptocephalus dorothae*	AF209065	AF209056	AF144218	-	-	-
Family Branchipodidae	*Parartemia contracta*	AF209059	AF209048	*	-	-	-
Family Chirocephalidae	*Artemiopsis stefanssoni*	AF209062	AF209053	*	-	-	-
	Eubranchipus sp.	AF209061	AF209052	*	-	-	-
Family Polyartemiidae	*Polyartemiella hazeni*	AF209063	AF209054	*	-	-	-
Subclass Calamanostraca							
Order Notostraca							
Family Triopsidae	*Lepidurus* sp.	AF209065	AF209058	AF144212	AF526293	AF494483	AF209047, AY137138
	Lepidurus couessi	*	*	*	-	-	-
	Triops sp.	*	*	*	U90058	AF494482	a,AY137137
	Triops australiensis	*	*	*	-	-	-
	Triops sp. nov.	*	*	*	-	-	-
Subclass Diplostraca							
Order Laevicaudata							
Family Lynceidae	*Lynceus* sp. 1	*	*	AF144215	AF526294	AF494479	a ,AY137136
	Lynceus sp. 2	*	*	*	-	-	-

Taxonomy	Species	COI	16S	18S	EF-1α	12S	28S
Order Spinicaudata							
Family Caenestheriidae	*Caenestheriella setosa*	*	*	*	-	-	-
	Caenestheriella sp.	*	*	*	-	-	-
Family Limnadiidae	*Limnadia* sp.	*	*	*	AF063412	AF494471	AF532886
"Cyclestherida"							
Family Cyclestheriidae	*Cyclestheria hislopi*	*	*	AF144209	AF526292	AF494478	AF532878
"Cladocera"							
Order Anomopoda							
Family Daphniidae	*Daphnia pulex*	NC_000844	*	AF014011	-	-	-
	Scapholeberis rammneri	*	*	*	AF526282	AF494465	AF532880
	Simocephalus vetulus	*	*	AF144216	AF526281	AY009492	AF532887
	Ceriodaphnia sp.	*	*	AF144208	AF526283	AF494466	AF532889
Family Bosminidae	*Bosmina* sp. 1	*	*	*	AF526284	AF494467	AF482744
	Bosmina sp. 2	*	*	*	-	-	-
Family Macrothricidae	*Ophryoxus gracilis*	*	*	*	-	-	-
	Ilyocryptus sp.	*	*	*	-	-	-
	Acantholeberis curvirostris	*	*	*	-	-	-
	Macrothrix sp.	*	*	*	-	-	-
	Drepanothrix dentata	*	*	*	-	-	-
Family Chydoridae							
Subfamily Chydorinae	*Chydorus sphaericus*	*	*	*	AF526286b	AF494469b	AF532891
	Alonella exigua	*	*	*	-	-	-
	Pleuroxus denticulatus	*	*	*	-	-	-
	Dunhevedia crassa	*	*	*	-	-	-
Subfamily Aloninae	*Alona setulosa*	*	*	*	-	-	-
	Camptocercus rectirostris	*	*	*	-	-	-
	Acroperus harpae	*	*	*	-	-	-
	Graptoleberis testudinaria	*	*	*	-	-	-
Subfamily Sayciinae	*Saycia cooki*	*	*	*	-	-	-
Subfamily Eurycerinae	*Eurycercus longirostris*	*	*	*	AF526285	AF494468	a
	Eurycercus glacialis	*	*	*	-	-	-
Family Moinidae	*Moina* sp. 1	*	*	*	-	-	-
	Moina sp. 2	*	*	*	-	-	-

Taxonomy	Species	COI	16S	18S	EF-1α	12S	28S
Order Onychopoda							
Family Polyphemidae	*Polyphemus pediculus*	*	*	*	-	-	-
Family Cercopagidae	*Cercopagis pengoi*	*	*	*	-	-	-
	Bythotrephes cederstroemi	*	*	AF144207	-	-	-
Family Podonidae	*Evadne spinifera*	*	*	*	AF526288	AY009498	AF532906,AY137167
	Podonevadne sp.	*	*	*	-	-	-
	Podon leuckarti	*	*	*	AF526287	AY009496	AF532901, AY137147
	Pleopis polyphemoides	*	*	*	-	-	-
Order Ctenopoda							
Family Sididae	*Sida crystallina*	*	*	*	AF526280	AY009489	AF532873
	Diaphanosoma sp.	*	*	AF144210	AF526279	AY009490	AF532910,AY137155
Family Holopedidae	*Holopedium gibberum*	*	*	*	-	-	-
Order Haplopoda							
Family Leptodoridae	*Leptodora kindtii*	*	*	AF144214	AF526278	AY009488	AF532877
Class Malacostraca							
Order Syncarida	*Anaspides tasmaniae*	*	*	*	-	-	-

a sequence was not deposited in GenBank; copied from the supplementary materials of Taylor et al (1999)

b the confamilial and closely-allied *Pseudochydorus globosus* was used to represent *Chydorus sphaericus* for EF-1α and 12S

Table 2. Support for hypotheses of branchiopod relationships. Examples of previous studies that support the hypotheses are given. Support from the different trees estimated in the present study is also given, where Y signifies support for the hypothesis and N signifies no support. Trees with strong support (posterior probabilities or bootstrap percentages > 90) from the present study are denoted by an asterisk (*).

Hypothesis	Examples of previous support	Support from present study					
		More taxa approach (MT dataset)			More genes approach (MG dataset)		
		MP	ML	BI	MP	ML	BI
Phyllopoda	1, 2, 3, 4, 5	N	N	N	-	-	-
Diplostraca	1, 2, 3, 5, 6	N	N	N	N	N	N
Conchostraca	1, 2, 4	N	N	U	N	N	N
Cladoceromorpha	4, 5, 6, 7, 8, 9, 10	Y	Y	Y*	Y	Y*	Y*
Cladocera	2, 3, 4, 5, 6, 11	Y	N	N	Y*	Y*	Y*
Spinicaudata + Cladoceromorpha	5, 6	N	N	N	Y	Y	Y*
Ctenopoda + Anomopoda + Onychopoda (=Eucladocera)	5, 12, 13, 14	Y	N	N	N	N	N
Anomopoda + Haplopoda + Onychopoda	11, 15	N	N	N	N	N	N
Ctenopoda + Anomopoda (= Calyptomera)	4, 5	N	N	N	Y	Y	Y*
Ctenopoda + Haplopoda + Onychopoda	6	N	N	N	N	N	N
Gymnomera	3, 10, 11, 15, 16	N	N	N	Y	Y	Y*
Calyptomera + Gymnomera	-	N	N	N	Y*	Y*	Y*
Moininae within Daphniidae	3, 17	N	N	N	-	-	-
Radopoda	18	N	N	N	-	-	-
Chydoridae	3, 17	N	N	N	-	-	-
Macrothricidae paraphyly	3, 13	Y	Y	Y*	-	-	-

Sources are as follows: 1. Schram 1986; 2. Walossek 1993; 3. Olesen 1998; 4. Negrea et al. 1999; 5. Spears and Abele 2000; 6. Braband et al. 2002; 7. Crease and Taylor 1998; 8. Ax 1999; 9. Taylor et al. 1999; 10. Swain and Taylor 2003; 11. Martin and Cash-Clark 1995; 12. Eriksson 1934; 13. Wingstrand 1978; 14. Bowman and Abele 1982; 15. Schwenk et al. 1998; 16. Richter et al. 2001; 17. Fryer 1995; 18. Dumont and Silva-Briano 1998.

Table 3. Sequence statistics for the more taxa (MT) and more genes (MG) datasets. Variable and parsimony informative characters and results of the g1 skewness test and RASA test for each dataset is given.

Dataset	Taxa	bp	Variable sites	Informative sites	g_1 statistic	*P*	t_{RASA}	*P*
MT dataset	56	1546	686	588	-0.52	<0.01	15.9	<0.001
MG dataset	17	4096	1543	1087	-0.94	<0.01	12.8	<0.001

Table 4. Base compositions of the 6 genes used in this study and results of chi-square tests for base homogeneity.

	bp	A	C	G	T	X^2	p
COI	639	0.248	0.194	0.205	0.354	306.5	<0.01
16S	353	0.288	0.157	0.239	0.316	95.8	1.00
18S	554	0.255	0.244	0.267	0.233	21.9	1.00
MT dataset	1546	0.260	0.203	0.235	0.302	154.5 (165)	0.71
28S	1300	0.239	0.245	0.337	0.180	27.5	0.99
12S	250	0.318	0.196	0.206	0.281	22.5	1.00
EF-1α	1000	0.237	0.292	0.254	0.216	85.6	<0.01
MG dataset	4096	0.252	0.237	0.266	0.245	66.5 (48)	0.04

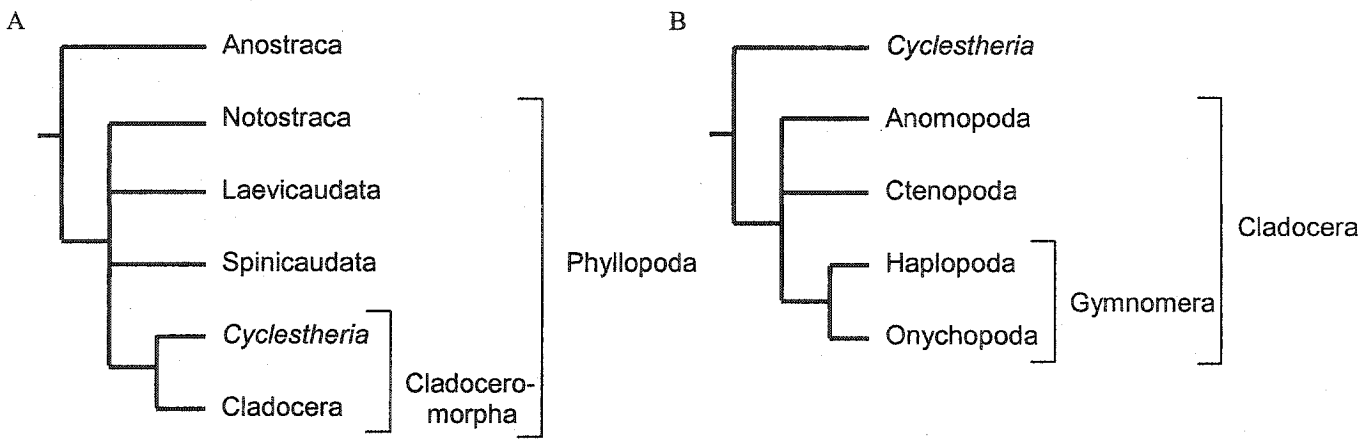

Fig. 1. Currently accepted phylogeny of the class Branchiopoda. (A) Relationships of the 'large branchiopod' orders and the Cladocera. (B) Relationships among the cladoceran orders. Affinities that have not been reliably resolved are drawn as polyphyletic.

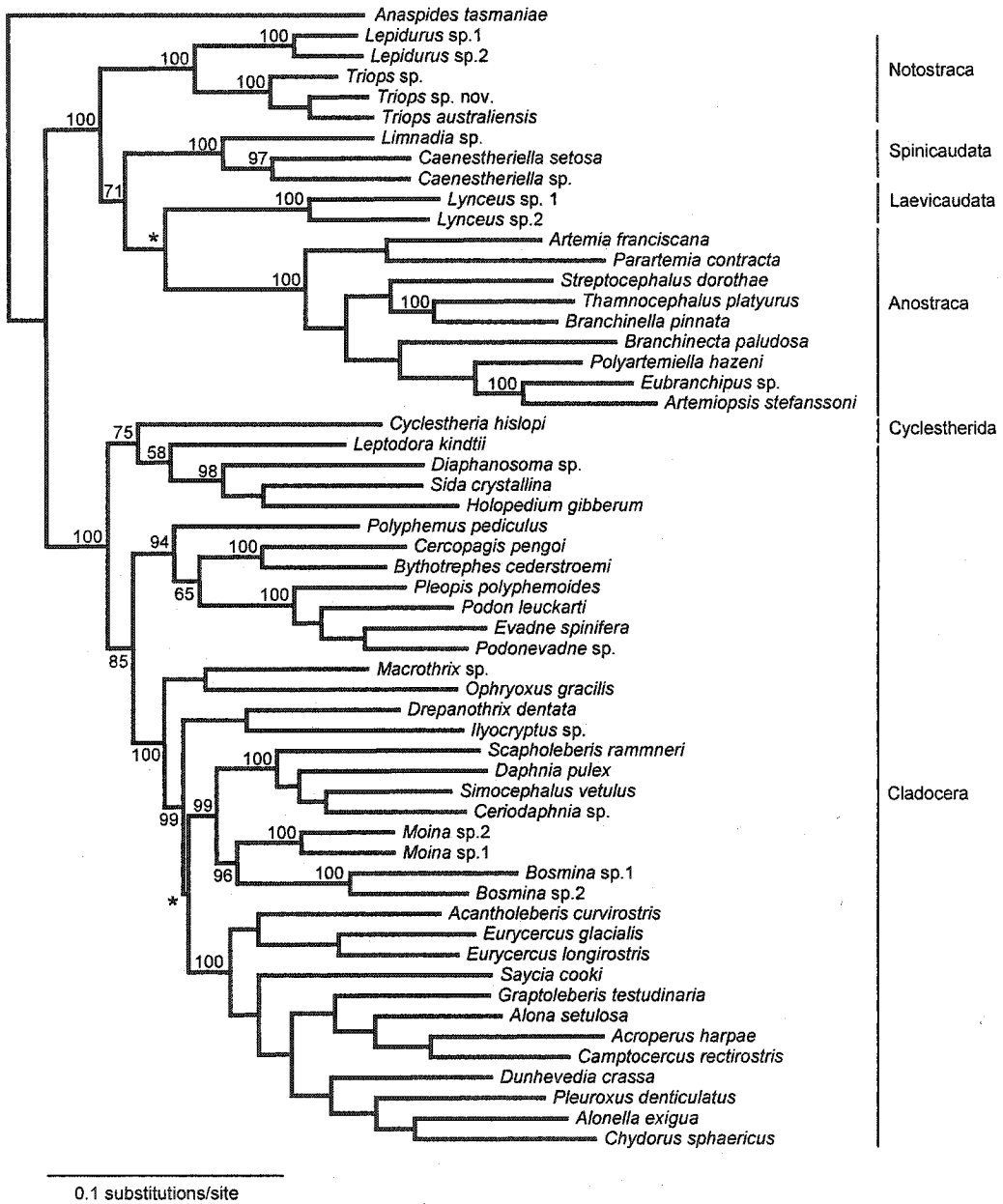

Fig. 2. Phylogenetic relationships of the Branchiopoda as determined by maximum likelihood analysis (-lnL = 35100.7) of the combined COI, 16S and 18S ('more taxa') dataset for 56 taxa. The Bayesian inference tree was of identical topology except for two nodes that were left unresolved and are marked by asterisks (*). The tree was rooted with the outgroup *Anaspides tasmaniae*. Branch lengths are proportional to reconstructed distances. Posterior probabilities are given for the nodes at the family level or above.

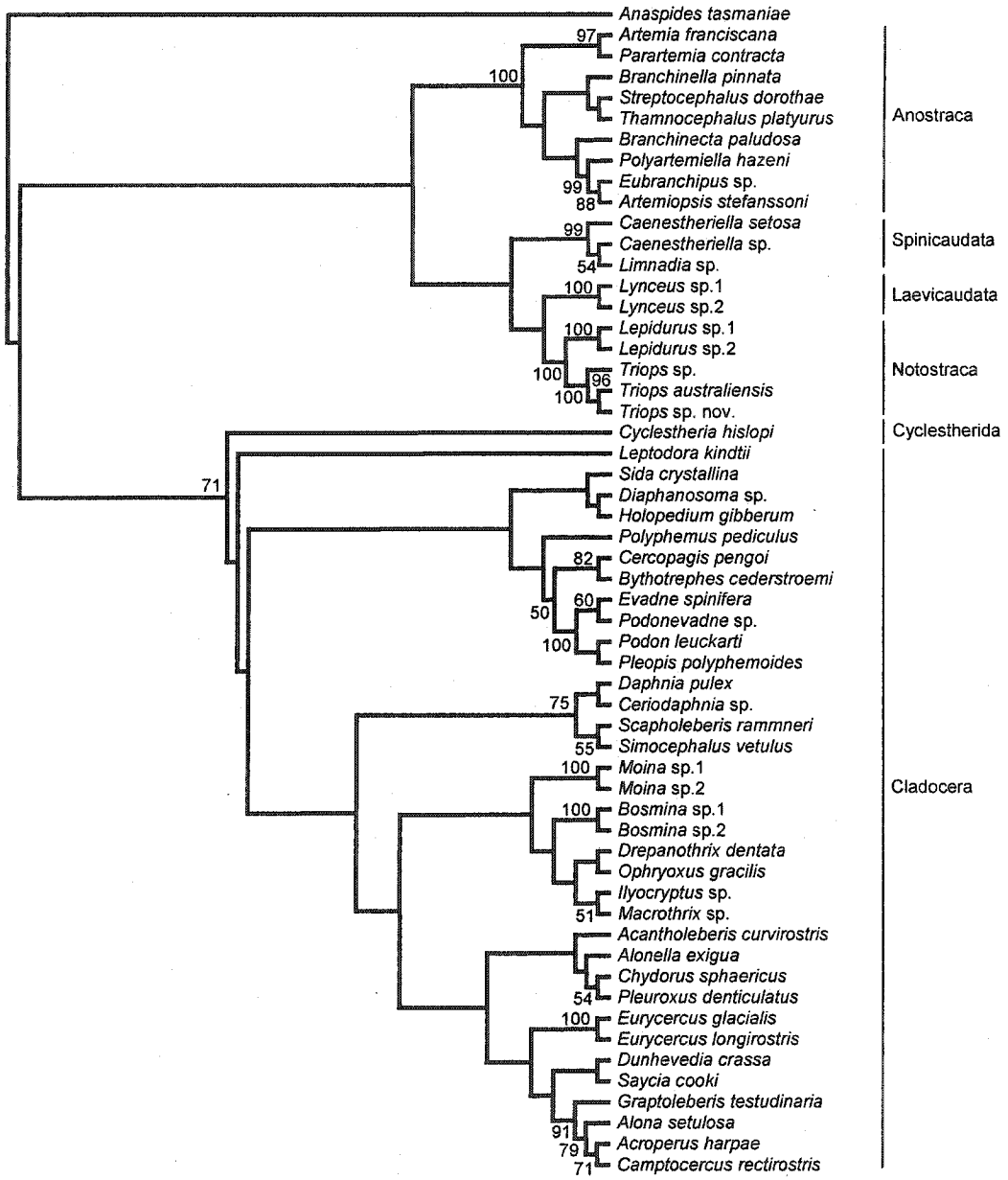

Fig. 3. Phylogenetic relationships of the Branchiopoda as determined by maximum parsimony of the combined COI, 16S and 18S ('more taxa') dataset for 56 taxa. The majority rule consensus cladogram of the 26 equally parsimonious trees (length = 6536) is shown. The tree was rooted with the outgroup *Anaspides tasmaniae*. MP bootstrap percentages are given for the resolved nodes with values > 50.

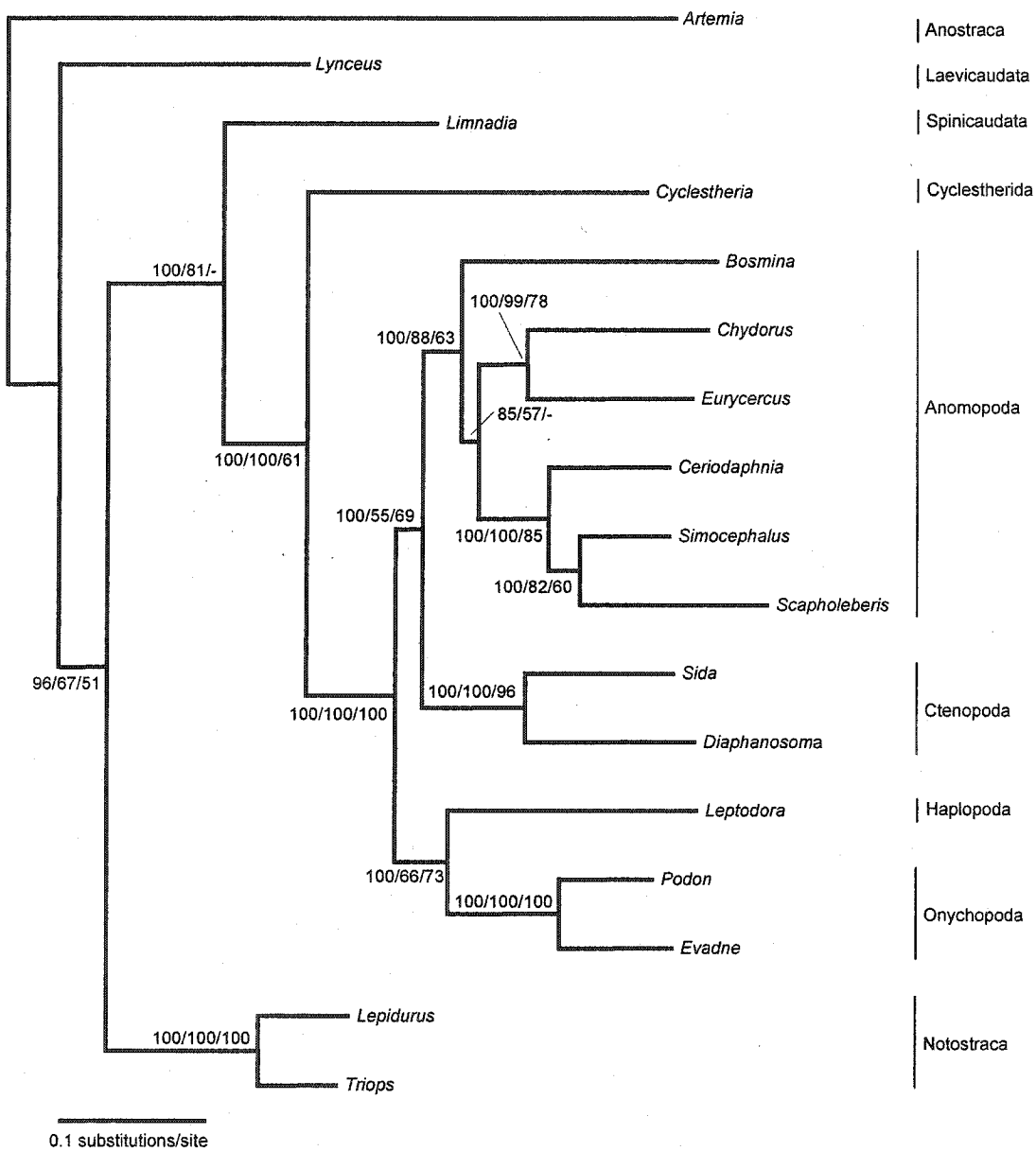

Fig. 4. Phylogenetic relationships of the Branchiopoda as determined by Bayesian inference of the combined COI, 16S, 18S, 28S, 12S, and EF-1α ('more genes') dataset for 17 taxa. Topologies of the maximum likelihood tree (-lnL = 29471.7) and the single most parsimonious tree (length = 4950) were identical. The tree was rooted with the anostracan *Artemia franciscana*. Branch lengths are proportional to reconstructed distances. Branch support values are given for all nodes and are given as BI posterior probabilities / ML bootstrap probabilities / MP bootstrap probabilities. A dash (-) indicates a MP bootstrap probability of < 50.

CHAPTER 2

UV exposure, the planktonic lifestyle, and accelerated rates of molecular evolution

Abstract

Increased exposure to ultraviolet (UV) radiation may cause tremendous damage to aquatic organisms. The most severe biological effect that UV radiation imposes on biological systems is damage to genomic DNA, and previous studies have linked this UV-induced mutagenesis to increased rates of molecular evolution. I tested for an association between UV exposure and rates of molecular evolution by investigating several planktonic animal groups under the assumption that pelagic organisms should have an elevated exposure to solar radiation in comparison to their benthic relatives. The results provided clear evidence that pelagic organisms in a diverse range of lineages have an increased rate of nucleotide substitution relative to their benthic allies. Furthermore, the results suggested that the increased mutagenesis may be linked to the formation of UV-induced photoproducts.

Introduction

Recent concerns over ozone depletion and the subsequent increase in UV radiation exposure have generated a flux of research examining the effects of UV radiation on the biosphere (e.g. Tevini 1993; Biggs and Joyner 1994). Ultraviolet-B (UV-B) radiation (280-315 nm), the most energetic and biologically destructive solar radiation (Molina and Molina 1986), has been shown to have adverse consequences for a wide range of organisms (e.g. Witkin 1976, Bernstein 1981, Hatnes and Kunz 1981, Friedberg et al. 1995). In the world's oceans and inland water bodies, UV-B radiation often penetrates deeply, potentially causing biological damage to depths below 30m (Smith and Baker 1979; Calkins and Thordardottir 1980; Karentz et al. 1991; Smith et al. 1992; Calkins 1982). Studies have demonstrated that elevated UV-B levels have a detrimental effect on both primary producers (e.g. diatoms) (reviewed in Weiler and Penhale 1994) and heterotrophs (ichthyoplankton and krill) (Malloy et al. 1997), confirming suggestions that changes in UV-B exposure may perturb ecosystems as a whole (Karentz 1991; Roberts 1989).

UV-B radiation affects cellular systems in numerous ways, including the induction of photoreactions in proteins, lipids, membranes, and nucleic acids (Sauerbier 1976; Kantor and Hull 1979; Harm 1980; Tevini 1993). This may have severe repercussions for the organism, impeding cellular processes such as replication, transcription, cell division, transport, and photosynthesis (Strid et al. 1994; Lesser 1996; Hoeijmakers 2001). The most severe long-term biological effect that UV-B radiation imposes on biological systems is damage to genomic DNA. DNA exposed to UV-B photons results in a variety of damaging photoproducts, including cross-links and strand

breaks, and the most significant in terms of prevalence and injury, the formation of cyclobutane pyrimidine dimers (CPDs) (Friedberg et al. 1995). CPDs and other DNA lesions may be repaired or tolerated (reviewed in Friedberg et al. 1995; Hoeijmakers 2001; Friedberg 2003), or prevented in advance by a variety of behavioural, physiological or chemical mechanisms (reviewed in Epel 2003), but such processes impose costs and may decrease fitness. Furthermore, prevention and repair is not without error; mutations persist, providing new genetic diversity.

This link between increased UV-B radiation and mutagenesis has led to the hypothesis that ultraviolet radiation may contribute to the heterogeneity in molecular evolutionary rates between taxonomic lineages. For example, molecular studies on foraminiferans (Pawlowski et al. 1997) and diatoms (Kooistra and Medlin 1996) have revealed accelerated rates of evolution in planktonic taxa compared to their benthic relatives. Sea urchins (Smith et al. 1992) and branchiopod crustaceans (Hebert et al. 2002) also display accelerated substitution rates that are consistent with a 'UV exposure hypothesis'. Lutzoni and Pagel (1997) not only demonstrated that 'UV-exposed' lineages of fungi have an increased molecular rate, but also that they have a higher proportion of nucleotide substitutions at AA and TT sites, signatures of the formation of CPDs. Beyond these few and unrelated groups, little work has been dedicated to investigating the apparent association between UV-B radiation and molecular rates of evolution.

The present study tests the hypothesis that increased UV-B radiation exposure is associated with a higher rate of molecular evolution. I test this hypothesis by examining several planktonic animal groups under the assumption that pelagic organisms should

have an elevated exposure to solar radiation in comparison to their benthic counterparts. The results of my comparative approach strongly support the hypothesis, and I find some preliminary evidence that the increased mutagenesis may be linked to CPD formation.

Materials and Methods

Taxonomic sample and independent comparisons

I selected the largest possible number of phylogenetically independent comparisons (Felsenstein 1985; Harvey and Pagel 1991) to test for an association between molecular rate of evolution and a planktonic lifestyle in the animal kingdom. Each comparison involved two groups of taxa to be compared and a third 'outgroup' taxon that was closely related to this monophyletic pair. This comparative approach involves the selection of comparisons that do not overlap in evolutionary history, allowing each comparison to contribute one data point for statistical analysis. This approach eliminates the type I and type II errors that can occur when phylogenetic bias is ignored (Harvey and Pagel 1991). I maximized the relatedness between the two clades being compared, and between this pair and the outgroup, to increase the power and resolution of the analyses (Muse and Weir 1992; Moran 1996). To choose the independent comparisons for this study, I used well-supported phylogenies to select conservative monophyletic pairs (see Bromham et al. 1996; Bromham 2002).

I selected independent contrasts that were composed of a planktonic clade, a benthic sister clade and a closely-related outgroup. Some planktonic groups were not included in this study due to the lack of consensus on their phylogenetic placement.

Examples of planktonic taxa not analyzed include the arrow worms (Chaetognatha), copepods (Crustacea: Copepoda) and bristleworms (Annelida: Polychaeta). In total, eight independent comparisons of a planktonic versus benthic clade were identified (Table 1) spanning 4 phyla and 5 classes.

Sequence data

For both the planktonic and benthic clades within each independent comparison, I obtained sequence data from one or two genes, and sequences for between one and four species for each gene. I attempted to maximize the amount of sequence data being compared, but were often limited by either the availability of specimens or sequences. Each comparison was composed of an equal number of planktonic and benthic taxa to prevent an inaccurate measurement of substitution rate that can be caused by a node density effect (see Bromham et al. 2002). Furthermore, I attempted to both maximize taxonomic breadth for each clade and to use the same taxa for both genes when possible.

I selected one mitochondrial and one nuclear gene for my analysis: the mitochondrial protein-coding gene cytochrome oxidase subunit 1 (COI) and the nuclear small subunit ribosomal RNA gene (18S). The sequence data for the two genes were obtained from Genbank or determined by the following procedure. Genomic DNA was extracted using 50 μl of proteinase K extraction buffer according to the protocol of Schwenk et al. (1998) or using the GenElute™ Genomic DNA Miniprep Kit (Sigma) following the manufacturer's instructions (final product suspended in 30 μl of water). For amplification of the two genes, I used the primer pairs LCO1490/HCO2198 (Folmer *et al.* 1994) for an approximately 680 base pair (bp) fragment of COI and 1F/5R (Giribet et al. 1996) for ~950 bp of 18S. The 50 μl PCR reactions contained 0.5-2.0 μl of DNA

template, 4.5 µl 10X PCR buffer, 0.5 µl (10 µM) of each primer, 2.5 µl (50mM) $MgCl_2$, 0.25 µl (0.2 mM) dNTPs, and 0.2 µl *Taq* DNA polymerase. The PCR conditions for COI consisted of 3 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 40-50°C, and 1 min at 72°, followed by a final step of 5 min at 72°C. The PCR conditions for 18S consisted of an initial step at 95 °C for 5 min, 35 cycles of 45 s at 94 °C, 30 s at 40-50 °C, and 1 min at 72 °C, followed by a final step of 5 min at 72 °C. PCR products were sequenced directly or were purified using the Qiaex II gel extraction kit (Qiagen Inc.) following the manufacturer's instructions. Products were sequenced using an ABI 377 automated sequencer and ABI prism BigDye terminator 3 and 3.1 sequencing kits (Applied Biosystems).

The sequence data for each comparison were aligned separately for each gene. Sequences for the protein-coding gene COI were easily aligned by eye, guided by the translated amino acids; this was performed in the program Bioedit (Version 5.0.9) (Hall 1999). The ribosomal 18S gene was initially aligned using ClustalX (Version 1.81) (Thompson et al. 1997), and then manually adjusted in BioEdit. To eliminate ambiguous positions in the 18S alignments, I used Gblocks (Version 0.91b) (Castresana 2000) with program defaults.

Relative rates analysis

For each comparison, I estimated the rate of nucleotide substitution for the planktonic group and the benthic group. This was performed by employing relative rate tests (RRTs) (Sarich and Wilson 1973) on each alignment. RRTs have been shown to perform relatively poorly at diagnosing rate constancy within a phylogeny, or significant departures from it, (Bromham et al. 2000), but they are a straightforward and effective

method for comparing rate differences between *a priori*-defined lineages (Mooers and Harvey 1994; Bromham et al. 1996; Bromham 2002; Bromham and Cardillo 2003). RRTs were computed with PHYLTEST (Kumar 1996) which calculates the two-cluster RRT of Takezaki, Rzhetsky, and Nei (1995), providing an estimate for the rate of each of the two groups in substitutions per site since their divergence. I selected the Kimura 2-parameter (K2P) (Kimura 1980) distance metric for all analyses. In addition, I calculated the substitution rates while accounting for among-site rate heterogeneity since failing to do so may lead to misleading results (Slowinski and Arbogast 1999) The gamma parameter was calculated for each alignment using DAMBE (Version 4.2.8) (Xia and Xie 2001) and then entered as the value for the K2P + gamma option in PHYLTEST.

Nucleotide and dinucleotide frequency analysis

As in Lutzoni and Pagel (1997) and Hebert et al. (2002), I explored the base composition of nucleotides and dinucleotides for evidence of CPDs induced by UV-B radiation. I computed the proportion of thymine residues and thymine dimers within the planktonic clade versus the benthic/sedentary clade using DAMBE (Xia and Xie 2001).

Tests of association

I performed one-sided sign tests and one-sided Wilcoxon signed ranks tests using SYSTAT (Version 8.0) (SPSS Inc.) to test for associations of a planktonic lifestyle with molecular rate and nucleotide and dinucleotide frequencies. For all tests, I used a significance level of $\alpha = 0.05$.

Results

A significant association between a planktonic lifestyle and an accelerated rate of molecular evolution was observed for both the COI and 18S genes. For COI, all eight of the relative rate calculations revealed acceleration in the planktonic lineage (Table 2). Repeating the RRTs with K2P + gamma distances to account for among site rate heterogeneity produced nearly identical results (data not shown). Comparisons for the 18S gene were slightly less significant with 5 of 6 comparisons showing accelerated rates in the planktonic clade (Table 2), regardless of the distance measure used.

If CPD formation contributed to the higher substitution rate in planktonic lineages, one would expect a higher incidence of both thymine bases and thymine dimers since the predominant mutations arising from CPD formation are C→T and CC→TT events (Brash et al. 1991; Ziegler et al. 1993). There was no overwhelming trend in the comparisons, but all of the contrasts favour the hypothesis (Table 3). In particular, five of six 18S comparisons display a higher content of thymine residues in the planktonic clades and four of five display a higher thymine dimer concentration. Although not conclusive, these results suggest that CPD formation has contributed to the rate acceleration seen in the planktonic lineages.

Discussion

Rate acceleration and UV radiation

The present study provides the first comprehensive test of the hypothesis that increased UV-B radiation exposure is associated with a higher rate of molecular evolution. My results provide clear evidence that pelagic organisms in a diverse range of lineages have an elevated rate of nucleotide substitution in comparison to their

benthic relatives. This acceleration was observed in both mitochondrial and nuclear DNA, as well as in a protein-coding and a ribosomal gene. The general evidence for an increased substitution rate, across both genomes and varied gene types, is expected by an UV-induced increase in mutation rate. Furthermore, the association was even observed in the freshwater comparison (Branchiopoda), where one might expect the effects of UV-radiation to be attenuated by dissolved organic compounds (Xenopoulos and Schindler 2001).

The present study reinforces prior results suggesting that UV-B radiation is a contributing factor to the enhanced substitution rate apparent in some lineages. The biased nucleotide and dinucleotide compositions in the pelagic groups may be the signatures of cyclobutane pyrimidine excision and thus, provide supporting evidence of UV-induced mutagenesis (Brash et al. 1991; Ziegler et al. 1993). Although my results are dependent on the phylogenetic relationships implied by the comparisons, the association is consistent with previous studies (Smith et al. 1992; Hebert et al. 2002) including those investigating even more divergent groups (foraminiferans, Pawlowski et al. 1997; diatoms, Kooistra and Medlin 1996). The generality of this association between a planktonic lifestyle and an accelerated rate of molecular evolution is so broad that is very unlikely to be artifactual.

Alternative reasons for rate acceleration in plankton

The acceleration of molecular substitution rate in plankton might also reflect the higher metabolic rate of plankton compared to benthic taxa. Increased metabolic rate leads to an increased generation of oxygen radicals, known to damage DNA, providing a mechanistic basis by which a higher metabolism can accelerate rates of molecular

evolution (Martin and Palumbi 1993; Rand 1994). Unfortunately, there are no comparative data on the relative metabolic rate in planktonic versus benthic animals. However, I might infer that it is higher in most planktonic groups, since metabolic rate is negatively correlated with depth of occurrence in marine organisms (reviewed in Childress 1995). But, regardless of whether or not this inference is correct, there is very limited empirical support for the metabolic rate hypothesis. The few studies that have revealed a positive correlation between metabolic rate and substitution rate (e.g. Martin and Palumbi 1993; Bleiweiss 1998) have been questioned for their failure to account for phylogenetic bias or multiple substitutions (Slowinski and Arbogast 1999; Bromham 2002), and, in at least one case, reanalyzing the data appropriately resulted in no association (Slowinski and Arbogast 1999). Furthermore, studies that have adequately accounted for phylogenetic nonindependence and multiple substitutions have demonstrated that no correlation exists, at least for birds (Mooers and Harvey 1994), rodents (Rowe and Honeycutt 2002), and other mammals (Bromham et al. 1996). Therefore, currently available data do not support increased metabolic rate as a plausible cause for the observed rate acceleration in plankton.

In addition, the open water environment might select for certain other characteristics in its pelagic inhabitants, which could, in turn, influence rates of molecular evolution. In contrast to the benthos, the pelagic environment is dominated by abiotic-biotic interactions and stochastic environmental conditions (Rigby and Milsom 2000). This has fostered the long-term coexistence of species due to an unusually low amount of inter- and intra-specific competition (Ghilarov 1984; Rigby and Milsom 2000). Additionally, pelagic organisms, particularly those with long-lived

planktonic larvae, generally have greater dispersal capabilities and larger geographic ranges. But these characteristics should in theory decrease the rate of molecular evolution in these lineages. It is benthic taxa, with greater competition and smaller and sub-divided populations that should tend to form genetically isolated populations with increased localized selection, potentially leading to accelerated rates of evolution (Ohta 1976; DeSalle and Templeton 1988). For this reason, and with a lack of support for the metabolic rate hypothesis, exposure to UV-B radiation remains the best explanation for the observed rate acceleration in planktonic organisms.

Future directions

To advance this work and further test the hypothesis that UV-exposed organisms have an elevated rate of molecular elevation, two lines of future work would be useful. First of all, an expansion of both taxa and sequence data would be valuable. Once the phylogenetic affinities of numerous planktonic lineages within the Metazoa are resolved (e.g. chaetognaths, copepods, polychaetes), it will be possible to increase the number of comparisons. In addition, comparisons could be added from outside the Metazoa, including the reanalysis of groups that have previously been investigated (Pawlowski et al. 1997; Kooistra and Medlin 1996). The analysis of additional genes (e.g. mitochondrial ribosomal and/or nuclear protein-coding genes) would also be valuable, providing a further evaluation of the extent of the UV-induced mutagenesis. Secondly, it would be worthwhile to concentrate on one or more of the rate accelerated planktonic groups identified in this study and increase the taxonomic sampling of both benthic and planktonic taxa. A comparison of the patterning of interspecific and intraspecific genetic diversity would enable a more precise indication of the extent of rate variation.

Furthermore, it would allow reconstruction of the evolution of each nucleotide and dinucleotide site (as in Lutzoni and Pagel 1997), enabling verification that substitutions are occurring more rapidly at dipyrimidine sites, and that substitutions that eliminate dipyrimidine sites are favoured. Carefully designed studies following these suggestions will further elucidate the cause and generality of this association.

Table 1. Phylogenetically independent comparisons used in this study.

#	Taxonomic group	Comparison	Groundplan biology	Phylogeny /biology reference(s)
1	Cnidaria	Hydromedusae	Marine/planktonic	Bridge et al. 1995
		Anthozoa	Marine/sedentary	Collins 2002
		Diptera (COI)/ Ctenophora (18S)	Outgroup	
2	Opisthobranchia	Thecosomata/ Gymnosomata	Marine/planktonic	Dayrat et al. 2001
		Aplysiidae	Marine/benthic	
		Architectibranchia (COI)/ Cephalaspidea (18S)	Outgroup	
3	Branchiopoda	Daphniidae	FW/planktonic	Chapter 1
		Chydoridae/ Macrothricidae	FW/benthic	Taylor et al. 1999
		Cyclestheriidae	Outgroup	
4	Amphipoda	Hyperiidea	Marine/planktonic	Berge et al. 2003
		Gammaridea	Marine/benthic	
		Caprellidea	Outgroup	
5	Dendrobranchiata	Sergestoidea	Marine/planktonic	Schram 2001
		Penaeoidea	Marine/benthic	
		Astacidea	Outgroup	
6	Eucarida	Euphausiacea	Marine/planktonic	Schram 2001
		Decapoda	Marine/benthic	
		Cumacea	Outgroup	
7	Ostracoda	Myodocopa	Marine/planktonic	Cohen et al 1998
		Podocopa	Marine & FW/benthic	Martin & Davis 2001
		Cirripedia	Outgroup	
8	Urochordata	Appendicularia	Marine/planktonic	Wada 1998
		Ascidiacea	Marine/sedentary	Stach & Turbeville 2002
		Cephalochordata	Outgroup	

Table 2. Relative rate calculations for 8 independent comparisons using nucleotide data. The ratios of the planktonic branch length to the benthic branch length (ΔR) were calculated for the genes COI and 18S. Results of the one-tailed sign tests and Wilcoxon signed-ranks tests of the null hypothesis for each gene are given, where the null hypothesis states that increases in molecular rate are equally likely to be associated with a planktonic or benthic lifestyle.

#	Taxonomic group	ΔR_{COI}	ΔR_{18S}
1	Cnidaria	1.16	1.04
2	Opisthobranchia	1.94	2.07
3	Branchiopoda	1.02	1.36
4	Amphipoda	1.46	-
5	Dendrobranchiata	1.52	-
6	Eucarida	1.23	0.92
7	Ostracoda	1.83	1.10
8	Urochordata	1.46	1.85
Supportive comparisons		8 of 8	5 of 6
Sign test		$p = 0.004$	$p = 0.110$
Wilcoxon test		$p = 0.006$	$p = 0.038$

Table 3. Nucleotide and dinucleotide frequency data for the COI and 18S comparisons. Thymine content (T) and thymine dinucleotide content (TT) was determined for the planktonic and benthic clade of each comparison for both genes. Results of the one-tailed sign test and Wilcoxon signed-ranks test of the null hypothesis are given, where the null hypothesis states that biases in T or TT content are equally likely to be associated with a planktonic or benthic lifestyle.

Taxonomic group	Contrast	COI		18S	
		T	TT	T	TT
1- Cnidaria	Hydromedusae	0.362	0.133	0.274	0.086
	Anthozoa	0.381	0.163	0.271	0.086
2- Opisthobranchia	Thecosomata/Gymnosomata	0.395	0.156	0.257	0.071
	Aplysiidae	0.411	0.179	0.250	0.068
3- Branchiopoda	Daphniidae	0.342	0.149	0.228	0.056
	Chydoridae/Macrothricidae	0.338	0.134	0.232	0.058
4- Amphipoda	Hyperiidea	0.433	0.204		
	Gammaridea	0.351	0.137		
5- Dendrobranchiata	Sergestoidea	0.380	0.150		
	Penaeoidea	0.353	0.137		
6- Eucarida	Euphausiacea	0.358	0.137	0.244	0.071
	Decapoda	0.337	0.130	0.241	0.069
7- Ostracoda	Myodocopa	0.344	0.134	0.252	0.070
	Podocopa	0.338	0.131	0.243	0.065
8- Urochordata	Appendicularia	0.607	0.150	0.270	0.089
	Ascidiacea	0.427	0.200	0.246	0.077
Supportive comparisons =		6 of 8	5 of 8	5 of 6	4 of 5
Sign test one-sided: p =		0.144	0.364	0.110	0.188
Wilcoxon one-sided: p =		0.062	0.5	0.058	0.04

CHAPTER 3

Fallout from an arms race: Accelerated and miniaturized genomes in metazoan parasites

Abstract

The transition to a symbiotic habit has been linked to accelerated rates of molecular evolution in numerous taxonomic lineages. This association has been shown for mutualistic symbionts, but it is unclear if it also exists for antagonistic symbionts, or parasites. I tested whether parasitic lineages have an increased rate of molecular evolution in comparison with their free-living relatives by performing a comprehensive comparative analysis of 21 phylogenetically independent contrasts. The results clearly demonstrated a generalized genomic acceleration due to the transition to a parasitic lifestyle, for both mitochondrial and nuclear DNA, and in both protein-coding and ribosomal genes. By exploring the patterns of sequence evolution, with reference to previous work, I inferred that the enhanced substitution rate is most likely a consequence of a genetic arms race between host and parasite. In addition to accelerated sequence evolution, this study provided evidence that parasites are experiencing other genome-wide changes, such as a reduction in DNA content.

Introduction

It has long been recognized, and is now generally accepted, that rates of molecular evolution vary widely across taxonomic lineages (e.g. Britten 1986; Vawter and Brown 1986; Li and Tanimura 1987; reviewed in Mindell and Thacker 1996; Li 1997). It is also well established that a variety of organismal characteristics – ecological, physiological, and genomic – are correlated with incidents of rate variation. Hence, an understanding of the interplay between several variables is required to explain the differences in rates of molecular evolution among lineages. The most significant and general traits identified, each with a sound mechanistic basis, are generation time (Wu and Li 1985, Li et al. 1996), population size (Ohta 1972), DNA repair efficiency (Britten 1986), and metabolic rate (Martin and Palumbi 1993; Rand 1994). Unfortunately, the comparative data necessary to determine the relative contributions of these variables to rate heterogeneity in species lineages is sparse for most taxonomic groups (e.g. invertebrates) (Bromham et al. 1996).

Deeper insights into the factors modulating rates of evolution can be gained by systematic comparisons between allied lineages differing in key biological attributes. One such pattern is the transition to a symbiotic existence. The adoption of this lifestyle has occurred repeatedly and has been linked to elevated rates of substitution in symbiotic organisms relative to their free-living counterparts. Most past work has concentrated on mutualistic symbionts, largely bacteria. Numerous proteobacterial mutualists (e.g. *Buchnera*) (Moran 1996; Itoh et al. 2002; Woolfit and Bromham 2003), other eubacterial endosymbionts (Woolfit and Bromham 2003), and a protist mitochondrion (Itoh et al. 2002) all clearly exhibit accelerated molecular rates. Moreover, this

association also seems to hold in eukaryotes, where basidiomycete (Lutzoni and Pagel 1997) and ascomycete fungi (Woolfit and Bromham 2003) and aphid 'inquilines' (Miller and Crespi 2003) provide further examples of increased substitution rates when mutualists are contrasted with their free-living sister groups. The proposed reasons for this association include Muller's ratchet (Moran 1996), the fixation of slightly deleterious mutations in smaller effective populations (Lambert and Moran 1998; Woolfit and Bromham 2003; Miller and Crespi 2003), and/or an enhanced mutation rate (Itoh et al. 2002).

There is also evidence to suggest that antagonistic symbionts (i.e. pathogens and parasites) show accelerated rates of molecular evolution. Parasitic bacteria *Wolbachia* (Moran 1996), parasitic plants (Nickrent and Starr 1994), lice (Hafner et al. 1994; Page et al. 1998; Johnson et al. 2003; Yoshizawa and Johnson 2003), parasitic wasps (Dowton and Austin 1995; Castro et al. 2002) and at least one parasitic fly lineage (Castro et al. 2002) display elevated rates of substitution. Once again, Muller's ratchet (Moran 1996) and slightly deleterious mutation theory have been proposed as an explanation for the rate acceleration (Dowton and Austin 1995; Page et al. 1998; Castro et al. 2002). However, this trend could also be explained by a "genetic arms race" (Dowton and Austin 1995; Page et al. 1998). The intimate and conflicting relationship between parasite and host may engender an acceleration of adaptive evolution, resulting from either the Red Queen process (Van Valen 1973), or through selection for an increased rate of mutation (Haraguchi and Sasaki 1996). Since these mechanisms are not mutually exclusive, they may jointly contribute to the likelihood of a general link between molecular rate acceleration and parasitic lifestyles.

The present study tests the generality of the association between accelerated rates of molecular evolution and a parasitic lifestyle by exploring multiple independent transitions to parasitism in the Metazoa. I perform a comparative analysis using phylogenetically independent contrasts of parasitic taxa versus free-living sister taxa. In addition, I explore other correlates of the transition to a parasitic lifestyle, including base composition bias and genome size variation. Previous work has identified a higher A+T content in the nuclear and/or mitochondrial genomes of numerous parasites, including endosymbiotic microorganisms (Rocha and Danchin 2002; Woolfit and Bromham 2003), hymenopterans (Dowton and Austin 1995), and nematodes (Hugall et al. 1997). Similarly, complete genome sequencing projects have revealed that genome reduction is apparent in many bacterial pathogens and endosymbionts (Moran 2002; Wernegreen 2002), the only known parasitic archaeon (Waters et al. 2003), and in parasitic microsporidians (e.g. Katinka et al. 2001). Testing these hypotheses may provide insight into the factors affecting rate variation across lineages, the evolution of mutualistic lifestyles, and evolution of the genome in general.

Materials & Methods

Taxonomic sample

The taxonomic sample was designed to enable a comparative analysis by the phylogenetically independent contrasts method (Felsenstein 1985; Harvey and Pagel 1991). This procedure selects a maximum number of independent comparisons such that no lineage is represented more than once, allowing each comparison to contribute one data point for statistical analysis. This method entails selecting two groups of taxa

to be compared, and a third 'outgroup' taxon that is more distantly related. To maximize the power and resolution of the analyses, the two clades being compared should be closely related to each other, as well as to the outgroup (Muse and Weir 1992; Moran 1996). To choose the independent comparisons for this study, I used published phylogenies and selected conservative and monophyletic pairs (see Bromham et al. 1996; Bromham 2002).

I selected independent contrasts that comprised a parasitic clade, a free-living sister clade and a free-living outgroup. I defined a parasite as an organism that both spends a segment of its lifecyle in or on its host (Poulin and Morand 2000), and decreases the fitness of its host. I treat a clade as parasitic if the majority of its constituents are parasitic. I was unable to incorporate certain parasitic groups into phylogenetically independent comparisons for two reasons. Firstly, there exists little consensus on the phylogenetic placement of several parasitic taxa (Zrzavy 2001). Taxa not included for this reason include the twisted-wing parasites (Strepsiptera), myzostomids (Myzostomida), tongue worms (Pentastomida), and the numerous parasitic lineages within the ticks and mites (Acari). Secondly, other taxa that have a well supported position in the metazoan tree lack a free-living sister clade for comparison. Examples of this case include horsehair worms (Nematomorpha), and fleas (Siphonaptera), sister taxa to the nematodes (Nematoda) and flies (Diptera), respectively. Nematoda and Diptera each contain several parasitic lineages, a majority of which are included in my analyses, precluding their use as sister taxa.

Sequence data

For each independent comparison, I obtained sequence data from between one and four genes, and for each gene, I employed between one and four sequences for both the parasitic and free-living clades. In all instances, I attempted to maximize the amount of sequence data compared, but were limited by the availability of specimens and sequences in online databases. Each comparison was composed of an equal number of parasitic and free-living taxa to prevent overestimation of the molecular rate in lineages with higher representation, which is linked to the node density effect (see Bromham et al. 2002). In addition to this, I attempted to maximize taxonomic breadth for each clade and to use the same taxa for multiple genes when possible.

The four loci chosen were the mitochondrial protein-coding gene *cytochrome oxidase subunit 1* (COI), the mitochondrial ribosomal gene 16S, the nuclear protein-coding gene *elongation factor 1 alpha* (EF-1a), and the nuclear ribosomal gene 18S. Sequences for the four genes were largely obtained from Genbank or from ongoing sequencing projects (see below). However, additional sequences were determined by the following procedure. Genomic DNA was extracted using 50 μl of proteinase K extraction buffer according to the protocol of Schwenk et al. (1998) or using the GenElute™ Genomic DNA Miniprep Kit (Sigma) following the manufacturer's instructions (final product suspended in 30 μl of water). The primer pairs for amplification of the three loci were as follows: LCO1490/HCO2198 (Folmer *et al.* 1994) for a 680 base pair (bp) fragment of COI; 16Sar/16Sbr (Palumbi *et al.* 1991) for a 570 bp fragment of 16S; and the primers 1F and 5R (Giribet et al. 1996) for approximately 950 bp of 18S. The 50μl PCR reactions contained 0.5-2.0 μl of DNA template, 4.5 μl

10x PCR buffer, 0.5 ul (10 μM) of each primer, 2.5 ul (50mM) $MgCl_2$, 0.25 ul (0.2 mM) dNTPs, and 0.2 ul *Taq* DNA polymerase. The PCR conditions for COI and 16S consisted of 1.5 min at 94°C, followed by 35 cycles of 45 s at 93°C, 1 min at 50°C, and 1 min at 72°, followed by one cycle of 5 min at 72°C. The PCR conditions for 18S consisted of 1 cycle at 94 °C, 35 cycles of 30 s at 93 °C, 30 s at 50 °C, and 3 min at 72 °C, followed by one cycle of 5 min at 72 °C. PCR products were excised from agarose gels and purified using the Qiaex II gel extraction kit (Qiagen Inc.), or were sequenced directly subsequent to amplification. Products were sequenced using an ABI 377 automated sequencer and ABI prism BigDye terminator 3 and 3.1 sequencing kits (Applied Biosystems).

Additional COI sequence data were obtained from ongoing nematode complete genome sequencing projects (see Parkinson et al. 2003). Contigs, or clusters of contigs, primarily derived from expressed sequence tags (ESTs) were queried for the COI gene. For *Caenorhabditis briggsae*, I blasted the COI sequence of *C. elegans* against the *C. briggsae* genomic DNA database on WormBase (http://www.wormbase.org). Two large contigs with significantly close identity to the *C. elegans* complete genome were selected and the COI gene (identical in the two contigs) was removed. I obtained the COI gene for 3 further species by querying two databases for clusters putatively identified as COI: *Ancylostoma caninum* from NemaGene (http://www.nematode.net) and *Haemonchus contortus* and *Pristionchus pacificus* from NemBase (http://www.nematodes.org). In all instances, the sequences were taken from highly redundant contigs or clusters which should remove the possibility they contain any sequencing errors which are inherent in the generation of ESTs (Parkinson et al. 2003).

The sequences for each sister taxon comparison, and for each gene, were aligned separately. COI and EF-1a sequences were aligned by eye in Bioedit (Version 5.0.9) (Hall 1999), whereas 16S and 18S were aligned using ClustalX (Version 1.81) (Thompson et al. 1997) and manually adjusted, when necessary, in BioEdit. Gblocks (Version 0.91b) (Castresana 2000) was then used to eliminate ambiguous positions in the ribosomal alignments.

Relative rates analysis

For each comparison, I estimated the rate of nucleotide substitution in the parasitic and free-living clades. This was performed by employing relative rate tests (RRTs) (Sarich and Wilson 1973) on each alignment. RRTs perform relatively poorly at diagnosing rate constancy within a phylogeny, as well as significant departures from it (Bromham et al. 2000), but they provide a straightforward and effective method for comparisons of rate differences between *a priori*-defined lineages (Mooers and Harvey 1994; Bromham et al. 1996; Bromham 2002; Bromham and Cardillo 2003). RRTs were computed with PHYLTEST (Kumar 1996) which calculates the two-cluster RRT of Takezaki et al. (1995), providing an estimate for the rate in each of the two groups in substitutions per site since their divergence from a common ancestor. I selected the Kimura 2-parameter (K2P) (Kimura 1980) distance metric for all analyses. In addition, I followed the advice of Slowinski and Arbogast (1999) and calculated the relative rates while accounting for among-site rate heterogeneity. This was accomplished by first calculating the gamma parameter for each alignment using DAMBE (Version 4.2.8) (Xia and Xie 2001) and then selecting the K2P + gamma option in PHYLTEST. For the two protein-coding genes COI and EF-1a, I also performed RRTs on the translated

amino acids and nonsynonymous sites. I performed the amino acid RRTs in PHYLTEST using a Poisson correction for multiple hits and computed the d_N (number of non synonymous substitutions per non synonymous site) RRTs in RRTree (Robinson-Rechavi and Huchon 2000) following the method of Robinson et al. (1998).

Base composition

To explore the rate variation in more detail, and to determine whether metazoan parasites share the biased A+T content of other symbionts (Moran 1996; Rocha and Danchin 2002; Woolfit and Bromham 2003), I compared the base composition in each clade. I tabulated the A+T content and A only content for all four genes using the program DAMBE.

Genome size variation

I also tested whether metazoan parasites possess smaller genome sizes than their free-living relatives since a trend towards genome reduction is evident in endosymbiotic, pathogenic, and parasitic microbes (e.g Moran 1996; Wernegreen 1996; Katinka et al. 2001; Waters et al. 2003). I used the Animal Genome Size Database (Gregory 2001) to collate data for haploid genome sizes (i.e. C-value), measured in picograms/cell, for 10 independent comparisons. In cases where more than one measurement existed for a species, I calculated the species average before calculating the clade mean.

Tests of association

I performed one-sided sign tests and one-sided Wilcoxon signed ranks tests using SYSTAT (Version 8.0) (SPSS Inc.) to test for associations of a parasitic lifestyle with molecular divergence rate, base composition, and genome size. For the molecular rate and genome size comparisons, I compared the ratios of branch lengths and C-values,

respectively. For the base composition comparisons, I compared the difference in A+T and A only contents. For all tests, I used a significance level of $\alpha = 0.05$.

Results

A total of 21 phylogenetically independent comparisons spanning 6 metazoan phyla were identified (Table 1). For the nucleotide sequence data, a significant association between parasitic lifestyle and rate of molecular evolution was revealed for three of the four genes (COI, 16S and 18S) (Table 2). The comparisons for the EF-1a gene were not significant, although 4 of 5 showed accelerated rates in the parasitic lineage. Repeating the relative rate tests while accounting for among site rate heterogeneity produced near identical results (data not shown). The results were similar for the relative rate tests using amino acid and nonsynonymous sequence data for the two protein-coding genes, COI and EF-1a. For the COI gene, 13 of 14 amino acid comparisons and 13 of 13 d_N comparisons displayed an increased substitution rate in the parasitic lineage. The EF-1a gene again did not show a significant association, but 3 of 5 and 4 of 5 comparisons, for amino acid and d_N data, respectively, showed a rate acceleration following a transition to a parasitic lifestyle.

The comparisons of base compositions in parasitic and free-living clades were inconclusive (Table 4). There was no association between A+T bias and a parasitic lifestyle, in contrast to the pattern seen in mutualistic symbionts (e.g. Rocha and Danchin 2002; Woolfit and Bromham 2003). If my analysis failed to detect a difference, possibly due to the low power of the tests or the narrow range of base composition values, there might actually be a G+C bias, since a majority of the

comparisons for three of the four genes possessed a higher G+C content in the parasitic clade relative to the free-living clade. There was also no association of adenine content and a parasitic existence, which has been observed in parasitic wasps (Dowton and Austin 1997) and might be expected if oxidative damage is responsible for the enhanced substitution rate in parasites (Wagner et al. 1992; Martin 1995).

Ten comparisons to test the association between small genome size and a parasitic lifestyle were identified (Table 5). Eight of the ten comparisons showed a reduced genome size in the parasitic lineage relative to the free-living lineage (one-tailed Wilcoxon signed-rank test, $Z = 1.58$; $p = 0.055$). The two comparisons that did not reveal the association (Copepoda: Siphonostomatoida vs. Harpacticoida and Isopoda: Cymothoidae vs. Sphaeromatidae) were based on the smallest number of measurements.

Discussion

Parasitism and rate acceleration

The present study convincingly demonstrates an association between accelerated rates of molecular evolution and a parasitic lifestyle. My analysis of 21 phylogenetically independent comparisons over a broad range of metazoan taxa revealed a significantly increased rate of sequence evolution in parasitic lineages relative to their free-living counterparts. This increased substitution rate was observed for both mitochondrial and nuclear DNA, and in both protein-coding and ribosomal genes. This study is not the first to demonstrate an acceleration of rates in metazoan parasites, but it is the first to show a generalized genomic acceleration correlated with the transition to a parasitic lifestyle.

The accelerated rates of molecular evolution in parasites could be due to Muller's ratchet, deleterious mutation accumulation in small populations, relaxed selective constraints, or to a genetic arms race. The results of my study, when coupled with earlier work, allow exclusion of the former three possibilities in favour of the latter.

Muller's ratchet (Muller 1964; Maynard Smith 1978) predicts elevated rates of molecular evolution in asexual lineages and has been proposed as the cause of rate acceleration in several proteobacterial endosymbionts such as *Buchnera* (Moran 1996). However, Muller's ratchet can be excluded as a contributing factor in the present study since the bulk of my parasitic lineages were sexual. Furthermore, there is no convincing evidence that transitions to asexuality always result in a higher substitution rate. Normark et al. (2003) suggest a number of situations that may actually lead to a decrease in substitution rate, and there are two previous examples of ancient asexuals that do not possess a higher rate of molecular evolution than related sexual lineages – the bdelloid rotifers (Mark Welch and Meselson 2001) and darwinulid ostracods (Schon and Martens 1998).

Increased rates of molecular evolution can be due to the increased rates of fixation of weakly deleterious mutations in small populations. This mechanism has been linked to the transition to a symbiotic existence, both in mutualists (Lambert and Moran 1998; Woolfit and Bromham 2003; Miller and Crespi 2003) and parasites (Dowton and Austin 1995; Page et al. 1998; Castro et al. 2002). Depending on their transmission dynamics and the intimacy of their relationship with the host, symbionts may experience frequent bottlenecks and strong population subdivision, leading to small effective population sizes. While there is compelling evidence for this mechanism in

endosymbionts (Lambert and Moran 1998; Woolfit and Bromham 2003), I find little evidence to suggest its importance in metazoan parasites. Firstly, there are no comparative data on the relative effective population sizes in parasites versus free-living animals. However, I might infer that no general association exists due to the diversity of host-parasite relationships and the variation in parasite reliance on host resources. Furthermore, it takes little gene flow to disrupt demic structure (Wright 1951), and ectoparasitic lineages rarely have mechanisms that foster strong imbreeding. Secondly, organisms with small effective population sizes typically exhibit high A+T bias, a reflection of drift outweighing translational selection (e.g. Clark et al. 1999; Woolfit and Bromham 2003). Yet my analysis revealed no evidence of A+T bias in parasitic lineages; if anything, they favoured the metabolically more expensive guanine and cytosine (Rocha and Danchin 2002).

Increased rates of sequence evolution could also be explained by a relaxation of selection in parasites versus their free-living allies. However, this would imply that the environments encountered by parasites are more stable than those encountered by free-living taxa. Given the relative hostility of certain aspects of the parasite's environment (e.g. host immune response) this is unlikely. If relaxation of selection occurs at all, it would probably be gene-specific following a lifestyle transition (Moran 1996). Secondly, relaxed selection should lead to increased A+T ratios as a consequence of weaker selection (Gojobori et al. 1982; Hu and Thilly 1994; Lutzoni and Pagel 1997). Yet there was no evidence of A+T bias in the parasitic taxa in comparison to their free-living relatives. These two observations suggest that relaxation of selection is unlikely to be a significant contributing factor to the rate acceleration in parasites.

A final cause for an association between accelerated rates and parasitism is a genetic arms race between host and parasite (Dowton and Austin 1995; Page et al. 1998). Because parasite and host are in endless conflict, an antagonistic coevolution should result from species continuously engaging in adaptation and counteradaptation (Ehrlich and Raven 1964; Van Valen 1973). This coevolutionary race may produce an elevated rate of molecular evolution in two ways.

Firstly, parasites should experience selective regimes with an increased strength of positive selection in comparison to those of their free-living cousins, increasing the rate of fixation of mutations. Prior studies have demonstrated positive selection in genes involved in host-parasite interaction (e.g surface proteins) (Endo et al. 1996; Haydon et al. 2001; Urwin et al. 2002; Jiggins et al. 2002), as well as increased positive selection in parasitic symbionts relative to mutualists (Jiggins et al. 2002). While genes involved in evading host immune responses are obvious targets for selection, other genes in parasitic lineages may also be under particularly intense selection. For example, hosts commonly shower parasites with reactive compounds such as oxygen radicals and nitric oxide (Nappi et al. 1995, 2000). Some of these reactive species damage parasite cell membranes, proteins, and nucleic acids (Nappi et al. 2000), but nitric oxide competes with oxygen for binding sites of cytochrome oxidase, the terminal enzyme in the electron transport chain. As such, it can have a potent inhibitory effect on cellular respiration (Brunori 2001). The COI gene used in this study may well be under strong positive selection. Aside from genes exposed to selection, mutations in neutral genes may be fixed due to their linkage with genes under selection. Large genomic regions, and certainly entire mitochondrial genomes, may be driven to fixation due to selective

sweeps motivated by selection on a single gene. This may be of significant importance considering the mounting evidence to suggest an intimate relationship between mitochondrial and nuclear genomes (Rand 2001; Sackton et al. 2003; Hebert et al. 2003). Therefore, my results of a generalized genomic acceleration in parasites, particularly at amino acid and non synonymous sites, are consistent with an arms race-related increase of positive selection.

A genetic arms race might also cause rate acceleration by elevating the mutation rate, which may work in tandem with positive selection. Theoretical work has suggested that antagonistic coevolution should select for an increase in the mutation rate in both host and parasite (Nee 1989; Haraguchi and Sasaki 1996). This response might be accomplished in various ways, including the fixation of mutations that compromise the efficacy of the DNA repair and replication machinery, or the loss of genes encoding them (Koonin et al 1996). In both instances, these alterations would make the machinery more error-prone (Bromham and Penny 2003). This mechanism is supported by studies of vertebrate immune systems that actually raise their mutation rate in response to parasite attack (e.g. Sale et al. 2001) and parasitic bacteria that display accelerated substitution rates in the absence of DNA repair genes (Koonin et al. 1996). Furthermore, a 'mutator' allele that enhances mutation rate can be selected for in laboratory strains of *E. coli* (Chao and Cox 1983), and the frequency of these 'mutator' clones is high in natural isolates of pathogenic bacteria (LeClerc et al. 1996), suggesting that mutation rate can be shaped by selection (Bromham and Penny 2003). This mechanism of selection for mutator genes is consistent with the generalized genomic acceleration that I observed in parasitic lineages. To thoroughly test this, it will be

necessary to compare the synonymous (d_S) substitutions in parasitic versus free-living taxa. Unfortunately, I was unable to compute the d_S in almost all comparisons due to saturation, a common problem in such studies (e.g. Itoh et al. 2002). Despite this, my results clearly exclude alternative hypotheses and favour a genetic arms race as the likely cause of the observed association between accelerated sequence evolution and a parasitic lifestyle.

Parasitism and genome reduction

The present study also provides the first comprehensive test of an association between genome reduction and a parasitic lifestyle. My results demonstrate a trend towards smaller genome size in parasitic lineages, a pattern seen previously in microbial mutualists and parasites (reviewed in Moran 2002; Wernegreen 2002). Two factors likely underpin genome shrinkage following a transition to parasitism. Firstly, simplification is a common theme in parasites, evident by varying degrees of reduced morphological organization, and may apply to genomic organization as well. Parasites are reliant on their hosts for resources, also to varying degrees, and may reduce their gene repertoire to a minimal set able to simply complete core processes such as growth and replication (Koonin et al. 1996; Moran 2002). Genes that are unused become inactivated by fixed mutations, and inactivated genes can subsequently be excised from the genome through deletional bias (Mira et al. 2001). This process, whereby mutation is biased towards deletion instead of insertion, may remove genes that are not effectively maintained by selection, as well as non-coding portions of the genome. Selection in favour of smaller genomes may be the second mechanism contributing to genome reduction. Genome shrinkage may be advantageous, most notably, to facilitate a faster

replication rate, which potentially leads to shorter generation times and more rapid development. Although genome size is undoubtedly modulated by a variety of factors (Gregory and Hebert 1999; Gregory 2001b), a 'race for replication' may have been of particular importance in the evolution of parasite genomes. Division rate may have been under intense selection, which in turn, may have been mediated by host-parasite genetic arms races.

Interestingly, there is anecdotal evidence to suggest that genome shrinkage may be occurring to the mitochondrial genome as well. *Plasmodium*, a blood parasite and the causative agent of malaria, has the smallest known mtDNA (6 kbp) and the least number of genes (5) (Feagin 1994; Burger et al. 2003b). Also, the ten smallest full mitochondrial genomes sequenced from metazoans (375 in total) are possessed by parasitic taxa (GenBank release 138). This pattern suggests an adaptive decrease in mitochondrial genome size has also occurred, since selection for smaller size has been demonstrated to be important in the evolution of animal mtDNA (see Rand 1993). It is certain that an investigation of both nuclear and mitochondrial genome sizes in parasitic taxa would be an intriguing avenue for future research.

Parasitism and other distinctive genetic features

Parasites may well experience other genome-wide changes in addition to accelerated sequence evolution and reduction in DNA content. Studies exploring the organization of mtDNA over a broad spectrum of organisms point towards, at least superficially, a pattern of peculiar features in parasitic lineages (Burger et al. 2003b). For example, *Dicyema*, a mesozoan endoparasite of cephalopods (Watanabe et al. 1999), trypanosomes, protozoan parasites of various plants and animals, (Lukes at el. 2002),

and *Amoebidium parasiticum*, an ectosymbiont of freshwater arthropods, but member of the mostly parasitic Ichthyosporea (Burger et al. 2003a), all possess multipartite mitochondrial genomes consisting of two or more distinct DNA molecules. This pattern extends into metazoan parasites, as the parasitic nematode *Globodera* is unique among metazoans in having a multipartite mtDNA structure (Armstrong et al. 2000). Gene rearrangements also appear to be occurring more readily in parasites. Dowton and Campbell (2001) noted that parasitic lineages appear at first glance to be more prone to mtDNA rearrangement events (e.g. Black and Roehrdanz 1998; Campbell and Barker 1998, 1999; Le at el. 2000; Shao et al. 2001). Since nucleotide substitution rates are correlated with mitochondrial gene rearrangement rates (Shao et al. 2003), this might provide the mechanism for their cause. Increased substitution rates may cause more errors in replication through mutations at initiation and termination sites, or by producing more double strand breaks, facilitating an increase in gene rearrangements (Boore 2000; Lavrov et al. 2002; Shao et al. 2003). Although unusual mtDNA features (and gene rearrangements) are not limited to parasitic taxa (Burger et al. 2003b), their high occurrence does warrant further study, particularly in light of other genomic changes identified in the present study.

Table 1. Phylogenetically independent comparisons of parasitic and free-living taxa used in the relative rate analysis. References used for the selection of comparisons are listed.

Phylum	#	Taxonomic group	Comparison	Biology*	Reference(s)
Cnidaria	1	Cnidaria	Narcomedusae	EC/EN	Collins 2002
			Rhopalonematidae	FL	
			Limnomedusae	OG	
Nematoda	2	Nematoda clade 1	Trichocephalida	EN	Blaxter et al. 1998
			Mononchida	FL	Dorris et al. 1999
			Rhabditoidea	OG	
	3	Nematoda clade 2	Triplonchida	EN	Blaxter et al. 1998
			Enoplida	FL	Dorris et al. 1999
			Rhabditoidea	OG	
	4	Nematoda clade 4a	Strongyloididae	EN	Blaxter et al. 1998
			Panagrolaimidae	FL	Dorris et al. 1999
			Rhabditoidea	OG	
	5	Nematoda clade 4b	Tylenchida	EN	Blaxter et al. 1998
			Cephaloboidea	FL	Dorris et al. 1999
			Rhabditoidea	OG	
	6	Nematoda clade 5	Strongylida	EN	Blaxter et al. 1998
			Rhabditoidea	FL	Dorris et al. 1999
			Diptera (COI)/ Enoplida (16S)/ Cephaloboidea (18S)	OG	
Syndermata	7	Syndermata	Acanthocephala	EN	Garey et al. 1998
			Rotifera	FL	Herlyn et al. 2003
			Platyhelminthes	OG	
Arthropoda	8	Cirripedia	Rhizocephala	EN	Martin & Davis 2001
			Thoracica	FL	
			Copepoda	OG	
	9	Maxillopoda	Branchiura	EC	Brusca & Brusca 1990
			Ostracoda	FL	Spears & Abele 1997
			Copepoda	OG	
	10	Copepoda	Siphonostomatoida/ Poecilostomatoida	EC	Huys & Boxshall 1991
			Harpacticoida	FL	
			Calanoida	OG	

Phylum	#	Taxonomic group	Comparison	Biology*	Reference(s)
	11	Isopoda	Bopyridae/ Cymothoidae Sphaeromatidae Gammaridae	EC/EN FL OG	Brusca & Wilson 1991 Dreyer & Wagele 2001
	12	Diptera 1	Oestroidea Muscoidea Ephydroidea	EN FL OG	McAlpine 1989 Wiegmann et al. 1993
	13	Diptera 2	Nycteribiidae Glossinidae Drosophilidae	EC FL OG	McAlpine 1989 Wiegmann et al. 1993
	14	Diptera 3	Pipunculidae Syrphidae Drosophilidae	EN FL OG	McAlpine 1989 Wiegmann et al. 1993
	15	Diptera 4	Conopoidea Tephritoidea Ephydroidea	EC/EN FL OG	McAlpine 1989 Wiegmann et al. 1993
	16	Diptera 5	Bombyliidae Therevidae Drosophilidae	EC FL OG	McAlpine 1989 Wiegmann et al. 1993
	17	Hymenoptera	Apocrita Tenthrediniodea Diptera	EN FL OG	Ronquist et al. 1999 Dowton & Austin 2001
	18	Psocodea	Phthiraptera Psocoptera Hemiptera	EC FL OG	Yoshizawa & Johnson 2003
	19	Platyhelminthes	Neodermata Turbellaria Polychaeta	EN FL OG	Littlewood et al. 1999 Lockyer et al. 2003
Annelida	20	Annelida	Euhirudinea Lumbricidae/Enchytraeidae Polychaeta	EC FL OG	Martin 2001 Siddall et al. 2001
Mollusca	21	Bivalvia	Unionoida Trigonioida Veneroida	EN FL OG	Adamkewicz et al. 1997 Hoeh et al. 1998

* EC: ectoparasitic; EN: endoparasitic; FL: free-living; OG: outgroup

Table 2. Relative rate calculations for 21 independent comparisons using nucleotide data. The ratios of the parasitic branch length to the free-living branch length (ΔR) were calculated for the four genes COI, 16S, EF-1a, and 18S. Results of the one-tailed sign tests and Wilcoxon signed-ranks tests of the null hypothesis for each gene are given, where the null hypothesis states that increases in molecular rate are equally likely to be associated with a parasitic or free-living lifestyle.

#	Taxonomic group	ΔR_{COI}	ΔR_{16S}	$\Delta R_{EF\text{-}1a}$	ΔR_{18S}
1	Cnidaria				5.12
2	Nematoda clade 1				1.45
3	Nematoda clade 2				1.10
4	Nematoda clade 4a				1.12
5	Nematoda clade 4b				0.70
6	Nematoda clade 5	1.02	0.60		0.76
7	Syndermata	1.33	0.79		2.16
8	Cirripedia	1.81	1.41		1.12
9	Maxillopoda	1.22			0.75
10	Copepoda	1.02	1.57		0.69
11	Isopoda	1.12	1.02		1.75
12	Diptera 1	1.19	1.39	1.58	1.49
13	Diptera 2		0.91		2.11
14	Diptera 3	1.32	5.17		
15	Diptera 4		1.02		
16	Diptera 5	0.83	1.18	0.72	1.33
17	Hymenoptera	1.21	1.08		
18	Psocodea	1.74	1.94	1.09	1.44
19	Platyhelminthes	2.33	1.08	1.08	1.76
20	Annelida	2.04	1.40	1.40	2.02
21	Bivalvia	1.26			1.04
Supportive comparisons		13 of 14	11 of 14	4 of 5	14 of 18
Sign test		$p = 0.001$	$p = 0.029$	$p = 0.188$	$p = 0.016$
Wilcoxon test		$p = 0.002$	$p = 0.037$	$p = 0.112$	$p = 0.008$

Table 3. Relative rate calculations for 14 independent comparisons using amino acid and codon sequence data. The ratios of the parasitic branch length to the free-living branch length were calculated for amino acid substitutions (ΔR_{AA}) and nonsynonymous substitutions (ΔdN) on the two protein-coding genes COI and EF-1a. Results of the one-tailed sign tests and Wilcoxon signed-ranks tests of the null hypothesis for each gene are given, where the null hypothesis states that increases in molecular rate are equally likely to be associated with a parasitic or free-living lifestyle.

#	Taxonomic group	$\Delta R_{AA\ COI}$	ΔdN_{COI}	$\Delta R_{AA\ EF\text{-}1a}$	$\Delta dN_{EF\text{-}1a}$
6	Nematoda clade 5	0.92	1.03		
7	Syndermata	2.36	1.41		
8	Cirripedia	3.27	1.92		
9	Maxillopoda	1.83	1.32		
10	Copepoda	1.19	1.00		
11	Isopoda	1.18	1.16		
12	Diptera 1	1.53	1.22	0.92	1.18
14	Diptera 3	1.78	1.45		
16	Diptera 5	2.00	2.20	0.67	0.74
17	Hymenoptera	1.71	1.46		
18	Psocodea	2.19	1.58	16.72	1.26
19	Platyhelminthes	2.97	1.72	1.22	1.05
20	Annelida	2.90	2.52	1.66	1.16
21	Bivalvia	2.23	1.06		
Supportive comparisons		13 of 14	13 of 13	3 of 5	4 of 5
Sign test		$p = 0.001$	$p < 0.0005$	$p = 0.5$	$p = 0.19$
Wilcoxon test		$p = 0.0005$	$p = 0.0005$	$p = 0.17$	$p = 0.25$

Table 4. Base composition for 21 independent comparisons of parasitic versus free-living metazoans. Adenine + thymine content (A+T), adenine only (A) was determined for the parasitic and free-living clade of each comparison for all 4 genes. Results of the two-tailed sign test and Wilcoxon signed-ranks test of the null hypothesis are given, where the null hypothesis states that biases in base composition are equally likely to be associated with a parasitic or free-living lifestyle.

#	Comparison	$A+T_{COI}$	A_{COI}	$A+T_{16S}$	A_{16S}	$A+T_{EF\text{-}1a}$	$A_{EF\text{-}1a}$	$A+T_{18S}$	A_{18S}
1	Cnidaria							50.4/55.9	25.6/27.6
2	Nematoda clade 1							49.9/52.4	25.7/26.0
3	Nematoda clade 2							51.6/52.2	26.4/26.2
4	Nematoda clade 4a							60.3/55.5	30.9/28.0
5	Nematoda clade 4b							53.5/53.9	28.0/27.4
6	Nematoda clade 5	69.4/70.3	23.8/26.6	80.1/77.5	36.0/36.7			53.7/53.3	26.5/25.7
7	Syndermata	60.4/63.8	22.8/20.9	64.9/68.1	29.5/31.8			54.8/53.4	27.5/26.6
8	Cirripedia	68.2/64.5	25.3/26.4	70.9/71.4	35.3/34.9			51.2/48.1	26.1/24.2
9	Maxillopoda	61.1/62.5	25.5/28.7					50.1/50.1	25.4/25.5
10	Copepoda	61.8/59.2	26.0/22.5	59.5/63.1	31.7/35.3			54.6/50.4	27.6/24.9
11	Isopoda	60.1/56.2	25.6/21.1	66.2/64.5	35.9/37.2			51.6/50.2	26.5/26.2
12	Diptera 1	70.2/71.6	31.4/32.0	74.5/75.4	34.7/35.5	53.1/49.3	26/24.4	58.2/58.9	28.1/28.5
13	Diptera 2			77.0/76.4	37.0/35.3			57.9/57.3	28.7/28.3
14	Diptera 3	71.4/71.6	31.3/32.9	79.7/83.6	41.5/41.5				
15	Diptera 4			81.0/81.8	39.4/41.4				
16	Diptera 5	69.9/66.0	29.8/30.3	74.7/75.0	34.7/35.4	55.3/58.3	29.7/28.9	58.8/59.5	29.3/30
17	Hymenoptera	74.6/69.4	33.7/33.7	78.7/74.6	39.0/34.7				
18	Psocodea	63.2/68.7	24.9/30.6	80.6/83.6	40.9/38.3	54.5/58.0	27.1/27.3	48.6/49.3	26.1/27.8
19	Platyhelminthes	66.6/65.1	20.7/23.4	68.1/74.4	29.2/37.2	52.5/56.1	27.3/29.7	50.6/55.6	25.4/29.0
20	Annelida	67.0/58.1	28.6/27.0	69.0/60.3	32.5/35.1	50.3/48.4	28.3/26.3	49.6/49.8	24.9/25.1
21	Bivalvia	58.5/62.2	18.0/23.0					50.2/50.9	25.4/25.4
Supportive comparisons		7 of 14	4 of 13	5 of 14	4 of 13	2 of 5	3 of 5	7 of 17	9 of 17
Sign test		1	0.267	0.424	0.267	1	1	0.629	1
Wilcoxon test		0.397	0.311	0.51	0.249	0.686	0.686	0.758	0.636

Table 5. Genome size variation for 10 phylogenetically independent comparisons. Genome size data was collated from the Animal Genome Size Database (Gregory 2001a). Results of the one-tailed sign test and Wilcoxon signed-ranks test of the null hypothesis are given, where the null hypothesis states that increases in genome size are equally likely to be associated with a parasitic or free-living lifestyle.

Biology	Comparison	# of species	Mean C-value	Ratio
Parasitic	Order Tylenchida	4	0.050	0.25
Free-living	Order Rhabditida	2	0.20	
Parasitic	Family Steinernematidae	1	0.23	0.88
Free-living	Family Panagrolaimidae	3	0.26	
Parasitic	Family Heterorhabditidae	1	0.039	0.24
Free-living	Family Rhabditidae	9	0.16	
Parasitic	Order Rhizocephala	1	0.67	0.47
Free-living	Order Thoracica	7	1.42	
Parasitic	Order Siphonostomatoida	1	0.58	2.32
Free-living	Order Harpacticoida	1	0.25	
Parasitic	Family Cymothoidae	1	8.82	1.84
Free-living	Family Sphaeromatidae	2	4.80	
Parasitic	Superfamily Oestroidea	3	0.68	0.8
Free-living	Superfamily Muscoidea	3	0.85	
Parasitic	Family Braconidae	2	0.16	0.8
Free-living	Family Apidae[a]	2	0.20	
Parasitic	Class Trematoda	4	1.1	0.52
Free-living	Class Turbellaria	57	2.13	
Parasitic	Class Hirudinea	7	0.39	0.50
Free-living	Family Lumbricidae	15	0.78	
Supportive comparisons				8 of 10
Sign test				$p = 0.055$
Wilcoxon test				$p = 0.055$

[a] secondarily free-living (see Dowton & Austin 2001)

General discussion and conclusions

General discussion and conclusions

Summary of results

My work has explored how lifestyle transitions influence the rates of molecular evolution in animals. Insights were gained by employing a phylogenetic comparative method to contrast relative rates of molecular divergence in groups of animals with lifestyle transitions in a diverse range of taxa. The results suggest that adopting either a planktonic or a parasitic lifestyle leads to an enhanced rate of molecular evolution. These results have broad implications for the application of molecular clocks, as well as for understanding both diversification patterns, and the factors that influence the tempo of animal molecular evolution.

My study of the phylogenetic relationships of the branchiopod crustaceans based on molecular data has provided the most resolved branchiopod phylogeny to date. The results support a number of longstanding hypotheses about affinities among orders, as well as relationships within the speciose Cladocera. A few affinities remain uncertain, but strategies have been proposed for their resolution. Furthermore, the study provided a phylogenetically independent planktonic-benthic comparison for analysis in Chapter 2.

My work also involved a comparative analysis examining the association between transitions from a benthic to planktonic lifestyle, and rates of molecular evolution. This study tested the hypothesis that increased exposure to UV-B radiation, coincident with planktonic lifestyles, leads to higher rates of molecular evolution. The results strongly supported this hypothesis by demonstrating that rates of nucleotide substitution have been consistently elevated in pelagic organisms relative to their benthic

allies and that increased mutagenesis is linked to the formation of UV-induced photoproducts.

Finally, my work has involved a comparative analysis of the association between transitions to parasitism and increased rates of sequence evolution. My studies revealed that parasitic lineages have accelerated rates of molecular evolution in comparison with their free-living relatives, a pattern that likely arises as a consequence of a genetic arms race between host and parasite. In addition to an accelerated rate of sequence evolution, this study provided evidence that parasites are experiencing other genome-wide changes, such as a reduction in DNA content.

The application of molecular clocks

The molecular clock concept has revolutionized the field of evolutionary biology by providing an invaluable tool for probing the timing of diversification events. If the rate of evolution is approximately constant between two taxa, the extent of genetic divergence between them can be used to estimate a timeframe for their separation from a common ancestor. Since it is now widely accepted that the molecular clock is often compromised due to selection and constraints on mutation (see Introduction), current studies should account for these deviations when estimating divergence dates. This is usually accomplished by either assuming a 'local molecular clock', which assumes that closely related taxa share a similar rate of molecular evolution, or by performing tests to detect departures from rate constancy and to remove rate-variable taxa (e.g. Takezaki et al. 1995; reviewed in Bromham et al. 2000). The present study has demonstrated that lineages that have undergone lifestyle transitions often experience accelerated rates of molecular evolution, suggesting that even local clocks must be applied cautiously.

Although the transitions examined in this study were deep (tens to hundreds of millions of years), more recent lifestyle transitions may also be linked to variable substitution rates, implying that closely related taxa could exhibit heterogeneous rates. Furthermore, the tests most commonly employed to diagnose rate heterogeneity have low power and perform poorly, particularly for short sequence lengths (< 1 kb), such as those typically used for inferring phylogenies (Bromham et al. 2000). These findings have important implications for the accuracy of past and current estimates of divergence dates, because the inclusion of rate accelerated taxa will lead to consistent overestimates. It is critical that future work accounts for rate heterogeneity, particularly when the taxonomic sample includes organisms with varying lifestyles (e.g. parasitic and free-living).

Adaptive zones and diversification

Simpson (1953) suggested that rapid rates of diversification often follow the adoption of a new lifestyle i.e. the transition into a new "adaptive zone". This proliferation may be due to pre-adaptations to newly available niches, or to the evolution of 'key innovations' to exploit newly available resources (Futuyma 1986). These transitions occur rarely since they require substantial evolutionary change, but often result in adaptive radiations when they do (Mitter et al. 1988; Schluter 2000; Miller and Crespi 2003). The adoption of a planktonic or a parasitic lifestyle would certainly qualify as a transition to an adaptive zone, but it is unclear if they have been accompanied by rapid diversification. While there is evidence of radiations in both planktonic (Rigby and Molsom 2000) and parasitic lineages (Poulin and Morand 2000), comparative analyses have yet to be performed to determine if these two groups are more diverse than their benthic and free-living relatives, respectively. However, a

number of studies have demonstrated a positive association between accelerated rates of molecular evolution and diversification (Mindell et al. 1990; Barraclough and Savolainen 2001; Webster et al. 2003). Genetic change is necessary for speciation (Harrsion 1991; Coyne 1992) and these studies would suggest that rates of speciation are accelerated by elevated rates of sequence evolution, most likely by increasing the rate at which populations diverge and reproductive isolation evolves (Barraclough and Savolainen 2001). Hence, planktonic and parasitic lineages, which have elevated rates of molecular evolution, may also be more taxonomically diverse than their non-planktonic and non-parasitic relatives. This inference has yet to be corroborated with a comparative test for an association between diversification and transitions to planktonic or parasitic lifestyles.

Factors influencing animal molecular evolution

Studies of molecular rate variation often consider four common factors known to influence the tempo of animal molecular evolution: generation time, DNA repair and replication efficiency, metabolic rate, and population size (see Introduction). Although it is possible that these factors have played a role in determining the rate of molecular evolution among the organisms examined in this study, the results suggest that other factors have played a larger role. Firstly, the acceleration of planktonic lineages in comparison with their benthic relatives is best explained by the 'UV-exposure' hypothesis. This is consistent with several past studies (Smith et al. 1992; Lutzoni and Pagel 1997; Pawlowski et al. 1997; Kooistra and Medlin 1996; Hebert et al. 2002) and may be a potentially important mechanism among other taxa (e.g. plants). Secondly, the enhanced substitution rate evident in parasitic lineages is likely a consequence of a

genetic arms race. Again, this is consistent with previous work (Dowton and Austin 1995; Page et al. 1998) and may have played an important role in the diversification of the 60+ lineages of parasitic metazoans (Poulin and Morand 2000) and other antagonistic symbionts (e.g. pathogens; parasitic plants). While these two factors will obviously only have an influence in restricted lineages, they nonetheless illuminate the need to consider alternatives in addition to the four common determinants of molecular evolution. Moreover, in light of the numerous studies that have failed to support these four hypotheses over a broad range of varied taxa, the current view of rates of molecular evolution may be biased because comparative data is only available for these four factors. Expanding our comparative data to consider the impacts of additional biological and environmental factors will undoubtedly expand our understanding of the modulation of molecular evolution.

References

Adamkewicz, S. L., M. G. Harasewych, J. Blake, D. Saudek, and C. J. Bul. 1997. A molecular phylogeny of the bivalve mollusks. Mol. Biol. Evol. **14**:619-629.

Armstrong, M. R., V. C. Blok, and M. S. Philips. 2000. A multipartite mitochondrial genome in the potato cyst nematode *Globodera pallida*. Genetics **154**:181-192.

Ax, P. 1999. Das System der Metazoa II - Ein Lehrbuch der Phylogenetischen Systematik. 1-383. G. Fischer Verlag, Stuttgart.

Barraclough, T. G. and V. Savolainen. 2001. Evolutionary rates and species diversity in flowering plants. Evolution 55:677-683.

Berge, J., G.A. Boxshall, and W. Vader. 2003. Cladistic analysis of the Amphipoda, with special emphasis on the origin of the Stegocephalidae. Proceedings of the Xth. International Colloquium on Amphipoda. Polskie Archivum Hydrobiologii.

Bernstein C. 1981. Deoxyribonucleic acid repair in bacteriophage. Microbiol. Rev. **45**:72-98.

Biggs, R. H., and M. E. B. Joyner. 1994. Stratospheric ozone depletion/ UV-B radiation in the biosphere. Springer-Verlag, Berlin.

Black IV, W.C., and R. L. Roehrdanz. 1998. Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes. Mol. Biol. Evol. **15**:1772-1785.

Blaxter, M. L., P. De Ley, J. R. Garey, L. X. Liu, P. Scheldeman, A. Vierstraete, J. R. Vanfleteren, L. Y. Mackey, M. Dorris, L. M. Frisse, J. T. Vida, and W. K. Thomas. 1998. A molecular evolutionary framework for the phylum Nematoda. Nature **392**:71-75.

Boore, J. L. 2000. The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals. Pp. 133-147 *in* D. Sankoff and J. Nadeau, eds. Comparative genomics. Kluwer Academic Publishers, Netherlands.

Bowman, T., and L. G. Abele. 1982. Classification of the Crustacea. Pp. 1-27 in L. G. Abele, ed. Biology of the Crustacea, vol. I. Academic Press, New York.

Braband, A., S. Richter, R. Hiesel, and G. Scholtz. 2002. Phylogenetic relationships within the Phyllopoda (Crustacea, Branchiopoda) based on mitochondrial and nuclear markers. Mol. Phylogenet. Evol. **25**:229–244.

Brash, D. E., J. A. Rudolph, J. A. Simon, A. Lin, G. J. McKenna, H. P. Baden, A. J. Halperin, and J. Pontén. 1991. A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. Proc. Natl. Acad. Sci. USA **88**:10124–10128.

Bridge, D., C. W. Cunningham, R. DeSalle, and L. W. Buss. 1995. Class-level relationships in the phylum Cnidaria: Molecular and morphological evidence. Mol. Biol. Evol. **12**:679-689.

Britten, R. J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. Science **231**:1393-1398.

Bromham, L. 2002. Molecular clocks in reptiles: life history influences rate of molecular evolution. Mol. Biol. Evol. **19**:302-309.

Bromham, L., and M. Cardillo. 2003. Testing the link between the latitudinal gradient in species richness and rates of molecular evolution. J. Evol. Biol. **16**:200-207.

Bromham, L., and D. Penny. 2003. The modern molecular clock. Nature Rev. Genet. **4**:216-224.

Bromham, L., A. Rambaut, and P. H. Harvey. 1996. Determinants of rate variation in mammalian DNA sequence evolution. J. Mol. Evol. **43**:610-621.

Bromham, L., A. Rambaut, M. D. Hendy, and D. Penny. 2000. The power of relative rates tests depends on the data. J. Mol. Evol. **50**:296-301.

Bromham, L., M. Woolfit, M. S. Lee, and A. Rambaut. 2002. Testing the relationship between morphological and molecular rates of change along phylogenies. Evolution **56**:1921-1930.

Brusca, R. C., and G. J. Brusca. 1990. Invertebrates. Sinauer, Sunderland, MA.

Brusca, R. C., and G. D. F. Wilson. 1991. A phylogenetic analysis of the Isopoda with some classificatory recommendations. Mem. Queensland Mus. **31**:143-204.

Burger, G., L. Forget, Y. Zhu, M. W. Gray, and B. F. Lang. 2003a. Unique mitochondrial genome architecture in unicellular relatives of animals. Proc. Natl. Acad. Sci. U.S.A. **100**:892-897.

Burger, G, M. W. Gray, and B. F. Lang. 2003b. Mitochondrial genomes - anything goes. Trends Genet. **19**:709-716.

Burt, A. 1989. Comparative methods using phylogenetically independent contrasts. Oxf. Surv. Evol. Biol. **6**:33-53.

Calkins, J., and T. Thordardottir. 1980. The ecological significance of solar UV radiation on aquatic organisms. Nature **283**:563-566.

Campbell, N. J. H., and S. C. Barker. 1998. An unprecedented major rearrangement in an arthropod mitochondrial genome. Mol. Biol. Evol. **15**:1786-1787.

Campbell, N. J. H., and S. C. Barker. 1999. The novel mitochondrial gene arrangement of the cattle tick, *Boophilus microplus*: five-fold tandem repetition of a coding region. Mol. Biol. Evol. **16**:732-740.

Capaldi, R. A. 1990. Structure and function of cytochrome c oxidase. Annu. Rev. Biochem. **59**:569-596.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. **17**:540-552.

Castro, L.R., A. D. Austin, and M. Dowton. 2002. Contrasting rates of mitochondrial molecular evolution in parasitic Diptera and Hymenoptera. Mol. Biol. Evol. **19**:1100-1113.

Chao, L., and E. C. Cox. 1983. Competition between high and low mutating strains of *Escherichia coli*. Evolution **37**:125-134.

Childress, J. J. 1995. Are there physiological and biochemical adaptations of metabolism in deep-sea animals? Trends Ecol. Evol. **10**:30-36.

Clark, M. A., N. A. Moran, and P. Baumann. 1999. Sequence evolution in bacterial endosymbionts having extreme base composition. Mol. Biol. Evol. **16**:1586-1598.

Cohen, A. C., J. W. Martin, and L. S. Kornicker. 1998. Homology of Holocene ostracode biramous appendages with those of other crustaceans: the protopod, epipod, exopod and endopod. Lethaia **31**:251-265.

Collins, A. G. 2002. Phylogeny of Medusozoa and the evolution of cnidarian life cycles. J. Evol. Biol. **15**:418-432.

Coyne, J. A. 1992. The genetics of speciation. Nature **355**:511-515.

Crandall, K. A., D. J. Harris, and J. W. Fetzner. 2000. The monophyletic origin of freshwater crayfishes estimated from nuclear and mitochondrial DNA sequences. Proc. R. Soc. Lond. Ser. B **267**:1679-1686.

Crease T., and J. Colbourne. 1998. Unusually long small subunit ribosomal RNA of the crustacean, *Daphnia pulex*: sequence and predicted secondary structure. J. Mol. Evol. **46**:307-313.

Crease, T. J., and D. J. Taylor. 1998. The origin and evolution of variable-region helices in V4 and V7 of the small-subunit ribosomal RNA of branchiopod crustaceans. Mol. Biol. and Evol. **15**:1430-1446.

Cristescu, M. E. A., and P. D. N. Hebert. 2002. Phylogeny and adaptive radiation in the Onychopoda (Crustacea: Cladocera): evidence from multiple gene sequences. J. Evol. Biol. **15**:838-849.

Cunningham, C. W. 1997. Is congruence between data partitions a reliable predictor of phylogenetic accuracy? Empirically testing an interative procedure for choosing among phylogenetic methods. Syst. Biol. **46**:464-478.

Dayrat, B., A. Tillier, G. Lecointre, and S. Tillier. 2001. New clades of euthyneuran gastropods (Mollusca) from rRNA sequences. Mol. Phylogenet. Evol. **19**:225-235.

De Rijk, P., J. Wuyts, Y. Van de Peer, T. Winkelmans, R. De Wachter. 2000. The European large subunit ribosomal RNA database. Nucleic Acids Res. **28**:177–178.

DeSalle, R., and A. R. Templeton. 1988. Founder effects accelerate the rate of mtDNA evolution of Hawaiian *Drosophila*. Evolution **42**:1076-1085.

Dickerson, R. E. 1971. The structure of cytochrome c and the rates of molecular evolution. J. Mol. Evol. **1**:26-45.

Dorris, M., P. De Ley, and M. L. Blaxter. 1999. Molecular analysis of nematode diversity and the evolution of parasitism. Parasitol. Today **15**:188-193.

Dowton, M., and A. D. Austin. 1995. Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the Hymenoptera. J. Mol. Evol. **41**: 958-695.

Dowton, M., and A. D. Austin. 1997. The evolution of strand-specific compositional bias. A case study in the hymenopteran mitochondrial 16S rRNA gene. Mol. Biol. Evol. **44**:109-112.

Dowton, M., and A. D. Austin. 2001. Simultaneous analysis of 16S, 28S, COI and morphology in the Hymenoptera: Apocrita – evolutionary transitions among parasitic wasps. Biol. J. Linn. Soc. **74**:87-111.

Dowton, M., and N. J. H. Campbell. 2001. Intramitochondrial recombination – is it why some mitochondrial genes sleep around. Trends Ecol. Evol. **16**:269-271.

Dreyer, H., and J. W. Wagele. 2001. Parasites of crustaceans (Isopoda:Bopyridae) evolved from fish parasites: Molecular and morphological evidence. Zool-Anal. Complex Sy. **103**:157-78.

Dumont, H. J., and M. Silva-Briano. 1998. A reclassification of the anomopod families Macrothricidae and Chydoridae, with the creation of a new suborder, the Radopoda (Crustacea : Branchiopoda). Hydrobiol. 384:119-149.

Ehrlich, P. R., and P. H. Raven. 1964. Butterflies and plants: a study in coevolution. Evolution **18**:586-608.

Endo, T., K. Ikeo, and T. Gojobori. 1996. Large-scale search for genes on which positive selection may operate. Mol. Biol. Evol. **13**:685-690.

Eriksson, S. 1934. Studien Ïber die Fangapparate der Branchiopoden nebst einigen phylogenetischen Bemerkungen. Zool. Bidrag Upps. **15**:23-287.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1994. Testing significance of incongruence. Cladistics **10**:315-319.

Feagin, J. E. 1994. The extrachromosomal DNAs of Apicomplexan parasites. Ann. Rev. Microbiol. **48**:81-104.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. **125**:1-15.

Folmer, O., M. Black, R. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. **3**:294-299.

Friedberg, E.C. 2003. DNA damage and repair. Nature **421**:436-439.

Friedberg, E. C., G. C. Walker, and W. Siede. 1995. DNA repair and mutagenesis. ASM Press, Washington D.C.

Fryer, G. 1995. Phylogeny and adaptive radiation within the Anomopoda: a preliminary exploration. Hydrobiol. **307**:57-68.

Futuyma, D. J. 1986. Evolutionary biology. 2nd ed. Sinnauer Associates, Sunderland, Massachusetts.

Garey, J. R., A. Schmidt-Rhaesa, T. J. Near, and S. A. Nadler. 1998. The evolutionary relationships of rotifers and acanthocephalans. Hydrobiol. **387**:83-91.

Ghilarov, A. M. 1984. The paradox of the plankton reconsidered; or why do species coexist? Oikos **43**:46-52.

Gillespie, J. H. 1991. The causes of molecular evolution. Oxford University Press, Oxford.

Giribet, G., S. Carranza, J. Baguna, M. Riutort, and C. Ribera. 1996. First molecular evidence for the existence of a Tardigrada + Arthropoda clade. Mol. Biol. Evol. **13**:76–84.

Giribet, G., G. D. Edgecombe, and W. C. Wheeler. 2001. Arthropod phylogeny based on eight molecular loci and morphology. Nature **413**:157-161.

Gojobori, T., W. H. Li, and D. Graur. 1982. Patterns of nucleotide substitution in pseudogenes and functional genes. J. Mol. Evol. **18**:360-369.

Gregory, T. R. 2001a. Animal Genome Size Database. http://www.genomesize.com.

Gregory, T. R. 2001b. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. Biol. Rev. **76**:65-101.

Gregory, T. R. and P. D. N. Hebert. 1999. The modulation of DNA content: proximate causes and ultimate consequences. Genome Res. **9**:317-324.

Gross, N. J., G. S. Getz, and M. Rabinowitz. 1969. Apparent turnover of mitochondrial deoxyribonucleic acid and mitochondrial phospholipids in the tissues of the rat. J. Biol. Chem. **244**:1552-1562.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis.

Hanner, R., and M. Fugate. 1997. Branchiopod phylogenetic reconstruction from 12S rDNA sequence data. J. Crust. Biol. **17**:174-183.

Hafner, M. S., P. D. Sudman, F. X. Villablanca, T. A. Spradling, J. W. Demastes, and S. A. Nadler. 1994. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science **265**:1087-1090.

Haraguchi Y., and A. Sasaki. 1996. Host-parasite arms race in mutation modifications: Indefinite escalation despite a heavy load? J. Theor. Biol. **183**:121-137.

Harrison, R. G. 1991. Molecular changes at speciation. Annu. Rev. Ecol. Syst. **22**:281-308.

Harm, W. 1980. Biological effects of ultraviolet radiation. Cambridge University Press, Cambridge.

Hart, R.W., and R. B. Setlow. 1974. Correlation between deoxyribonucleic acid excision repair and life span in a number of mammalian species. Proc. Nat. Acad. Sci, USA **71**:2169-2173.

Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology. Oxford University Press, Oxford.

Haydon, D. T., A. Bastos, A. Samuel, and N. Knowles. 2001. Evidence for positive selection in foot-and-mouth-disease-virus capsid genes from field isolates. Genetics **157**:7-15.

Haynes, R. H., and Kunz, B. A. 1981. DNA repair and mutagenesis in yeast. Pp. 317-414 *in* Strathern J. N., Jones E.W. and Broach J. R., eds. The molecular biology of yeast *Saccharomyces*. Life cycle and inheritance. Cold Spring Harbor Laboratory Press, New York.

Hebert, P. D. N., S. Ratnasingham, and J. R. deWaard 2003. Barcoding animal life: COI divergences among closely related species. Proc. R. Soc. Lond. Ser. B **270**:S96-S99.

Hebert, P. D. N., E. A. Remigio, J. K. Colbourne, D. J. Taylor, and C. C. Wilson. 2002. Accelerated molecular evolution in halophilic crustaceans. Evolution **56**:909-926.

Herlyn, H., O. Piskurek, J. Schmitz, U. Ehlers, and H. Zischler. 2003. The syndermatan phylogeny and the evolution of acanthocephalan endoparasitism as inferred from 18S rDNA sequences. Mol. Phylogenet. Evol. **26**:155–164.

Hillis, D. M., and J. P. Huelsenbeck. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. J. Hered. **83**:189-195.

Hoeh, W. R., M. B. Black, R. Gustafson, A. E. Bogan, R. A. Lutz et al. 1998. Testing alternative hypotheses of *Neotrigonia* (Bivalvia: Trigonioida) relationships using cytochrome c oxidase subunit I DNA sequences. Malacologia **40**:267-278.

Hoeijmakers, J. H. 2001. Genome maintenance mechanisms for preventing cancer. Nature **411**:366-374.

Huelsenbeck, J. P., and K. A. Crandall. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu. Rev. Ecol. Syst. **28**:437-466.

Huelsenbeck, J. P., and Ronquist, F. 2001. MRBAYES. Bayesian inference of phylogeny. Bioinformatics **17**:754-755.

Hugall A., J. Stanton, and C. Moritz. 1997. Evolution of the AT-rich mitochondrial DNA of the root knot nematode, *Meloidogyne hapla*. Mol. Biol. Evol. **14**:40-48.

Huys, R., and G. A. Boxshall. 1991. Copepod evolution. The Ray Society, London.

Hu, G., and W. G. Thilly. 1994. Evolutionary trail of the mitochondrial genome as based on human 16S rDNA pseudogenes. Gene **147**:197-204.

Itoh, T., W. Martin, and M. Nei. 2002. Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. Proc. Natl. Acad. Sci. USA **99**:12944-12948.

Jiggins, F. M., G. D. D. Hurst, and Z. Yang. 2002. Host-symbiont conflicts: positive selection on the outer membrane protein of parasite but not mutualistic Rickettsiaceae. Mol. Biol. Evol. **19**:1341-1349.

Johnson K. P., R. H. Cruickshank, R. J. Adams, V. S. Smith, R. D. Page, and D. H. Clayton. 2003. Dramatically elevated rate of mitochondrial substitution in lice (Insecta: Phthiraptera). Mol. Phylogenet. Evol. **26**:231-242.

Johnson, K. P., and J. Seger. 2001. Elevated rates of nonsynonymous substitution in island birds. Mol. Biol. Evol. **18**:874-881.

Karentz, D. 1991. Ecological considerations of Antarctic ozone depletion. Antarctic Sci. **3**:3-11.

Karentz, D., J. E. Cleaver, and D. L. Mitchell. 1991. DNA damage in the Antarctic. Nature **350**: 28.

Katinka M. D, S. Duprat, E. Cornillot, G. Metenier, F. Thomarat, et al. 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature **414**:450–453.

Kimura, M., and Ohta, T. 1973. Eukaryotes-prokaryotes divergence estimated by 5S ribosomal RNA sequences. Nature **229**:467-469.

Kluge, A. J. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. **38**:7-25.

Kohne, D. E. 1970. Evolution of higher-organism DNA. Quart. Rev. Biophys. **33**:327-375.

Kooistra, W. H. C. F., and L. K. Medlin. 1996. Evolution of the diatoms (Bacillariophyta) IV. A reconstruction of their age from small subunit rRNA coding regions and the fossil record. Mol. Phylogenet. Evol. **6**:391-407.

Koonin, E. V., A. R. Mushegian, and K. E. Rudd. 1996. Sequencing and analysis of bacterial genomes. Curr. Biol. **6**:404-416.

Korber, B., M. Muldoon, J. Theiler, F. Gao, R. Gupta, A. Lapedes, B. H. Hahn, S. Wolinsky, and T. Bhattacharya. 2000. Timing the ancestor of the HIV-1 pandemic strains. Science **288**:1789-1796.

Kumar, S. 1996. PHYLTEST. Version 2.0. Pennsylvania State University, University Park.

Kunkel, T. A. 1992. DNA replication fidelity. J. Biol. Chem. **267**:18251-18254.

Lambert, J. L., and N. A. Moran. 1998. Deleterious mutations destabilize ribosomal RNA of endosymbionts. Proc. Natl. Acad. Sci. USA **95**:4458-4462.

Lavrov, D. V., J. L. Boore, and W. M. Brown. 2002. Complete mtDNA sequences of two
millipedes suggest a new model for mitochondrial gene rearrangements: Duplication and non-random loss. Mol. Biol. Evol. **19**:163-169.

LeClerc, J. E., B. Li, W. L. Payne, and T. A. Cebula. 1996. High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science **274**:1208-1211.

Le, T. H., D. Blair, T. Agatsuma, P. F. Humair, N. J. H. Campbell, et al. 2000. Phylogenies inferred from mitochondrial gene orders-a cautionary tale from the parasitic flatworms. Mol. Biol. Evol. **17**:1123-1125.

Lesser, M.P. 1996. Elevated temperatures and ultraviolet radiation cause oxidative stress and inhibit photosynthesis in symbiotic dinoflagellates. Limnol. Oceanogr. **41**:271-283.

Li, W. H. 1997. Molecular evolution. Sinauer Associates, Sunderland, MA.

Li, W. H., and M. Tanimura. 1987. The molecular clock runs more slowly in man than in apes and monkeys. Nature **326**:93-96.

Li, W. H., D. L. Ellsworth, J. Krushkal, B. H. Chang, and D. Hewett-Emmett. 1996. Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. Mol. Phylogenet. Evol. **5**:182-187.

Littlewood, D. T. J., K. Rohde, R. A. Bray, and E. A. Herniou. 1999. Phylogeny of the Platyhelminthes and the evolution of parasitism. Biol. J. Linn. Soc. **68**:257-287.

Lockyer, A. E., P. D. Olson, and D. T. J. Littlewood. 2003. Utility of complete large and small subunit rRNA genes in resolving the phylogeny of the Neodermata (Platyhelminthes): implications and a review of the cercomer theory. Biol. J. Linn. Soc. **78**:155-173.

Lukes, J., D. L. Guilbride, J. Votypka, A. Zíková, R. Benne, and P. T. Englund. 2002. Kinetoplast DNA network: evolution of an improbable structure. Eukaryot. Cell **1**:495-502.

Lutzoni, F., and M. Pagel. 1997. Accelerated molecular evolution as a consequence of transitions to mutualism. Proc. Natl. Acad. Sci. USA **94**:11422-11427.

Lynch, M. 1996. Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. Mol. Biol. Evol. **13**:209-220.

Lyons-Weiler, J., G. A. Hoelzer, and R. J. Tausch. 1996. Relative Apparent Synapomorphy Analysis (RASA) I: the statistical measurement of phylogenetic signal. Mol. Biol. Evol. **13**:749-757.

Malloy, K. D., M. A. Holman, D. Mithchell, and H. W. Detrich III. 1997. Solar UVB-induced DNA damage and photoenzymatic DNA repair in antarctic zooplankton. Proc. Natl. Acad. Sci. USA **94**:1258-1263.

Mark Welch, D. B., and M. Meselson. 2001. Rates of nucleotide substitution in sexual and anciently asexual rotifers. Proc. Natl. Acad. Sci USA **98**:6720-6724.

Martin, A. P. 1995. Metabolic rate and directional nucleotide substitution in animal mitochondrial DNA. Mol. Biol. Evol. **12**:1124-1131.

Martin, A. P., G. J. P. Naylor, and S. R. Palumbi. 1992. Rates of mitochondrial DNA evolution in sharks are slow compared to mammals. Nature **357**:153-155.

Martin, A. P., and S. R. Palumbi. 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc. Natl. Acad. Sci. USA **90**:4087-4091.

Martin, J. W., and G. E. Davis. 2001. An updated classification of the Recent Crustacea. Natural History Museum of Los Angeles County Science Series **39**. Los Angeles, CA.

Martin, P. 2001. On the origin of the Hirudinea and the demise of the Oligochaeta. Proc. Royal Soc. London Ser. B **268**:1089-1098.

Maynard Smith, J. 1978. The evolution of sex. Cambridge University Press, Cambridge.

McAlpine, J. F., and D. M. Wood (eds.). 1989. Manual of Nearctic Diptera, Vol. 3. Research Branch, Agriculture Canada, Monograph 32.

Mercer, J. M., and V. L. Roth. 2003. The effects of Cenozoic global change on squirrel phylogeny. Science **299**:1568-1572.

Miller, D. G., and B. Crespi. 2003. The evolution of inquilinism, host-plant use and mitochondrial substitution rates in *Tamalia* gall aphids. J. Evol. Biol. **16**:731-743.

Mindell, D. P., J. W. Sites, and D. Graur. 1990. Mode of allozyme evolution: Increased genetic distance associated with speciation events. J. Evol. Biol. **3**:125-131.

Mindell, D. P., and C. E. Thacker. 1996. Rates of molecular evolution: phylogenetic issues and applications. Annu. Rev. Ecol. Syst. **27**:279-303.

Mira, A., H. Ochman, and N. A. Moran. 2001. Deletional bias and the evolution of bacterial genomes. Trends Genet. **17**:589-596.

Mitter, C., B. Farrell, and B. Wiegmann. 1988. The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? Am. Nat. **132**:107-128.

Molina, L. T. and M. J. Molina. 1986. Absolute absorption cross sections of ozone in the
185 to 350 nm wavelength range. J. Geophys. Res. **91**:14501-14508.

Mooers, A. O., and P. H. Harvey. 1994. Metabolic rate, generation time and the rate of molecular evolution in birds. Mol. Phylogenet. Evol. **3**:344-350.

Moran, N. A. 1996. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. Proc. Natl. Acad. Sci. USA **93**:2873-2878.

Moran, N. A. 2002. Microbial minimalism: genome reduction in bacterial pathogens. Cell **108**:583-586.

Muller, H. J. 1964. The relation of recombination to mutational advance. Mut. Res. **1**:2-9.

Muse, S. V., and B. S. Weir. 1992. Testing for equality of evolutionary rates. Genetics **132**:269-276.

Nappi, A. J., E. Vass, F. Frey, and Y. Carton. 1995. Superoxide anion generation in *Drosophila* during melanotic encapsulation of parasites. Eur. J. Cell. Biol. 68:450-456.

Nappi, A. J., E. Vass, F. Frey, and Y. Carton. 2000. Nitric oxide involvement in *Drosophila* immunity. Nitric Oxide 4:423-430.

Nee, S. 1989. Antagonistic coevolution and the evolution of genotype randomization. J. Theor. Biol. **140**:499-518.

Negrea, S., N. Botnariuc, and H. J. Dumont. 1999. Phylogeny, evolution and classification of the Branchiopoda (Crustacea). Hydrobiol. **412**:191–212.

Nickrent, D. L., and E. M. Starr. 1994. High rates of nucleotide substitution in nuclear small-subunit (18S) rDNA from holoparasitic flowering plants. J. Mol. Evol. **39**:62-70.

Normark, B. B, O. P. Judson, and N. A. Moran. 2003. Genomic signatures of ancient asexual lineages. Biol. J. Linn. Soc. **79**:69-84.

Nunn, G. B., and S. E. Stanley. 1998. Body size effects and rates of cytochrome b evolution in tube-nosed seabirds. Mol. Biol. Evol. **15**:1360-1371.

Ohta, T. 1972. Population size and the rate of evolution. J. Mol. Evol. **1**:305-314.

Ohta, T. 1976. Role of slightly deleterious molecular evolution and polymorphism. Theor. Popul. Biol. **10**:254-275.

Ohta, T. 1992. The nearly neutral theory of molecular evolution. Ann. Rev. Ecol. Syst. **23**:263-86.

Ohta, T. 1993. An examination of the generation time effect on molecular evolution. Proc. Natl. Acad. Sci. USA **90**:10676-10680.

Ohta T., and M. Kimura. 1971. On the rate of molecular evolution. J. Mol. Evol. **1**:1-17.

Olesen, J. 1998. A phylogenetic analysis of the Conchostraca and Cladocera (Crustacea, Branchiopoda, Diplostraca). Zool. J. Linn. Soc. **122**:491-536.

Olesen, J. 1999. Larval and post-larval development of the branchiopod clam shrimp *Cyclestheria hislopi* (Baird, 1859) (Crustacea, Branchiopoda, Conchostraca, Spinicaudata). Acta Zool. **80**:163-184.

Olesen, J. 2000. An updated phylogeny of the Conchostraca-Cladocera clade (Branchiopoda, Diplostraca). Crustaceana **73**:869-886.

Olesen, J., S. Richter, and G. Scholtz. 2003. On the ontogeny of *Leptodora kindtii* (Crustacea, Branchiopoda, Cladocera), with notes on the phylogeny of the Cladocera. J. Morph. **256**: 235-259.

Page, R. D. M., P. L. M. Lee, S. A. Becher, R. Griffiths, and D. H. Clayton. 1998. A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. Mol. Phylogenet. Evol. **9**:276-293.

Palumbi, S. R. 1996. Nucleic acids II: the polymerase chain reaction. Pp. 205–247 *in* D. Hillis and C. Moritz B. K. Mable, eds. Molecular systematics. Sinauer, Sunderland, MA.

Parkinson J., M. Mitreva, N. Hall, M. Blaxter, and J. P. McCarter. 2003. 400000 nematode ESTs on the net. Trends Parasitol. **19**:283-286.

Pawlowski, J., I. Bolivar, J. Fahrni, C. de Vargas, M. Gouy, and L. Zaninetti. 1997. Extreme differences in rates of molecular evolution of foraminifera revealed by comparison of ribosomal DNA sequences and the fossil record. Mol. Biol. Evol. **14**:498-505.

Poulin, R., and S. Morand. 2000. The diversity of parasites. Q. Rev. Biol. **75**:277-293.

Posada D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics **14**:817-818.

Rand, D. M. 1993. Endotherms, ectotherms and mitochondrial genome-size variation. J. Mol. Evol. **37**:281-295.

Rand, D. M. 1994. Thermal habit, metabolic rate, and the evolution of mitochondrial DNA. Trends Ecol. Evol. **9**:125-131.

Rand, D. M. 2001. The units of selection on mitochondrial DNA. Annu. Rev. Ecol. Syst. **32**:415-448.

Remigio, E.A., and P. D. N. Hebert. 2000. Affinities among anostracan (Crustacea: Branchiopoda) families inferred from phylogenetic analyses of multiple gene sequences. Mol. Phylogenet. Evol. **17**:117-128.

Remsen, J., and R. DeSalle. 1998. Character congruence of multiple data partitions and the origin of the Hawaiian Drosophilidae. Mol. Phylogenet. Evol. **9**:225-235.

Richter S., A. Braband, N. Aladin, and G. Scholtz. 2001. The phylogenetic relationships of predatory water-fleas (Cladocera: Onychopoda, Haplopoda) inferred from 12S rDNA. Mol. Phylogenet. Evol. **19**:105-113.

Rigby, S., and C. V. Milsom. 2000. Origins, evolution and diversification of zooplankton. Annu. Rev. Ecol. Syst. **31**:293-313.

Roberts, L. 1989. Does the ozone hole threaten Antarctic life? Science **244**:288-289.

Robinson-Rechavi, M., and D. Huchon. 2000. RRTree: Relative-Rate Tests between groups of sequences on a phylogenetic tree. Bioinformatics **16**:296-297.

Robinson, M., M. Gouy, C. Gautier, and D. Mouchiroud. 1998. Sensitivity of the relative
-rate test to taxonomic sampling. Mol. Biol. Evol. **15**:1091-1098.

Rocha, E., and A. Danchin. 2002. Base composition bias might result from competition for metabolic resources. Trends Genet. 18:291-294.

Ronquist, F., A. P. Rasnitsyn, A. Roy, K. Eriksson, and M. Lindren. 1999. Phylogeny of the Hymenoptera: A cladistic reanalysis of Rasnitsyn's (1988) data. Zool. Scripta **28**:13-50.

Rosenberg, M. S., and S. Kumar. 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. Proc. Natl. Acad. Sci. USA **98**:10751-10756.

Rowe, D. L., and R. L. Honeycutt. 2002. Phylogenetic relationships, ecological correlates, and molecular evolution within the Cavioidea (Mammalia, Rodentia). Mol. Biol. Evol. **19**:263-277.

Sacherova, V., and P. D. N. Hebert. 2003. The evolutionary history of the Chydoridae (Crustacea: Chydoridae). Biol. J. Linn. Soc. **79**:629-643.

Sackton, T. B., R. Haney, and D. M. Rand. 2003. Cytonuclear coadaptation in *Drosophila*: Disruptions of cytochrome c oxidase activity in backcross genotypes. Evolution **57**:2315-2325.

Sale, J. E., M. Bemark, G. T. Williams, C. J. Jolly, M. R. Ehrenstein, et al. 2001. *In vivo* and *in vitro* studies of immunoglobulin gene somatic hypermutation. Phil. Trans. R. Soc. Lond. B **356**:21-28.

Sanders H. L. 1963. Significance of the Cephalocarida. Pp. 163-175 in H. B. Whittington and W. D. I. Rolfe (eds.). Phylogeny and evolution of Crustacea. Museum of Comparative Zoology, Cambridge, Massachusetts.

Sarich, V. M., and A. C. Wilson. 1973. Generation time and genomic evolution in primates. Science **179**:1144-1147.

Schluter, D. 2000. The ecology of adaptive radiation. Oxford University Press, Oxford.

Schmitz, J., and R. F. A. Moritz. 1998. Sociality and the rate of rDNA sequence evolution in wasps (Vespidae) and honeybees (*Apis*). J. Mol. Evol. **47**:606-612.

Schon, I., and K. Martens. 1998. Opinion: DNA repair in ancient asexuals -- a new solution for an old problem? J. Nat. Hist. **32**:943-948.

Schram, F. R. 1986. Crustacea. Oxford University Press, Oxford.

Schram, F. R. 2001. Phylogeny of decapods: moving towards a consensus. Hydrobiol. **449**:1-20.

Schwenk, K., A. Sand, M. Boersma, M. Brehm, E. Mader, D. Offerhaus, and P. Spaak. 1998. Genetic markers, genealogies and biogeographic patterns in the Cladocera. Aquat. Ecol. **32**:37-51.

Shao, R., N. J. H. Campbell, and S. C. Barker. 2001. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera) Mol. Biol. Evol. **18**:858-865.

Shao, R., M. Dowton, A. Murrell, and S. C. Barker. 2003. Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. Mol. Biol. Evol. **20**:1612-1619.

Shigenaga, M. K., C. J. Gimeno, and B. C. Ames. 1989. Urinary 8-hydroxy-2'-deoxyguanosine as a biological marker of *in vivo* oxidative DNA damage. Proc. Natl Acad. Sci. USA **86**:9697-9701.

Siddall, M. E., K. Apakupakul, E. M. Burreson, K. A. Coates, C. Erseus, S. R. Gelder, M. Kallersjo, and H. Trapido-Rosenthal. 2001. Validating Livanow: Molecular data agree that leeches, branchiobdellidans, and *Acanthobdella peledina* form a monophyletic group of oligochaetes. Mol. Phylogenet. Evol. **21**:346-351.

Simpson, G. G. 1953. The major features of evolution. Columbia University Press, New York.

Singer, C. E. and B. N. Ames. 1970. Sunlight ultraviolet and bacterial DNA base ratios. Science **170**:822-826.

Slowinski, J. B., and B. S. Arbogast. 1999. Is there an inverse relationship between body size and the rate of molecular evolution? Syst. Biol. **48**:396-399.

Smirnov, N. N. 1992. Mesozoic Anomopoda (Crustacea) from Mongolia. Zool. J. Linn. Soc. **104**:97-116.

Smith, R. C., and K.S. Baker 1979. Penetration of UV-B and biologically effective dose-rates in natural waters. Photochem. Photobiol. **29**:311-323.

Smith, A. B., B. Lafay, and R. Christen. 1992. Comparative variation of morphological and molecular evolution through geologic time: 28S ribosomal RNA versus morphology in echinoids. Phil. Trans. Roy Soc. London **338**:365-382.

Smith, A. B., and K. J. Peterson. 2002. Dating the time of origin of major clades: molecular clocks and the fossil record. Annu. Rev. Earth Planet. Sci. **30**:65-88.

Smith, R., B. Prezelin, K. Baker, R. Bidigare, N. Boucher, T. Coley, D. Karentz, S. MacIntyre, H. Matlick, D. Menzies, M. Ondrusck, Z. Wan, and K. Waters. 1992. Ozone depletion: ultraviolet radiation and phytoplankton biology in Antarctic waters. Science **255**:952-959.

Spears, T., and L. G. Abele. 1997. Crustacean phylogeny inferred from 18S rDNA. Pp. 169-187 *in* Arthropod relationships. R. A. Fortey and R. H. Thomas, eds. Systematics Association Special Volume Series 55. Chapman & Hall, London.

Spears, T., and L. G. Abele. 2000. Branchiopod monophyly and interordinal phylogeny inferred from 18S ribosomal DNA. J. Crust. Biol. **20**:1-24.

Stach, T., and C. Turbeville. 2002. Tunicate phylogeny inferred from morphological and molecular characters. Mol. Phylogenet. Evol. **25**:408-428.

Strid, A., W. S. Chow, and J. M. Anderson. 1994. UV-B damage and protection at the molecular level in plants. Photosynth. Res. **39**:475-489.

Swain, T. D., and D. J. Taylor. 2003. Structural rRNA characters support monophyly of raptorial limbs and paraphyly of limb specialization in water fleas. Proc. R. Soc. Lond. B Biol. Sci. **270**:887-896.

Swofford, D. L. 2001. PAUP* . Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Takezaki, N., A. Rzhetsky, and M. Nei. 1995. Phylogenetic test of molecular clock and linearized trees. Mol. Biol. Evol. **12**:823-833.

Tasch, P. 1969. Branchiopoda. Pp. 128-191 in R. C. Moore (ed.). Treatise on invertebrate palaeontology, part R: Arthropoda 4, vol. 1. University of Kansas Press, Lawrence, Kansas.

Taylor, D. J., T. J. Crease, and W. M. Brown. 1999. Phylogenetic evidence for a single longlived clade of crustacean cyclic parthenogens and its implications for the evolution of sex. Proc. R. Soc. Lond. B Biol. Sci. **266**:791-797.

Taylor, D. J., T. L. Finston, and P. D. N. Hebert. 1998. Biogeography of a widespread freshwater crustacean: pseudocongruence and cryptic endemism in the North American *Daphnia laevis* complex. Evolution **52**:1648-1670.

Tevini, M. 1993. UV-B radiation and ozone depletion: Effects on humans, animals, plants, microorganisms, and materials. Lewis Publishers, Boca Raton, FL.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**:4876-4882.

Trudgill D. L., and Blok V. C. 2001. Apomictic, polyphagous root-knot nematodes: exceptionally successful and damaging biotrophic root pathogens. Annu. Rev. Phytopathol. **39**:53-77.

Urwin R., E. C. Holmes, A. J. Fox, J. P. Derrick, and M. C. Maiden. 2002. Phylogenetic evidence for frequent positive selection and recombination in the meningococcal surface antigen PorB. Mol Biol Evol. **19**:1686-1694.

Van de Peer, Y., P. De Rijk, J. Wuyts, T. Winkelmans, and R. De Wachter. 2000. The European small subunit ribosomal RNA database. Nucleic Acids Res. **28**:175-176.

Van Valen, L. 1973. A new evolutionary law. Evol. Theory **1**:1-30.

Vawter, L., and W. Brown. 1986. Nuclear and mitochondrial comparisons reveal extreme rate variation in the molecular clock. Science **234**:194-196.

Wada, H. 1998. Evolutionary history of free-swimming and sessile lifestyles in urochordates as deduced from 18S rDNA molecular phylogeny. Mol. Biol. Evol. **15**:1189-1194.

Wagner J. R., C. C. Hu, and B. N. Ames. 1992. Endogenous oxidative damage of deoxycytidine in DNA. Proc. Natl. Acad. Sci. USA **89**:3380-3384.

Walossek, D. 1993. The Upper Cambrian *Rehbachiella* and the phylogeny of Branchiopoda and Crustacea. Fossils and Strata **32**:1-202.

Walossek, D., 1995. The Upper Cambrian *Rehbachiella*, its larval development, morphology and significance for the phylogeny of Branchiopoda and Crustacea. – Hydrobiol. **298**:1-13.

Watanabe, K. I., Y. Bessho, M. Kawasaki, and H. Hori. 1999. Mitochondrial genes are found on minicircle DNA molecules in the mesozoan animal *Dicyema*. J. Mol. Biol. **286**:645-650.

Waters, E., M. J. Hohn, I. Ahel, D. E. Graham, M. D. Adams, et al. 2003. The genome of *Nanoarchaeum equitans*: Insights into early archaeal evolution and derived parasitism. Proc. Natl. Acad. Sci. U. S. A. **100**:12984-12988.

Webster, A. J., R. J. H. Payne, and M. Pagel. 2003. Molecular phylogenies link rates of evolution and speciation. Science **301**:478.

Wernegreen, J.J. 2002. Genome evolution in bacterial endosymbionts of insects. Nat. Rev. Genet. 3:850-861.

Wiegmann, B.M., C. Mitter, and F.C. Thompson. 1993. Evolutionary origin of the Cyclorrhapha (Diptera): tests of alternative morphological hypotheses. Cladistics **9**:41-81.

Weiler, C.S. and P.A. Penhale. 1994. Ultraviolet radiation in Antarctica: Measurements and biological effects. American Geophysical Union Antarctic Research Series Vol. 62, 257 pp.

Wingstrand, K. G. 1978. Comparative spermatology of the Crustacea Entomostraca. I. Subclass Branchiopoda. Kolgl. Danske. Vid Selskab Biol. Skr. **22**:1-66.

Witkin E.M. 1976. Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. Bacteriol. Rev. **40**:869-907.

Woolfit, M., and L. Bromham. 2003. Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. Mol. Biol. Evol. **20**:1545-1555.

Wright, S. 1951. The genetical structure of populations. Ann. Eugenics **15**:323-354.

Wu, C. I., and W.H. Li. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. USA **82**:1741-1745.

Xia, X., and Z. Xie. 2001. DAMBE: Data analysis in molecular biology and evolution. J. Hered. **92**:371-373.

Xenopoulos, M. A., and D. W. Schindler. 2001. Physical factors determining ultraviolet radiation flux into ecosystems. Pp 36-62 *in* CS Cockell and AR Blaustein eds. Ecosystems, evolution and ultraviolet radiation. Springer-Verlag, New York.

Yoshizawa, K., and K. P. Johnson. 2003. Phylogenetic position of Phthiraptera (Insecta: Paraneoptera) and elevated rate of evolution in mitochondrial 12S and 16S rDNA. Mol. Phylogenet. Evol. **29**:102-114.

Ziegler, A., D.J. Leffell, S. Kunala, H.W. Sharma, M. Gailani, J.A. Simon, A.J. Halperin, H.P. Baden, P.E. Shapiro, A.E. Bale, and D.E. Brash. 1993. Mutation hot spots due to sunlight in the p53 gene of nonmelanoma skin cancers. Proc. Natl. Acad. Sci. USA **90**:4216–4220.

Zuckerkandl, E., and L. Pauling. 1962, Molecular disease, evolution, and genetic heterogeneity: Pp. 189–225 in M. Kasha and B. Pullman, eds. Horizons in biochemistry. Academic Press, New York.

Zuckerkandl, E., and L. Pauling. 1965. Evolutionary divergence and convergence in proteins. Pp. 97-166 in V. Bryson, and H. J. Vogel (eds). Evolving genes and proteins. Academic Press, New York.

Zwickl, D. J., and D. M. Hillis. 2002. Increased taxon sampling greatly reduces phylogenetic error. Syst. Biol. **51**:588-598.

Appendix 1. Taxonomic sample and sources of sequence data for Chapter 2.

Taxonomic group	Species	Description	Genbank Accession Nos.	
			COI	18S
1- Cnidaria				
Hydromedusae	*Carybdea*	Marine/pelagic	-	-
	Porpita		*	AF358086
	Physalia physalis		*	AF358065
	Haliclystus		-	-
	Cyanea		*	-
	Aurelia		AY319474	AY039208
	Eugymnanthea japonica		-	-
Anthozoa	*Acropora*	Marine/sedentary	NC_003522	-
	Metridium		NC_000933	AF052889
	Sarcophyton		AF064823	-
	Xenia		-	AF052931
	Pocillopora		AY139813	-
	Anemonia		-	X53498
	Outgroup: *Beroe*	Marine/planktonic	-	AF293696
2- Opisthobranchia				
Thecosomata	*Limacina*	Marine/planktonic	*	-
	Creseis chierchae		*	-
	Creseis acicula		*	-
Gymnosomata	*Clione*		*	*
Aplysiidae	*Aplysia*	Marine/benthic	AF343432	AJ224918
	Dolabella auricularia		AF156148	-
	Notarchus indicus		AF156151	-
	Stylocheilus longicauda		AF156156	-
	Outgroup: *Pupa/ Haminoea*		AB028237	AF249221
3- Branchiopoda				
Daphniidae	*Scapholeberis*	FW/planktonic	*	*
	Ceriodaphnia		*	*
	Daphnia		*	*
Chydoridae/ Macrothricidae	*Chydorus*	FW/benthic	*	*
	Eurycercus		*	*
	Ilyocryptus		*	*
	Outgroup: *Cyclestheria*		*	*

Taxonomic group	Species	Description	Genbank Accession Nos.	
			COI	18S
4- Amphipoda				
Hyperiidea	*Themisto gaudichaudii*	Marine/planktonic	*	-
	Themisto libellula		*	-
	Hyperia galba		*	-
	Hyperiella		*	-
Gammaridea	*Weyprechtia*	Marine/benthic	*	-
	Rhachotropis aculeata		AY271853	-
	Gammarus lacustrus		*	-
	Diporeia hoyi		*	-
	Outgroup: *Aeginina*		*	-
5- Eucarida				
Sergestoidea	*Lucifer*	Marine/planktonic	*	-
Penaeoidea	*Penaeus*	Marine/benthic	NC_002184	-
	Outgroup:*Homarus americanus*		*	-
6- Peracarida				
Euphausiacea	*Euphausia*	Marine/pelagic	*	AY141010
	Thysanoessa		*	AY141011
	Nyctiphanes australis		*	-
Decapoda	*Jaxea nocturna*		AF436030	AF436006
	Lithodes aequispinus		AF425308	AF439385
	Homarus americanus		*	-
	Outgroup: *Cumacea*		AF520450	Z22519
7- Ostracoda				
Myodocopa	*Euconchoecia*	Marine/planktonic	*	AF363296
	Skogsbergia		*	AF363297
	Melavargula japonica		-	AF363300
Podocopa	*Cypridopsis*	FW/benthic	*	AF363307
	Mytilocypris ambiguosa		*	-
	Eurycypris pubera		-	*
	Cyprinotus incongruens		-	*
	Outgroup: *Chaemosipho/Balanus*		*	AF201663

Taxonomic group	Species	Description	Genbank Accession Nos.	
			COI	18S
8- Urochordata				
Appendicularia	*Oikopleura dioica*	Marine/planktonic	AY116609	AB013014
	Oikopleura sp.		-	AB013015
Ascidiacea	*Ciona intestinalis*	Marine/sedentary	NC_004447	AB013017
	Halocynthia roretzi		-	AB013016
	Outgroup: *Branchiostoma*		NC_000834	M97571

* denotes sequences generated in the present study

Appendix 2. Taxonomic sample and sources of sequence data for Chapter 3.

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
1- Cnidaria					
Narcomedusae	*Polypodium hydriforme*	-	-	-	U37526
	Solmissus marshalli	-	-	-	AF358060
Rhopalonematidae	*Crossota rufobrunnea*	-	-	-	AF358063
	Pantachogon haeckeli	-	-	-	AF358062
	Outgroup: *Maeotias*	-	-	-	AF358056
2- Nematoda clade 1					
Trichocephalida	*Trichinella spiralis*	-	-	-	U60231
	Trichuris muris	-	-	-	AF036637
Mononchida	*Mylonchulus arenicolus*	-	-	-	AF036596
	Clarkus papillatus	-	-	-	AY146547
	Outgroup: *Caenorhabditis*	-	-	-	AY268117
3- Nematoda clade 2					
Triplonchida	*Trichodorus primitivus*	-	-	-	AF036609
	Paratrichodorus anemones	-	-	-	AF036600
	Paratrichodorus pachydermus	-	-	-	AF036601
	Paratrichodorus minor	-	-	-	AJ438052

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
Enoplida	*Adoncholaimus* sp.	-	-	-	AF036642
	Prismatolaimus intermedius	-	-	-	AF036603
	Enoplus meridionalis	-	-	-	Y16914
	Pontonema vulgare	-	-	-	AF047890
	Outgroup: *Caenorhabditis*	-	-	-	AY268117
4- Nematoda clade 4a					
Strongyloididae	*Strongyloides stercoralis*	-	-	-	AF279916
	Strongyloides ratti	-	-	-	U81581
	Strongyloides cebus	-	-	-	AJ417025
	Parastrongyloides trichosuri	-	-	-	AJ417024
Panagrolaimidae	*Panagrolaimus* sp.	-	-	-	U81579
	Panagrellus redivivus	-	-	-	AF036599
	Plectonchus sp.	-	-	-	AF202154
	Turbatrix aceti	-	-	-	AF202165
	Outgroup: *Caenorhabditis*	-	-	-	AY268117
5- Nematoda clade 4b					
Tylenchida	*Meloidogyne incognita*	-	-	-	AY268120
	Meloidogyne javanica	-	-	-	AY268121
	Meloidogyne arenaria	-	-	-	U42342
	Heterodera glycines	-	-	-	AY043247
Cephaloboidea	*Zeldia punctata*	-	-	-	U61760
	Acrobeles ciliatus	-	-	-	AF202148
	Acrobeloides sp.	-	-	-	AF034391
	Cephaloboides sp.	-	-	-	AF083027

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
	Outgroup: *Caenorhabditis*	-	-	-	AY268117
6- Nematoda clade 5					
Strongylida	*Necator americanus*	NC_003416	NC_003416	-	-
	Haemonchus contortus	HCC00291[a]	-	-	-
	Ancylostoma caninum	AC01024[b]	-	-	-
	Ostertagia ostertagi	-	-	-	AF036598
	Parafilaroides sp.	-	-	-	U81590
	Otostrongylus sp.	-	-	-	U81589
Rhabditoidea	*Caenorhabditis elegans*	NC_001328	NC_001328	-	AY268117
	Caenorhabditis briggsae	c009001201[c]	-	-	U13929
	Pristionchus pacificus	PP00640[b]	-	-	AF083010
	Outgroup: *Drosophila/Adoncholaimus/Zeldia*	NC_001709	AF317080	-	M21017
7- Syndermata					
Acanthocephala	*Moniliformis moniliformis*	AF416998	AF108107	-	Z19562
	Oncicola sp.	AF417000	AF325127	-	AF064818
	Oligacanthorhynchus tortuosa	AF416999	AF325128	-	AF064817
	Macracanthorhynchus ingens	AF416997	AF325129	-	AF001844
Rotifera	*Synchaeta* cf. *pectinata*	AF499088	AF499048	-	-
	Keratella hiemalis	AF499077	AF499045	-	-
	Brachionus plicatilis	AF499054	AF364510	-	U49911
	Asplanchna sp.	AF499052	AF499036	-	AF092434
	Lecane bulla	-	-	-	AF154566
	Philodina roseola	-	-	-	AF154567
	Outgroup: *Pseudostylochus*	AB049114	AB049114	-	D17563

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
8- Cirripedia					
Rhizocephala	*Sacculina*	AY117692	-	-	AY265367
	Sylon	PABA2	PABA2	-	-
	Polyascus gregaria	-	-	-	AY265363
	Bosmaella japonica	-			AY265369
	Heterosaccus	AY117691	-	-	AY265359
Thoracica	*Balanus glandula*	-	-	-	AF201663
	Semibalanus balanoides	AF242699	-	-	-
	Lepas	-	PABA5	-	L26516
	Chthamalus	AF234819	-	-	L26515
	cf Chamaesipho	MxCht000	-	-	-
	Verruca spengleri	-	-	-	AF022230
	Outgroup: *Calanus/Eurytemora*	MxCal000	AF315006	-	AF367719
9- Maxillopoda					
Branchiura	*Argulus*	-	-	-	M27187
	Dolops	MxArg001	-	-	-
Ostracoda	*Cypridopsis*	MxCyp000	-	-	AF363307
	Outgroup: *Calanus*	MxCal000	-	-	AF367719
10- Copepoda					
Siphonostomatoida/ Poecilostomatoida	*Lepeophtheirus salmonis*	PACO2	-	-	PACO2
	Unidentified (002)	MxPc?002	MxPc?002	-	MxPc?002
	Unidentified (000)	MxPc?000	-	-	-
	Stellicola	MxLch000	-	-	-

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
Harpacticoida	*Tigriopus japonicus*	NC_003979	NC_003979	-	AF363306
	Cletocamptus helobius	AF315014	-	-	-
	Coullana	AF315015	-	-	-
	Cancrincola plumipes	-	-	-	L81938
	Outgroup: *Calanus*	MxCal000	AF315006	-	AF367719
11- Isopoda					
Bopyridae/Cymothoidae	*Bopyroides hippolytes*	MaBop000	PAIS1	-	-
	Olencira praegustator	AF255791	AF259547	-	-
	Lironeca vulgaris	AF260843	AF260852	-	-
	Rocinella angustata	MaAgd000	-	-	-
	Hemiarthrus abdominalis	-	-	-	AF255684
	Probopyrus pacificiensis	-	-	-	AF255683
	Anilocra physodes	-	-	-	AF255686
	Riggia paranensis	-	-	-	AF255685
Sphaeromatidae	*Exosphaeroma* sp.	MaSpm000	-	-	-
	Sphaeramene polytylotus	AF260846	AF259540	-	-
	Gnorimosphaeroma oregonense	AF260845	AF260866	-	-
	Sphaeroma	AF255785	AJ388074	-	AF255694
	Cymodoce tattersalli	-	-	-	AF255695
	Lekanesphaera hookeri	-	-	-	AF279600
	Campecopea hirsute	-	-	-	AF279601
	Outgroup: *Gammarus*	MaGam000	AF228046	-	AF202982

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
12- Diptera 1					
Oestroidea	*Peleteria biangulata*	-	-	AF364388	-
	Chrysomya chloropyga	NC_002697	NC_002697	-	-
	Cochliomyia hominivorax	NC_002660	NC_002660	-	-
	Hypoderma	AF257116	AF322438	-	AF322427
	Dexiinae	AF456872	-	AF364378	-
	Nemoraea pellucida	-	-	-	AF322418
Muscoidea	*Musca* sp.	-	AF322436	AF503149	AF322423
	Delia	AF325362	AF086867	-	-
	Fannia armata	AF104623	-	-	-
	Fucellia maritima	-	-	AF508545	-
	Scathophaga stercoraria	-	PAIN4/Test#4	-	PAIN4/Test#4
	Lasiomma seminitidum	AF104624	-	-	-
	Chiastocheta pellmyri	AF358428	-	-	-
	Outgroup: *Drosophila*	NC_001709	NC_001709	NT_033778	M21017
13- Diptera 2					
Nycteribiidae	*Penicillidia*	-	AF322435	-	AF322420
Glossinidae	*Glossina*	-	AF072373	-	AF322431
	Outgroup: *Drosophila*	-	NC_001709	-	M21017
14- Diptera 3					
Pipunculidae	*Jassidophaga*	AY261685	AF154786	-	-
	Chalarus	-	AF154756	-	-
Syrphidae	*Eristalis tenax*	AY261703	AY123344	-	-
	Paragus	-	AF154800	-	-

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
	Outgroup: *Drosophila*	NC_001709	NC_001709	-	-
15- Diptera 4					
Conopoidea	*Physocephala sagittaria*	-	AY123349	-	-
	Neoconops longicornis	-	AF154768	-	-
Tephritoidea	*Ceratitis capitata*	-	NC_000857	-	-
	Anastrepha ludens	-	AB035102	-	-
	Outgroup: *Drosophila*	-	NC_001709	-	-
16- Diptera 5					
Bombyliidae	*Heterotropus*	-	-	AF148082	-
	Comptosia	AF456871	AF456847	-	-
	Villa	AF456869	AF456846	-	-
	Anthrax	-	-	-	PAIN8/Test#18
Therevidae	*Ectinorhynchus*	AF456868	AF456856	AF148098	-
	Cerdistus	AF456874	AF456850	-	-
	Laphria	-	AY123342	-	AF286293
	Outgroup: *Drosophila*	NC_001709	NC_001709	NT_033778	M21017
17- Hymenoptera					
Apocrita	*Xorides*	-	AF003520	-	-
	Apis mellifera	NC_001566	NC_001566	-	-
	Snellenius	AF102701	-	-	-
Tenthredinoidea	*Perga condei*	U16713	U06953	-	-
	Phylacteophaga froggatti	U16714	U06954	-	-

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
	Outgroup: *Drosophila*	NC_001709	NC_001709	-	-
18- Psocodea					
Phthiraptera	*Heterodoxus macropus*	NC_002651	NC_002651	-	-
	Brueelia sp.	-	-	-	AF385039
	Haematomyzus elephantis	AF385009	-	AF320405	AY077778
	Hoplopleura sp.	AF545717	-	-	AY077773
	Bovicola sp.	AF545680	-	-	AY077769
Psocoptera	*Lepidopsocid*	NC_004816	NC_004816	-	-
	Echmepteryx hageni	AY275298	-	AY314832	AY275348
	Valenzuela sp.	AY275283	-	-	AF423793
	Compsocus elegans	AY275296	-	-	AY275346
	Outgroup: *Triatoma/Gerris*	NC_002609	NC_002609	AF251080	AJ243328
19- Platyhelminthes					
Neodermata	*Schistosoma mansoni*	NC_002545	NC_002545	Y08487	M62652
	Echinococcus	NC_000928	-	AF093621	U27015
	Fasciola hepatica	NC_002546	-	-	AJ004969
	Hymenolepis diminuta	NC_002767	-	AF124802	AF124475
	Neomicrocotyle pacifica	-	-	AF288070	-
	Haplobothrium globuliforme			AF124798	
Turbellaria	*Pseudostylochus*	AB049114	AB049114	-	D17563
	Bipalium adventitium	AF178306	-	-	X91402
	Dugesia japonica	D49916	-	D49924	AF013153
	Mesostoma	AJ405988	-	AF288069	AJ270157
	Schmidtea polychroa	-	-	AJ250914	-
	Prostheceraeus vittatus	-	-	AJ250909	-

Taxonomic group	Species	Genbank accession numbers / Sample codes			
		COI	16S	EF-1a	18S
	Outgroup: *Platynereis/Nereis*	NC_000931	NC_000931	U90064	Z83754
20- Annelida					
Euhirudinea	*Hirudo*	AF003272	AF315058	U90063	Z83752
	Placobdella parasitica	AF003261	-	AF005502	AF115990
	Haemadipsa	-	-	-	AF099944
	Haemopis sanguisuga	-	-	-	X91401
Lumbricidae/ Enchytraeidae	*Lumbricus terrestris*	NC_001673	NC_001673	AF005501	AJ272183
	Lumbricus rubellus	-	-	-	Z83753
	Enchytraeus	-	-	AF063418	Z83750
	Eisenia fetida	-	-	-	X79872
	Fridericia tuberose	AF064047	-	-	-
	Outgroup: *Platynereis/Nereis*	NC_000931	NC_000931	U90064	Z83754
21- Bivalvia					
Unionoida	*Lampsilis cardium*	AF120653	-	-	AF120537
Trigonioida	*Neotrigonia bednalli*	U56850	-	-	AF120538
	Outgroup: *Dreissena*	AF120663	-	-	AF120552

[a] cluster retrieved from NemBase (http://www.nematodes.org)
[b] cluster retrieved from NemaGene (http://www.nematode.net)
[c] contig retrieved from WormBase (http://www.wormbase.org)