THE UNIVERSITY OF HULL

The genetics and evolutionary dynamics of sexual system evolution in tadpole shrimps

being a Thesis submitted for the Degree of Ph.D.

in the University of Hull

by

Thomas Charles Mathers, BSc (Hons)

September, 2013

Acknowledgements

This study was funded by a NERC CASE Ph.D. Studentship (NE/G012318/1) in partnership with the Natural History Museum, London.

Diego Fontaneto, Don Dasis, Marco Seminara, Dani Boix, Anna Badosa, Andy Green and Kennet Lundin all kindly provided samples. Thorid Zierold, in addition to providing advice on the rearing, sexing and taxonomy of tadpole shrimps was of considerable help when sampling in Germany, organising access to several key sites. Rich Gill assisted with sampling in Germany, providing moral support, especially in Usingen.

RAD library construction and sequencing was performed by The GenePool, Edinburgh and funded with an NBAF small projects grant (NBAF400).

I would like to thank my supervisors Africa Gómez, Rob Hammond, Bernd Hänfling and Ronald Jenner for their support, encouragement and enthusiasm. I also owe Rob a personal debt of gratitude for inspiring me to study evolution.

Thanks to the Hull Evolutionary Biology Group members for all the coffee times. In particular I would like to thank Steve Moss who helped throughout with bioinformatics support and Dan for always saying yes to the pub.

Last but not least I would like to thank my wife Hannah for supporting me throughout.

Contents

Abstract 1

Chapter 1: General introduction

1.1 Evolutionary transitions between sexual systems 2

1.2 Sex chromosome evolution 5

1.3 Thesis aims and outline 7

1.4 Tadpole shrimps – a model to study the evolution of sexual systems in animals 8

1.5 Molecular divergence dating 10

1.6 Using molecular phylogenetic trees to infer character evolution 12

1.7 Restriction site associated DNA sequencing 13

Chapter 2: Multiple global radiations in tadpole shrimps challenge the concept of 'living fossils'

2.1 Abstract 16

2.2 Introduction 16

2.3 Materials and methods 18

2.4 Results and discussion 22

Chapter 3: High lability of sexual system over 250 million years of evolution in morphologically conservative tadpole shrimps

3.1 Abstract 26

3.2 Introduction 27

3.3 Materials and methods 29

3.4 Results 34

3.5 Discussion 42

Chapter 4: RAD sequencing reveals chromosomal ZW sex determination in androdioecious and gonochoric populations of *Triops cancriformis*

4.1 Abstract 47
4.2 Introduction 48
4.3 Materials and methods 51
4.4 Results 55
4.5 Discussion 73

Chapter 5: General discussion and concluding remarks

5.1 Tadpole shrimp phylogeny 80
5.2 The timing and tempo of tadpole shrimp diversification 81
5.3 Labile sexual systems 81
5.4 Sex determination and sex chromosome evolution 82
5.5 Concluding remarks 83

References 84

Glossary 111

Appendix 1: Additional information for Chapter 2 112

Appendix 2: Additional information for Chapter 3 127

Appendix 3: Additional analysis for Chapter 3 145

Appendix 4: Additional information for Chapter 4 149

Appendix 5: Consensus sequences of sex-linked RAD loci identified in Chapter 4 154

"In an old stereotype (not followed nearly so often as mythology proclaims), the natural history essay restricts itself to describing the peculiarities of animals – the mysterious ways of the beaver, or how the spider weaves her subtle web. There is exultation in this and who shall gainsay it? But each organism can mean so much more to us. Each instructs; its form and behaviour embodies general messages if only we can learn to read them. The language of this instruction is evolutionary theory. Exultation and explanation."

(S. J. Gould, 1980[1])

[1]Gould, S. J., (1980) *The Panda's Thumb - More Reflections in Natural History*. W. W. Norton and Co.

Abstract

Sexual reproduction is found throughout the eukaryote tree of life and fundamentally affects the organisms that practice it. In particular, the sexual system of an organism can influence genetic diversity, population structure, genome structure, the evolutionary potential of species and even cause speciation. Understanding the effects of sexual reproduction, and the effects of transitions between its various forms, is therefore a central theme in biology and essential to understanding how populations and species evolve.

In this thesis I investigate the genetics and evolutionary dynamics of sexual system evolution in tadpole shrimps (Notostraca), a group of branchiopod crustaceans with diverse sexual systems. Gonochorism (dioecy), hermaphroditism and androdioecy – the presence of self-fertile hermaphrodites and males within a population – are all found, with this variation occurring both within and between species. In contrast to their variable sexual system, tadpole shrimps exhibit high levels of morphological conservatism, appearing to have changed little since the Early Devonian.

I establish the first resolved phylogeny of Notostraca and find that tadpole shrimps have undergone at least two bouts of diversification in their evolutionary history, with extant species being younger than the fossil record would suggest. Analysis of sexual system across the phylogeny reveals the labile nature of sexual systems within Notostraca with multiple transitions having occurred between separate and combined sexes, driven by selection for reproductive assurance. Finally I use restriction site associated DNA sequencing (RAD-seq) to identify over 1000 high coverage, novel genomic markers for *Triops cancriformis*, a mixed mating tadpole shrimp. I confirm a Z W chromosomal sex determination system for *T. cancriformis*, conserved in gonochoric and androdioecious populations and identify significant alterations in the structure of the W and Z chromosomes associated with the transition in sexual system. The results presented within this thesis establish tadpole shrimps as an ideal model to study the evolution of sexual systems and the genomic effects of repeated transitions between sexual systems.

Chapter 1

General introduction

Sex is important. The majority of eukaryotes do it and phylogenetic evidence suggests that meiosis and sexual reproduction evolved early during eukaryote evolution (Dacks & Roger 1999; Ramesh et al. 2005). Furthermore, sex has been maintained in these lineages despite its costs compared to asexual reproduction (Lehtonen et al. 2012; Maynard Smith 1978; Williams 1975). As such, sexual reproduction pervades the life histories of many of the organisms found on earth and its effects influence many aspects of their biology including behaviour, morphology, genetic diversity and genome evolution. For example, in systems with differentiated sexes, sexual selection can act to drive the evolution of costly secondary sexual characters (reviewed by Clutton-Brock 2007) and sexually antagonistic polymorphisms – beneficial to one sex but harmful to the other – are predicted to shape genome evolution and drive the evolution of sex chromosomes (Charlesworth & Charlesworth 1980; Charlesworth et al. 2005; Jordan & Charlesworth 2012; Rice 1987a; van Doorn & Kirkpatrick 2010). In systems with combined sexes, sexually antagonistic selection is reduced, but self-fertilisation, which often accompanies simultaneous hermaphroditism (Jarne & Auld 2006), can reduce genetic diversity (within populations), effective population size and the efficacy of selection (Charlesworth & Wright 2001; Glémin 2007). Understanding the effects of sexual reproduction, and the effects of transitions between its various forms, are therefore a central theme in biology and essential to understanding how populations and species evolve.

1.1 Evolutionary transitions between sexual systems

Hermaphroditism, where an organism is able to produce both male and female gametes, is common amongst plants and animals and transitions between separate and combined sexes have occurred many times (Barrett 2002; Barrett 2010; Bawa 1980; Charnov 1982; Charnov et al. 1976; Darwin 1876; Jarne & Auld 2006). These transitions are not predicted to occur in a single step. Instead, they are thought to go

through an intermediate stage, either in the form of androdioecy, where males coexist with hermaphrodites, or gynodioecy, where females coexist with hermaphrodites, as shown in Figure 1.1 (Charlesworth & Charlesworth 1978). Additionally, in some cases trioecy may arise where males, females and hermaphrodites are all found; although this sexual system is predicted to be unstable (Wolf & Takebayashi 2004) and is rarely found (Weeks 2012).

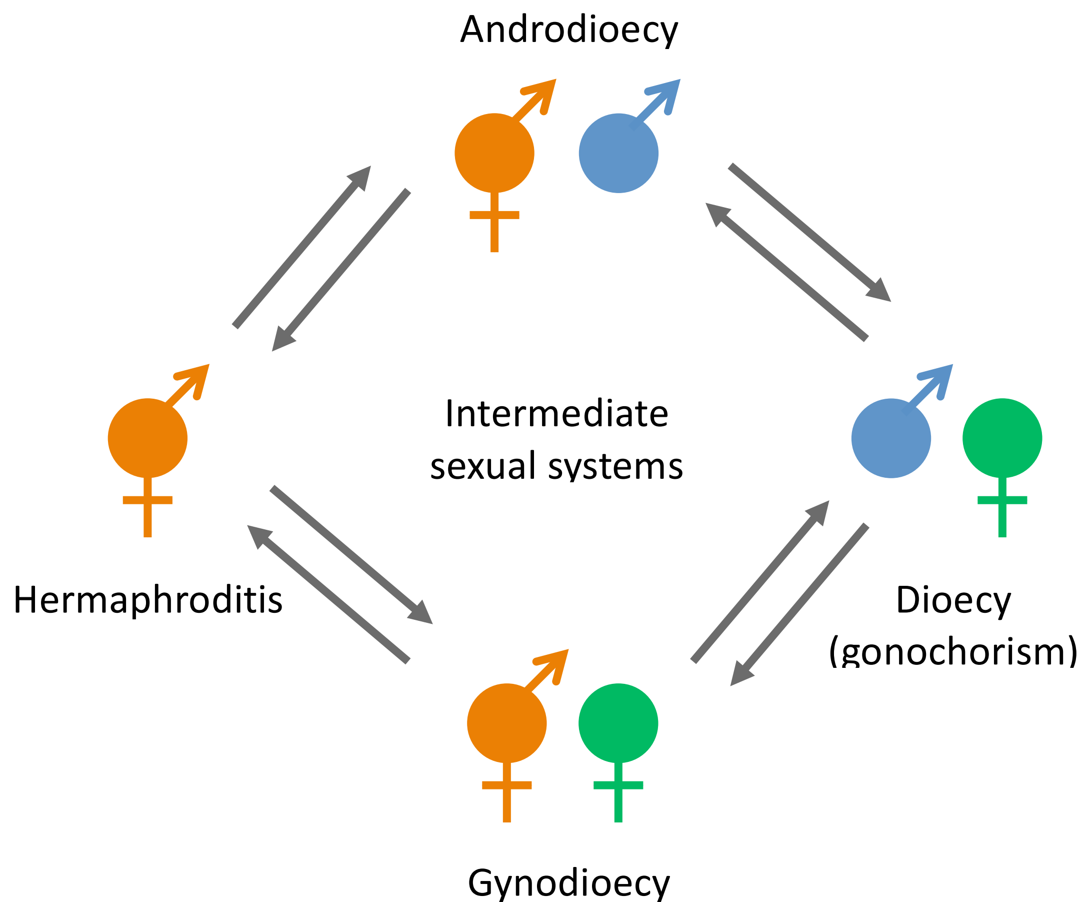


Figure 1.1: Possible evolutionary pathways between combined and separate sexes or vice versa (after Weeks et al. 2006a). The two most common intermediate sexual systems are shown; gynodioecy (females and hermaphrodites) and androdioecy (males and hermaphrodites).

The stability of intermediate sexual systems, especially androdioecy, has been the subject of much debate (e.g., Charlesworth 1984) due to the relative rarity of these systems and theoretical predictions that they are difficult to evolve – unisexuals are required to make more than twice the genetic contribution to offspring as hermaphrodites to spread in an outcrossing population (Charlesworth & Charlesworth 1978; Charnov et al. 1976; Lloyd 1975; Ross & Weir 1976). Androdioecy has only been

confirmed in a handful of plant and animal species (Pannell 2002a; Weeks 2012; Weeks et al. 2006a) and, whilst gynodioecy is more common in plants (Webb 1999), it is still relatively rare compared to the frequency of hermaphroditism and dioecy (Charlesworth 2006; Lloyd 1975) and considerably rarer than androdioecy in animals (Weeks 2012). Under certain conditions, however, both androdioecy and gynodioecy have been shown to be stable, although expectations for their evolution and maintenance are different in plants and animals (Otto et al. 1993; Pannell 1997; Pannell 2002a; Pannell 2008; Weeks 2012).

In plants, the most common transitions in sexual system are from hermaphroditism to dioecy, or from self-incompatible hermaphroditism to self-compatible hermaphroditism (i.e., from outcrossing to selfing) through the breakdown of self-incompatibility mechanisms (Charlesworth 2006; Jarne & Charlesworth 1993; Pannell 2002a; Weiblen et al. 2000). Models for the evolution of gynodioecy and androdioecy in plants have therefore generally considered transitions from a hermaphroditic progenitor. In these situations gynodioecy is thought to be more likely to evolve than androdioecy because, where selfing occurs, males suffer reduced mating opportunities as fewer ovules are available for fertilisation (Charlesworth & Charlesworth 1978; Lloyd 1975). Females on the other hand, do not suffer this cost in a gynodioecious population as their ovules can be fertilised by pollen from hermaphrodites. This situation is compounded in populations where selfing results in inbreeding depression as the relative fertility advantage required by females to invade the population may be reduced due to the fitness benefits of outcrossing compared to selfing, further increasing the chance of gynodioecy evolving over androdioecy (Sakai et al. 1997).

In animals, transitions between sexual systems have occurred in different circumstances to plants. In most cases hermaphroditism is thought to be a derived state (Eppley & Jesson 2008; Ghiselin 1969) and as such, transitions from dioecy to hermaphroditism are more common than from hermaphroditism to dioecy (Weeks 2012; Weeks et al. 2006a). Hermaphroditism may be selected for in dioecious species to provide reproductive assurance through self-fertilisation, particularly in situations where mates are scarce such as in fluctuating metapopulations with regular extinction and re-colonisation and during range expansions (Pannell 1997; Pannell 2000; Pannell 2002a; Pannell & Barrett 1998; Pannell & Dorken 2006). In sexually dimorphic animals

this is expected to lead to the preferential evolution of androdioecy over gynodioecy as an intermediate sexual system due to developmental constraints causing hermaphrodites to be strongly female biased, favouring the maintenance of male rather than female unisexuals (Weeks et al. 2006a; Weeks et al. 2009). This hypothesis assumes it is simpler to evolve sperm production in female gonads for self-fertilisation than to evolve the suite of secondary sexual characters (mating behaviours, sperm delivery etc.) associated with full male function and is supported by the much greater number of androdioecious animal species than gynodioecious ones (Weeks 2012). In cases where animals have evolved separate sexes from hermaphroditism, selection pressures are also thought to be different in animals compared to plants, leading to the evolution of androdioecy rather than gynodioecy as an intermediate strategy (Weeks 2012). Instead of separate sexes being selected for as a means to avoid inbreeding as in plants (Barrett 2010), unisexual individuals fill a unique ecological niche, as is the case for the evolution of dwarf complemental males in androdioecious pedunculate barnacles (Yusa et al. 2012).

1.2 Sex chromosome evolution

The evolution of sexual reproduction and the division of labour into separate sexes is considered to be one of the major evolutionary transitions (Szathmáry & Maynard Smith 1995) and it is often accompanied by the evolution of genetic sex determination and sex chromosomes. Sex chromosomes have evolved independently numerous times in plants and animals (Charlesworth 1996; Ellegren 2011; Fraser et al. 2004; Fraser & Heitman 2005; Vyskot & Hobza 2004) and are particularly interesting to biologists due to their role in sex determination and because they present an excellent opportunity to study a range of evolutionary processes across diverse taxa (Bachtrog et al. 2011). Fundamentally, sex chromosomes are linkage groups whose inheritance is associated with a particular sex (Bachtrog et al. 2011). The various forms of chromosomal sex determination found in nature are classified depending on which sex contains two types of sex chromosome (the heterogametic sex) and which sex contains one type of sex chromosome (the homogametic sex). Typically, chromosomal sex determination systems are either of the type XY where males are heterogametic or ZW where females are heterogametic (Graves 2008; Vyskot & Hobza 2004), although in some cases alternative systems arise such as the the XO XY system found in the rodent

Microtus oregoni (Charlesworth & Dempsey 2001) and the nematode *Caenorhabditis elegans* (reviewed by Meyer 2000).

Sex chromosomes originate from ordinary autosomes and evolve in a highly dynamic process (Fridolfsson et al. 1998; Graves 2008; Lahn & Page 1999; Pease & Hahn 2012; Skaletsky et al. 2003). This process has been reviewed in detail several times (e.g., Bachtrog 2006; Bachtrog et al. 2011; Bergero & Charlesworth 2009; Charlesworth et al. 2005; Livernois et al. 2011). Briefly, following the evolution of genetic sex determination, selection favours suppressed recombination in the sex determining region to prevent recombination between sex determining and sexually antagonistic loci. In the case of a transition from hermaphroditism to dioecy, this could be in the form of a male suppressing mutation that causes gynodioecy (females and hermaphrodites) followed by a female suppressing mutation in the remaining hermaphrodites leading to dioecy, or vice versa for an androdioecious pathway to dioecy. These sex determining genes are expected to be tightly linked. Initially, recombination is suppressed only in the sex determining region but subsequently cessation of recombination can spread across much of the sex specific chromosome, ultimately leading to morphologically distinct (heterogametic) pairs of chromosomes (reviewed by Bachtrog 2013). It is expected that sex specific chromosomes will become enriched for sexually antagonistic genes that are beneficial to the sex that carries them, promoting the evolution of reduced recombination (Charlesworth & Charlesworth 1980; Fisher 1931; Rice 1984; Rice 1987a). This can occur through inversions, transpositions and translocations of material from autosomes (Bachtrog 2013; Bergero & Charlesworth 2009; Charlesworth et al. 2005). In the absence of recombination the efficacy of selection is reduced and the sex specific chromosome degenerates, losing genes that do not have primary or secondary sex specific function and in many cases accumulating transposable elements (Bachtrog 2013; Graves 2006; Rice 1994). Degeneration of the sex specific chromosome is due to a combination of Muller's ratchet (Charlesworth & Charlesworth 1997) – the accumulation of deleterious mutations that cannot be purged due to the absence of recombination –, genetic hitchhiking (Rice 1987b) and the reduced ability of the sex specific chromosome to adapt (Bachtrog & Charlesworth 2002). Together these forces shape

the evolution of the sex specific chromosome producing a morphologically differentiated chromosome.

Sex chromosomes are not always heteromorphic with homomorphic sex chromosomes having evolved in many groups (Stöck et al. 2011). Sex chromosomes can be homomorphic due to a recent origin, meaning there has been insufficient time for degeneration of the sex specific chromosome, as in the incipient sex chromosome system found in the strawberry, *Fragaria virginiana* (Spigler et al. 2008; Spigler et al. 2010). Alternatively, occasional recombination between sex chromosomes, as in tree frogs (Guerrero et al. 2012; Stöck et al. 2011), or the evolution of sex biased gene expression, as in emus (Vicoso et al. 2013), has reduced genetic degeneration of the Y or W sex chromosome preventing morphological divergence. Studying these young or unusual sex chromosomes can expand our understanding of the evolutionary processes that take place during sex chromosome evolution.

1.3 Thesis aims and outline

In the research presented within this thesis I investigate evolutionary transitions in sexual system in tadpole shrimps, a group of branchiopod crustaceans with diverse sexual systems. Gonochorism (also known as dioecy), obligatory self-fertile hermaphroditism and the mixed mating system androdioecy are all found in tadpole shrimps (Longhurst 1955; Sassaman 1991; Sassaman 1997; Zierold *et al* 2007). Furthermore, this variation occurs both within and between species making tadpole shrimps an ideal system to study both the evolution of sexual systems and the influence of variable sexual systems on sex chromosome evolution. Currently, several key obstacles stand in the way of gaining a good understanding of sexual system evolution in the group:

1) There is no resolved phylogeny of the order (e.g., Vanschoenwinkel et al. 2012) and whilst fossil evidence suggests lineages are ancient (Barnard 1929; Fryer 1988; SunoUchi et al. 1997), the widespread presence of cryptic species (see section below on tadpole shrimps) and high levels of morphological conservatism within the order have hampered the establishment of a timeframe for tadpole shrimp evolution.

2) Although sexual system is known to be variable in tadpole shrimps (Sassaman 1991), the extent of this variability is unknown and has not been placed in a phylogenetic context.

3) With the exception of two expressed sequence tag (EST) libraries for the European tadpole shrimp *Triops cancriformis* (Meusemann et al. (2010); Ronald Jenner (Natural History Museum, London) and Matthew Wills (University of Bath), unpublished), there are no resources available to investigate the genomic impacts of transitions between sexual systems.

4) Whilst genetic sex determination has been established in *Triops,* with males being recessive to hermaphrodites (Sassaman 1989; Sassaman 1991), the specifics are unclear.

The key aim of this thesis is therefore to rectify these issues and develop our understanding of how sexual systems evolve. To this end I will use relaxed molecular clock divergence dating to establish the timing and tempo of diversification within tadpole shrimps (Chapter 2), comparative phylogenetic methods to investigate the evolution of sexual systems within the group (Chapter 3) and genomic approaches to elucidate the genetic basis of sex determination (Chapter 4) and the genomic effects of changes in sexual system in the mixed mating European tadpole shrimp *Triops cancriformis*.

In the remaining sections of this introduction I will provide an overview of the biology and evolution of tadpole shrimps and then introduce the main methodologies used within this thesis to meet the aims outlined above.

1.4 Tadpole shrimps – a model to study the evolution of sexual systems in animals

Notostraca, or tadpole shrimps as they are commonly known, are an ancient order of branchiopod crustaceans whose morphology has remained stable for millions of years, with fossils from as far back as the early Devonian bearing a striking resemblance to extant species (Barnard 1929; Fayers & Trewin 2002; Fryer 1988; SunoUchi et al. 1997). Tadpole shrimps are globally distributed, being found on every continent except for Antarctica and comprise of two extant genera, *Triops* and *Lepidurus*, distinguished by the presence of a distinctive supra-anal plate found in the latter but not the former

(Brendonck et al. 2008; Fryer 1988; Linder 1952; Longhurst 1955c). Both genera inhabit ephemeral freshwater to brackish habitats and survive periods of drought through highly resistant diapausing cysts (Brendonck 1996; Fryer 1996; Longhurst 1955c).

In 1955, Longhurst (1955c) reviewed Notostraca and identified 9 species. Subsequently, several cryptic species complexes have been uncovered (King & Hanner 1998; Korn et al. 2010; Korn & Hundsdoerfer 2006; Macdonald et al. 2011; Mantovani et al. 2009; Rogers 2001; Vanschoenwinkel et al. 2012) and it now appears that most of the species identified by Longhurst (1955c) are in fact endemic species complexes and the true taxonomic diversity of the order is uncertain. Additionally, the extent of cryptic species and high levels of morphological conservatism within the order have hampered attempts to infer a robust phylogeny of the order, with previous attempts being incomplete and unresolved (e.g., Mantovani et al. 2004; Vanschoenwinkel et al. 2012), leaving fundamental questions such as the monophyly of the two genera unknown.

Notostracan fossils are found as early as the Devonian (Fayers & Trewin 2002) and fossils attributed to the extant genera *Triops* and *Lepidurus* have been found as far back as the Triassic and Jurassic respectively (Barnard 1929; Gall & Grauvogel-Stamm 2005; Gore 1986; Haughton 1924; Trusheim 1938). This striking morphological conservatism has led them to be referred to as 'living fossils' (Fryer 1988; King & Hanner 1998; Mantovani et al. 2008), a controversial term coined by Darwin to describe species that have changed very little over long periods of time (Darwin 1859).

In contrast to their conserved morphology, tadpole shrimps display a diverse range of sexual systems, with variation being found both within and between species (Longhurst 1954; Longhurst 1955b; Longhurst 1955c; Macdonald et al. 2011; Sassaman 1991; Sassaman et al. 1997; Zierold et al. 2007; Zierold et al. 2009). This variation in sexual system includes unisexual populations with sexually dimorphic males and females found in an approximate 50:50 ratio (termed dioecy or gonochorism), obligatory selfing hermaphroditism and androdioecy. Significantly, within the European tadpole shrimp *Triops cancriformis*, all three of these sexual systems are found, with variation between populations, as can be seen in Figure 1.2 (Zierold et al. 2007; Zierold et al. 2009). This mixed mating system makes *T. cancriformis* - and the

Notostraca at large - an ideal system for studying the evolution of sexual systems and the effects that they have both at a population and genome level.

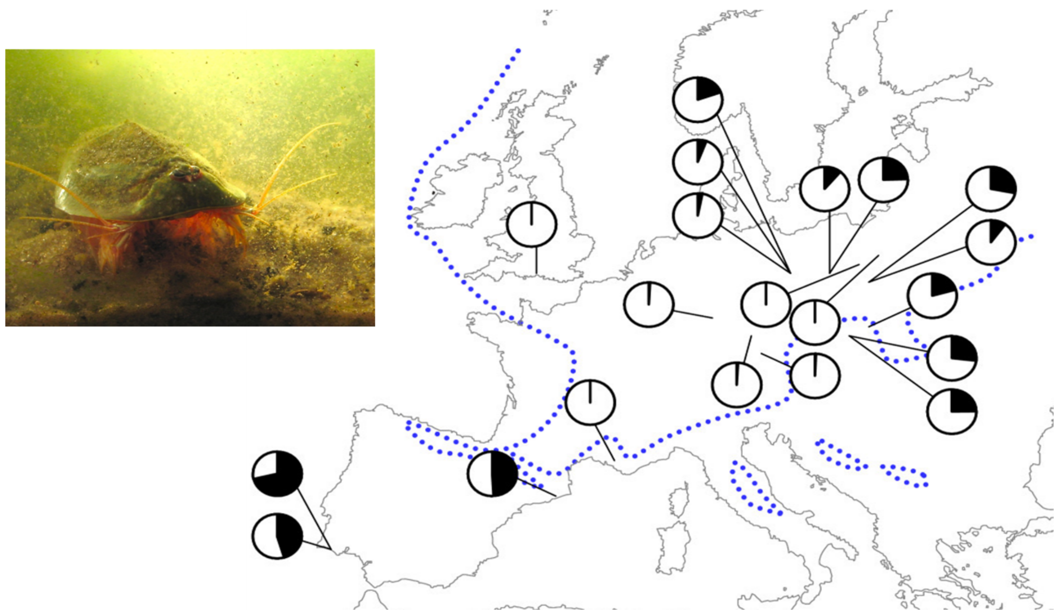


Figure 1.2: Sex ratio variation in the European tadpole shrimp *Triops cancriformis* (adapted from Zierold et al. 2007). Pie charts show male proportions (black sections) for populations of *T. cancriformis* sampled by Zierold et al. (2007). All sample sizes >25. Gonochoric (dioecious) populations are found in the Iberian peninsula with androdioecious and hermaphroditic populations found in central and Eastern Europe. Inlay shows photograph of adult *T. cancriformis* by Africa Gomez, used with permission.

1.5 Molecular divergence dating

The term 'molecular evolutionary clock' was first coined by Zuckerkandl and Pauling (1965) and refers to the hypothesis that molecular evolution occurs at an approximately uniform rate through time (Morgan 1998). This means that a timescale can be placed on evolutionary events by calibrating the molecular clock. Originally applied to protein evolution and later developed for DNA sequence data, molecular clocks are now widely used in evolutionary biology (reviewed by Bromham & Penny 2003) and were central to the development of the neutral theory of evolution (Kimura 1968; Kimura 1969; Kimura 1984; Kimura & Ohta 1971).

Molecular clocks are commonly used in phylogenetics to estimate divergence times of species. By measuring the sequence divergence between species, in combination with a model of sequence evolution and applying a calibration – either based on the fossil record or from geological events – divergence times can be estimated based on the

assumption of a molecular clock (Bromham & Penny 2003; Kumar 2005). This principle has been used to date important evolutionary events such as the diversification of metazoan phyla (Bromham et al. 1998; Peterson & Butterfield 2005; Wray et al. 1996), the radiation of placental mammals (Bininda-Emonds et al. 2007; dos Reis et al. 2012), the divergence of humans and chimps (Hasegawa et al. 1985; Kumar et al. 2005; Wilkinson et al. 2011) and even the origins of the HIV virus (Korber et al. 2000). Molecular clocks are not always uniform across taxa, however, leading to violations of a strict molecular clock model (Britten 1986; Lanfear et al. 2010). Additionally, in some cases molecular divergence times have conflicted with interpretations from the fossil record causing disagreement between palaeontologists and molecular biologists (Morris 1998; Pulquério & Nichols 2007). More recently, a holistic approach to divergence dating has been advocated (Parham et al. 2012) in order to increase accuracy, with palaeontologists and molecular biologists working closely together to combine accurate and comprehensive fossil calibrations with cutting edge phylogenetic techniques, including relaxed molecular clocks that allow the rate of molecular evolution to vary across taxa and using the latest models of DNA sequence evolution (e.g., Joyce et al. 2013).

Early methods for divergence dating simply took the sequence divergence between the taxa of interest and divided by a calibration rate to produce an estimate of divergence time (Bromham & Penny 2003; Kumar 2005). However, this approach has many sources of error that can lead to inaccurate dating. Significantly, the effects of mutational saturation (where more than one mutation has occurred at the same position, leading to an underestimation of the true level of divergence) are not taken into account along with other factors such as inaccuracy of the phylogeny and uncertainty in fossil calibrations. To combat this, modern divergence dating techniques use Bayesian methods that can incorporate prior information into the estimates of model parameters and uncertainty in fossil calibrations (Kumar 2005; Yang & Rannala 2006). Bayesian methods have several advantages for divergence time estimation: they incorporate complex models of sequence evolution across multiple genetic markers and can account for heterogeneity in rates across partitions (Brandley et al. 2011); they permit multiple fossil calibrations with realistic prior probabilities that model uncertainty in their placement (Yang & Rannala 2006); they allow the use of

relaxed molecular clocks that permit variation in the rate of molecular evolution across taxa (Drummond et al. 2006; Huelsenbeck et al. 2000; Kishino et al. 2001; Thorne & Kishino 2002; Thorne et al. 1998); and they can estimate divergence times over a posterior sample of trees (Yang & Rannala 2006) or simultaneously with phylogeny estimation (Drummond & Rambaut 2007; Drummond et al. 2006; Drummond et al. 2012) to take into account phylogenetic uncertainty. A key area of development has been how fossil information is more realistically incorporated into divergence dating analysis with researchers realising the importance of using appropriate prior distributions that include soft bounds and assessing how their calibration priors interact with each other to form the overall time prior for the tree (Hug & Roger 2007; Inoue et al. 2010; Warnock et al. 2012). Additionally, phylogenomic approaches to divergence dating are now providing unparalleled resolution to answer key questions on the timing and tempo of diversification of major clades in the tree of life. For example dos Reis et al. (2012) recently used a dataset including 36 animal nuclear genomes to date the radiation of placental mammals and Rehm et al. (2011) have recently used transcriptome data to date the arthropod tree.

1.6 Using molecular phylogenetic trees to infer character evolution

Phylogenetic trees reconstruct evolutionary history and are now ubiquitous in modern biology (Yang & Rannala 2012). A key area of phylogenetics that has received considerable attention and development over the last two decades is the use of molecular phylogenetic methods to reconstruct character evolution and to investigate correlated evolution between traits in a phylogenetically controlled way, i.e. to account for the non-independence of samples due to the shared ancestry of species when making comparisons of traits (Felsenstein 1985; Pagel 1999a). Methods for ancestral character state reconstruction range from simple parsimony-based approaches that rely on inferring the minimum number of transitions along a tree, irrespective of branch lengths (Maddison et al. 1984), to complex maximum likelihood and Bayesian approaches which can be used to test alternative models of character evolution (Pagel 1994; Pagel 1999b; Pagel et al. 2004). These methods have been used to investigate a diverse range of evolutionary questions such as the tempo of morphological evolution in mammals (Venditti et al. 2011), tests of Dollo’s law (Goldberg & Igić 2008; Wiens 2011), the origins of eusociality in apid bees (Cardinal &

Danforth 2011) and even the tempo of human language evolution (Atkinson et al. 2008).

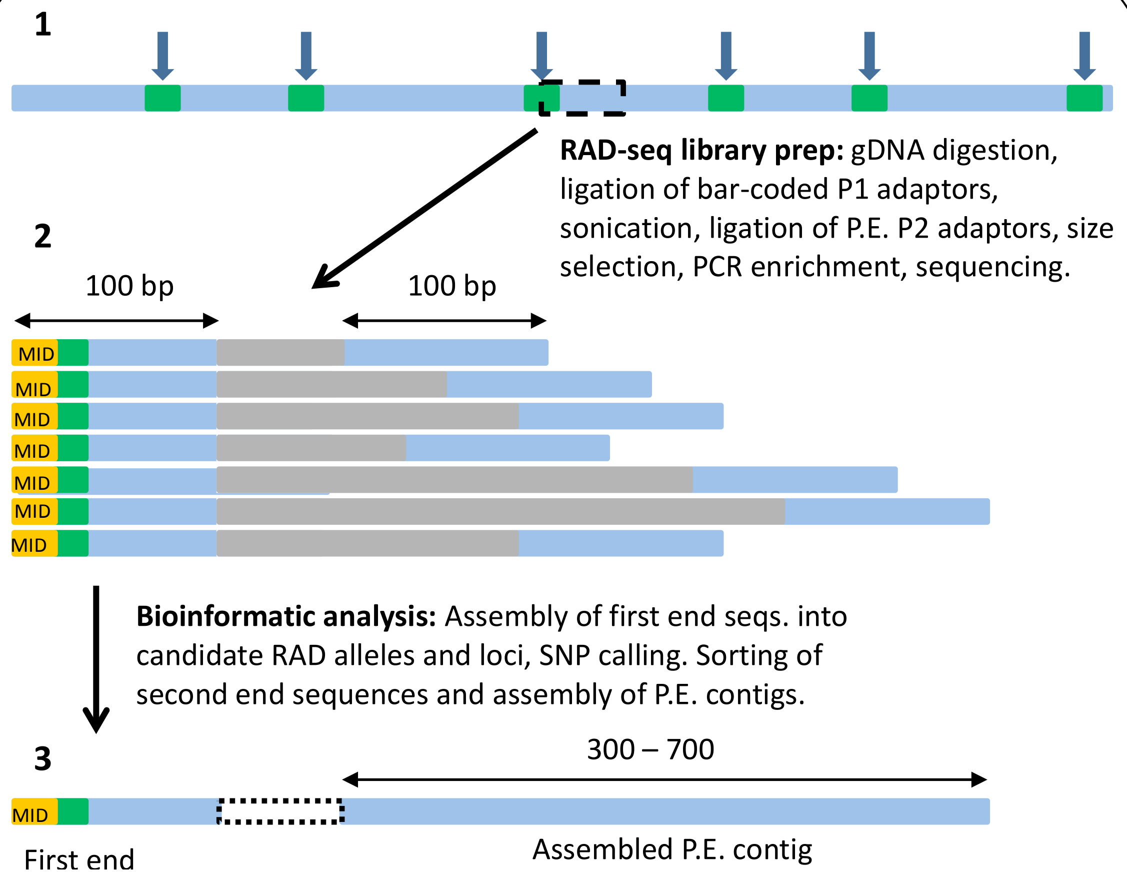
1.7 Restriction site associated DNA sequencing

The development of high throughput second generation sequencing has facilitated a step change in the scale of genetic marker discovery and has revolutionised the study of both model and non-model organisms (Ekblom & Galindo 2010; Mardis 2008; Soon et al. 2013). However, despite falling costs, whole genome sequencing of many individuals is not always tractable. One solution is genome complexity reduction and the use of reduced representation genomic libraries for genotyping by sequencing (Narum et al. 2013). Reduced representation sequencing provides a means to obtain thousands of genetic markers across the genomes of many organisms with a reduced sequencing effort and cost. This can be achieved through transcriptome sequencing, the use of capture probes or by targeting DNA associated with restriction sites (reviewed by Davey et al. 2011). Of these methods restriction site associated DNA sequencing (RAD-seq) has proved particularly useful due to its repeatability and suitability for non-model organisms where other genomic resources are unavailable (Davey & Blaxter 2010; Reitzel et al. 2013).

RAD-seq can generate thousands of single nucleotide polymorphism (SNP) markers in tens to hundreds of individuals (Baird et al. 2008; Hohenlohe et al. 2010). It has been used for the generation of linkage maps (Baird et al. 2008; Baxter et al. 2011; Richards et al. 2013), genome wide association studies (Hecht et al. 2013), population genetics (Catchen et al. 2013a; Hohenlohe et al. 2010; Keller et al. 2012; Reitzel et al. 2013), phylogenomics (Wagner et al. 2013), phylogeography (Emerson et al. 2010) and genome scaffolding (Dasmahapatra et al. 2012). Furthermore, sophisticated bioinformatic pipelines have been developed that can take the user from raw sequence data to genotypes and provide outputs for downstream analysis programs (Baxter et al. 2011; Catchen et al. 2013b; Catchen et al. 2011; Chong et al. 2012; Senn et al. 2013; Willing et al. 2011).

In RAD-seq, high throughput massively parallel Illumina sequencing is used to sequence bar-coded short reads associated with restriction sites known as RAD tags (Baird et al. 2008). A typical RAD analysis is summarised in Box 1. On average, RAD tags

are distributed evenly across the genome and, unless there is a polymorphism at a cut site, should be found in the majority of individuals of a species providing a repeatable framework for genotyping (Baird et al. 2008; Baxter et al. 2011; Hohenlohe et al. 2010). The frequency of RAD tags depends on the type of restriction enzyme used, with more frequent cutters providing greater resolution across the genome at the expense of sequence coverage (given an equivalent sequencing effort). Initially, RAD-seq was developed to identify markers close to the restriction site in what is known as the first end (see Box 1 and Baird et al. 2008). However, with the development of paired-end Illumina sequencing, where both ends of the DNA fragment are sequenced and known to be from the same fragment, RAD-seq was modified to enable the assembly of pared-end contigs (Etter et al. 2011b; Willing et al. 2011). Paired-end RAD-seq takes advantage of the random shearing process in the RAD protocol (Box 1). Whilst first end reads for each restriction site in the genome align perfectly, second end reads are staggered due to the random shearing process with their maximum distance from the restriction site determined by a size selection step in the RAD protocol. As such, paired-end reads for each RAD tag can be assembled into contigs and used for downstream applications such as the design of PCR primers, blast searching or the calling of additional SNP markers (Senn et al. 2013).



Box 1: Overview of paired-end (P.E.) RAD-seq library construction and analysis. For full details see Baird et al. (2008) and Etter et al. (2011b). 1) Genomic DNA (blue) is digested with a restriction enzyme individually for each sample to be included in the RAD library. Restriction sites across the genome are shown by blue arrows and green squares. Following DNA digestion a P1 adaptor containing a unique barcode (known as a molecular identifier (MID) (yellow block)), a PCR primer and an Illumina sequencing primer is ligated to the sticky end left by the restriction enzyme. DNA is then pooled, sonicated and size selected on an agarose gel – typically a band of 300 – 700 bp is cut out. Following the size selection the sequence ends are blunted and a P2 adapter is ligated. For P.E. RAD-seq the P2 adapter includes an Illumina sequencing primer allowing both ends of each fragment to be sequenced. The RAD library is then PCR amplified to enrich for fragments containing both a P1 and a P2 adapter. 2) Raw sequence data is assembled into candidate RAD alleles and loci. If the RAD-seq analysis program Stacks (Catchen et al. 2013b; Catchen et al. 2011) is being used individual genotypes are called simultaneously. Sequences can either be aligned to a reference genome or assembled *de novo* based on sequence similarity. In the absence of a reference genome P.E. sequences are sorted and assembled into consensus contigs using a short read assembler such as Velvet (Zerbino & Birney 2008). 3) Following assembly, first and second end (P.E.) contigs and genotype calls for each individual can be used in downstream analyses dependant on the experimental design.

* The size of the P.E. contig is dependent on the size selection and sequence coverage.

Chapter 2

Multiple global radiations in tadpole shrimps challenge the concept of 'living fossils'[1]

2.1 Abstract

'Living fossils', a phrase first coined by Darwin, are defined as species with limited recent diversification and high morphological stasis over long periods of evolutionary time. Morphological stasis, however, can potentially lead to diversification rates being underestimated. Notostraca, or tadpole shrimps, is an ancient, globally distributed order of branchiopod crustaceans regarded as 'living fossils' because their rich fossil record dates back to the early Devonian and their morphology is highly conserved. Recent phylogenetic reconstructions have shown a strong biogeographic signal, suggesting diversification due to continental breakup, and widespread cryptic speciation. However, morphological conservatism makes it difficult to place fossil taxa in a phylogenetic context. Here we reveal for the first time the timing and tempo of tadpole shrimp diversification by inferring a robust multilocus phylogeny of Branchiopoda and applying Bayesian divergence dating techniques using reliable fossil calibrations external to Notostraca. Our results suggest at least two bouts of global radiation in Notostraca, one of them recent, so questioning the validity of the 'living fossils' concept in groups where cryptic speciation is widespread.

2.2 Introduction

There has been much debate about the tempo and mode of the diversification of life (Eldredge & Gould 1972; Reznick & Ricklefs 2009; Rhodes 1983). Recently, this debate

[1] A modified version of this chapter was published as: Mathers T. C., Hammond R. L., Jenner R. A., Hänfling B. and Gómez A. (2013) Multiple global radiations in tadpole shrimps challenge the concept of 'living fossils'. *PeerJ* 1:e62 http://dx.doi.org/10.7717/peerj.62.

has been informed by dating using relaxed molecular clocks and diversification analyses; techniques which have revealed disparate patterns of speciation with early bursts (Burbrink & Pyron 2010), recent radiations (Nagalingum et al. 2011) and density dependency (Phillimore & Price 2008) being demonstrated. One extreme and often controversial pattern of diversification is found in 'living fossils', a concept introduced by Charles Darwin in *The Origin of Species* when dealing with the perplexing nature of the platypus and lungfish, relicts of once diverse groups (Darwin 1859). Since Darwin's first use, the 'living fossil' term has been applied to groups which appear to have diversified little and are morphologically stable over long periods of evolutionary time, with examples including cycads, tuatara, coelacanths, horseshoe crabs and *Ginkgo biloba*. However, morphological stasis can obscure the patterns of species diversification, and recent time-calibrated phylogenetic analysis of some 'living fossils' has indeed revealed that extant species are in fact only recently diverged (Kano et al. 2012; Nagalingum et al. 2011).

Notostraca, or tadpole shrimps, are an ancient, globally distributed order of branchiopod crustaceans with a rich fossil record dating back to the early Devonian (Fayers & Trewin 2002). The order has two extant genera, *Triops* and *Lepidurus*, in the family Triopsidae, with a yet undefined number of species. The nomenclature and systematic position of some ancient extinct Notostraca lineages is, however, problematic (Hegna & Dong 2010). Partly, this is due to the fact that throughout their evolutionary history, tadpole shrimps have maintained an extremely conserved yet complex *bauplan* with fossils indistinguishable from extant species of *Triops* found as early as the Triassic (Gall & Grauvogel-Stamm 2005; Gore 1986; Trusheim 1938) and of *Lepidurus* in the Jurassic (Barnard 1929; Haughton 1924). This striking morphological conservatism has led them to be referred to as 'living fossils' (Fryer 1988; King & Hanner 1998; Mantovani et al. 2008).

Phylogenetic reconstructions of extant Notostraca show a strong biogeographic signal (Mathers et al. 2013; Vanschoenwinkel et al. 2012). In *Triops*, species complexes are largely restricted to single continents, while *Lepidurus* lineages show high levels of endemism (Rogers 2001), patterns that suggest ancient radiation with diversification through continental break-up. However, the extreme morphological conservatism of this order hampers both the taxonomy of extant species and the phylogenetic

placement of fossil taxa, with little known about the timing and tempo of notostracan diversification. Genetic analyses have revealed widespread cryptic species (King & Hanner 1998; Korn et al. 2010; Korn & Hundsdoerfer 2006; Macdonald et al. 2011; Vanschoenwinkel et al. 2012), further illustrating the difficulty of inferring past and present diversity. To rectify this difficulty we infer a robust phylogeny of all known notostracan species from both extant genera and seven branchiopod outgroups. Our analysis uses all available Notostraca sequence data for seven genes, and Bayesian relaxed clock dating techniques, with multiple branchiopod fossil calibrations, to estimate divergence times.

2.3 Materials and methods

Species delimitation

As Notostraca is known to contain cryptic species complexes (e.g., King & Hanner 1998), and in order to follow the same criterion for species selection for the multilocus analysis, we delimited species using a generalised mixed Yule coalescent (GMYC) model (Pons et al. 2006) fitted to an ultrametric phylogeny based on all available cytochrome oxidase I (COI) sequences from GenBank. 270 sequences were aligned with Muscle (Edgar 2004) and phylogeny estimated with BEAST v1.7.4 (Drummond et al. 2012) under a constant population size coalescent tree model and GTR $+\Gamma$ substitution model. A strict molecular clock was used with the substitution rate fixed to 1 to provide branch lengths relative to an arbitrary time scale. The Markov chain Monte Carlo (MCMC) chain was run for 9,000,000 iterations with the first 500,000 iterations removed as burnin. Effective sample sizes (ESS's) of parameters (all greater than 200) and appropriate burnin were checked using Tracer v1.5 (Rambaut & Drummond 2007). From this a maximum clade credibility tree using median heights was made. We then fitted a single threshold GMYC model to the COI tree to delimit species from populations. A total of 34 species of Notostraca were identified in this analysis (Appendix 1, Table 1).

Multilocus phylogenetic analysis of Branchiopoda

We constructed a multilocus alignment containing representatives of all known species of Notostraca. Single representatives of each phylogenetic species identified by the

GMYC analysis were selected for inclusion in our phylogenetic analysis. In addition, four Notostraca lineages (*T. gadensis, T. cf. granarius* (Tunisia)*, L. bilobatus* and *L. cryptus*) which did not have COI data available but were represented by other genes, were also used in our multilocus phylogenetic analysis. The species status of these lineages has been confirmed in regional studies of cryptic diversity (King & Hanner 1998; Korn et al. 2010; Korn & Hundsdoerfer 2006; Rogers 2001). We also included seven representatives of the other branchiopod orders so that robust fossil calibrations could be applied for the dating analysis. We included sequences for the genes 12S, 16S, 28S, cytochrome oxidase I (COI), Elongation Factor 1-alpha (EF1), RNA Polymerase II and Glycogen Synthase (see Appendix 1, Table 2 for accession numbers).

Sequences were aligned using MUSCLE (Edgar 2004) with final adjustments by eye. Introns in the nuclear protein coding genes were identified and removed based on alignment with available Notostraca mRNAs. Translation was checked in MEGA 5 (Tamura et al. 2011). Overall, sequences for 45 taxa (38 notostracan and 7 branchiopod outgroups) were concatenated for analysis with the alignment containing 5793 positions and 52% missing data (Appendix 1, Table 2; the alignment file is available in Dryad DOI: 10.5061/dryad.77bt2).

Optimum partitioning schemes and substitution models for our phylogenetic analysis and divergence time estimation were identified using PartitionFinder (Lanfear et al. 2012). PartitionFinder uses a heuristic search algorithm, starting with a fully partitioned analysis (gene and codon position where appropriate), and identifies the best fit partitioning scheme and substitution models based on Bayesian Information Criterion (BIC). Due to the restricted model choice available in RAxML we conducted separate PartitionFinder analyses for the phylogenetic analysis and divergence time estimation. For the phylogenetic analysis we restricted model choice options to GTR or GTR + Γ whereas for divergence time estimation we allowed models to be selected from the full suite available in BEAST. We excluded models with proportion of invariant sites (+I) as rate heterogeneity is accounted for by the gamma shape parameter (+Γ). Optimum partitioning schemes and substitution models for both analyses are given in Appendix 1, Tables 3 and 4.

Branchiopod phylogeny was estimated using Bayesian and maximum likelihood (ML) methods with partitions and substitution models set to those identified by PartitionFinder (Appendix 1, Table 3). Bayesian analysis was performed with MrBayes v3.2 (Ronquist et al. 2012). Model parameters between partitions were unlinked. Two independent MCMC chains were run for 10,000,000 iterations each, sampling every 5,000 iterations. The first 25% of each run was discarded as burnin with the remaining samples pooled and used to create a maximum clade credibility tree. Maximum likelihood phylogenetic analysis was performed using RAxMLHPC-PTHREADS v7.0.4 (Stamatakis 2006). An initial ML search using GTR + Γ_4 was performed onto which 100 rapid bootstraps were drawn.

Bayesian relaxed clock divergence dating

We estimated Bayesian divergence times with BEAST v1.7.4 (Drummond et al. 2012) using an uncorrelated lognormal relaxed clock (Drummond et al. 2006) and a Yule speciation prior. XML files for all BEAST runs were created using BEAUTi v1.7.4 (Drummond et al. 2012). Topology was constrained to that of the unconstrained RAxML analysis. We used the best fit partitioning scheme identified by PartitionFinder (Appendix 1, Table 4) and estimated substitution model parameters independently for each partition. Initial runs were conducted using substitution models identified by PartitionFinder, however, this resulted in poor mixing of some GTR model parameters for partitions 1, 2 and 5, so subsequent runs were performed using a simpler HKY + Γ model for these partitions.

Five branchiopod fossils representing the oldest known occurrences of their respective crown groups were used to calibrate the molecular clock with minimum age constraints (Table 2.1). Lognormal prior distributions were used to specify the level of uncertainty in the placement of these fossil calibrations as they reflect the likely scenario that the true date of divergence of a given node was some time before the earliest known fossil belonging to that clade (Ho & Phillips 2009). Lognormal distributions have 3 parameters – mean, standard deviation and offset. We set the offset to correspond to the minimum age of the node as determined by the fossil record, we then specified mean and standard deviations that resulted in 95% of the distribution falling between the age of the fossil and the age of the next oldest fossil

(at a lower taxonomic level) for that group. This gives a prior distribution, which assigns the majority of the probability close to the age of the oldest known fossil but gives a long tail to account for uncertainty in the proximity of the fossil to the true date of divergence. As Bayesian divergence dating benefits from at least one maximum age constraint we conservatively constrained the root of the tree to a maximum age of 558 Mya, the age of the oldest bilaterian fossil (Fedonkin et al. 2007). To investigate the interactions between the fossil calibrations we ran the BEAST analysis using just the priors (no sequence data) to generate a marginal prior density of divergence times for each node in the tree. We compared this with the fossil calibration density for each calibrated node and found no substantial differences between them, indicating that fossil calibrations do not adversely interact with each other.

Table 2.1: Fossils used to calibrate divergence time analysis in BEAST. Age constraints are treated as hard bounds unless otherwise stated. Node numbers indicate phylogenetic placement of fossil calibrations in Figure 1.

Node	Fossil taxa	Geological period	Minimum age (Mya)	Maximum age (Mya)	Reference
1	Oldest Bilateria eg. *Kimberella*	Ediacaran	-	558	(Fedonkin et al. 2007)
1	*Rehbachiella*	Upper Cambrian	-	500*	(Waloßek 1995)
1	Undescribed anostracan	Base Ordovician	488	-	(Harvey et al. 2012)
2	*Castracollis*	Pragian, Early Devonian	410	-	(Fayers & Trewin 2002)
3	*Ebullitiocaris elatus*	Carboniferous	300	-	(Womack et al. 2012)
4	*Daphnia* and *Ctenodaphnia sp.*	Jurassic / Cretaceous	145	-	(Kotov & Taylor 2011)

* Soft maximum.

We ran two independent BEAST MCMC chains for 50,000,000 iterations, sampling every 5,000 iterations. 10,000,000 iterations were removed as burnin from each run. Convergence of the two runs and the ESS of parameter estimates (all greater than 250) where assessed using Tracer v1.5 (http://tree.bio.ed.ac.uk/software/tracer/). A posterior sample of 8000 trees from one of the runs was used to construct a maximum clade credibility time tree for Notostraca and our selected outgroups.

Diversification analysis

Patterns of diversification through time within Notostraca were investigated using LASER (Rabosky 2006) based on the BEAST time tree with outgroups pruned. Using LASER we compared constant rate and variable rate speciation models using the dAICrc test statistic which compares Akaike information criterion (AIC) scores between the best constant rate and variable rate models tested.

2.4 Results and discussion

The most complete taxon sampling to date coupled with the inclusion of multiple nuclear and mitochondrial markers allowed us to generate a robust phylogeny of 38 extant Notostraca species. ML and Bayesian inference gave highly congruent phylogenetic trees with most branches highly supported (Appendix 1, Figures 1 and 2). The recovered relationships between branchiopod orders are in agreement with recently published arthropod phylogenies (Regier et al. 2010; von Reumont et al. 2012), providing a solid platform for divergence dating analysis.

Our robust time-calibrated phylogeny of Branchiopoda (Figure 2.1) clearly shows that Notostraca have a pattern of diversity incompatible with Darwin's original usage of the term 'living fossil' as relics of once diverse groups, and importantly reveals cryptic patterns of diversification. Our analysis – using outgroup fossil calibrations – estimates an ancient divergence of *Triops* and *Lepidurus* during the Jurassic, 184 Mya (95% confidence interval 132 – 259 Mya), which agrees with the earliest fossils assigned to *Lepidurus* (Barnard 1929; Gand et al. 1997), and with a sister relationship of Notostraca to the extinct order Kazacharthra of the Late Triassic/Early Jurassic (Olesen 2009). The initial radiation of extant Notostraca was not, however, due to continental break up as the timing and pattern of diversification within the genera substantially postdates the break-up of Pangaea 160 – 138 Mya (Scotese 2001). Furthermore, the current species distributions of *Triops* and *Lepidurus* are likely to have resulted from a second global radiation of the order, possibly following considerable levels of extinction. We conclude this because fossil Notostraca, attributed to *Triops* and *Lepidurus*, have been found in modern day North and South America, Europe, Africa and Antarctica, implying a global distribution by the early Jurassic (Gall & Grauvogel-Stamm 2005; Gand et al. 1997; Garrouste et al. 2009; Gore 1986; Haughton 1924; Trusheim 1938) yet our LASER

analysis shows a significant increase in the rate of diversification of Notostraca about 73 Mya (Figure 2.2), close to the time of the Cretaceous-Palaeogene mass extinction event. It is this second radiation that resulted in the current global distribution of extant *Triops* and *Lepidurus*.

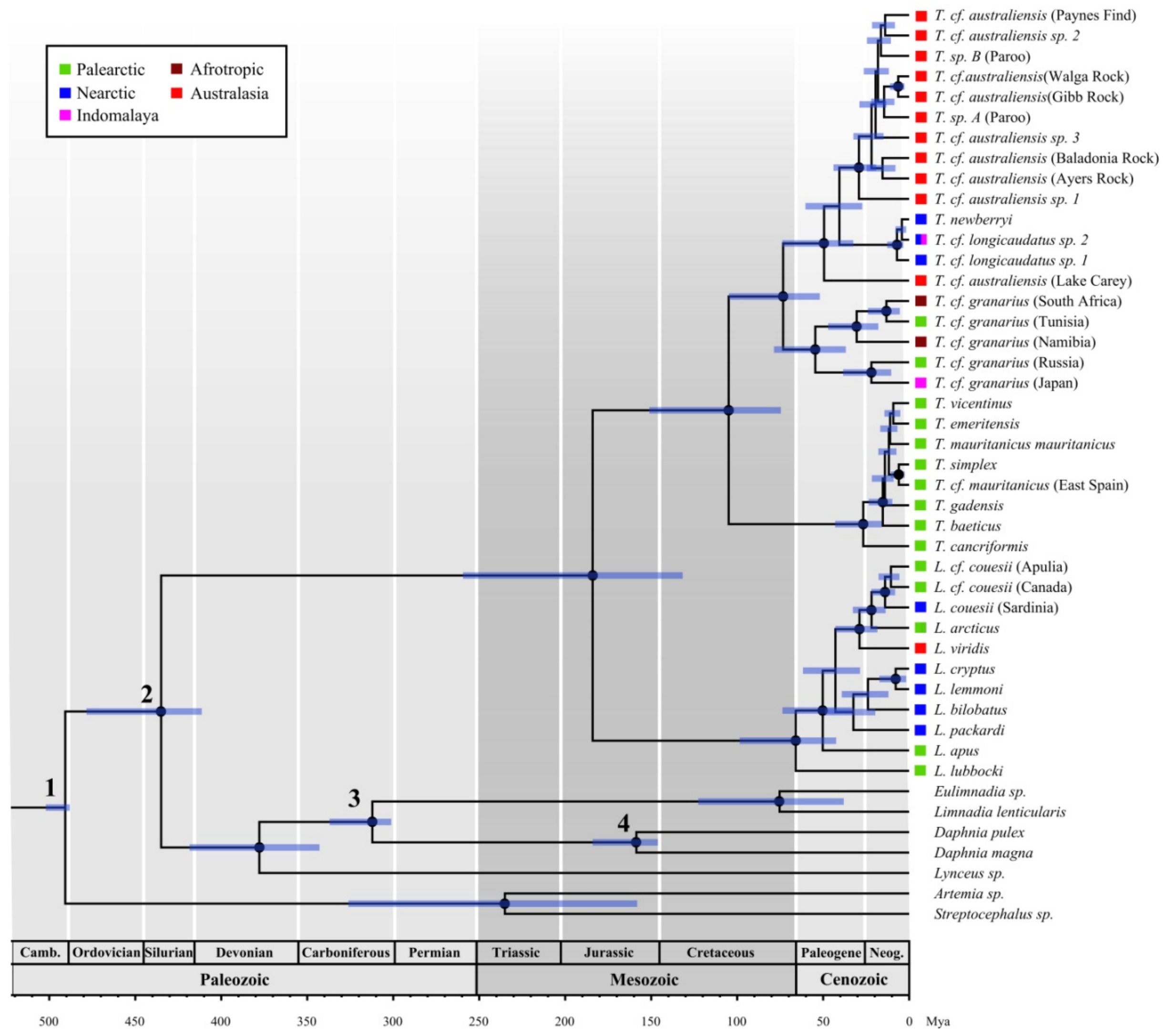


Figure 2.1: Time calibrated phylogeny of 38 notostracan species and seven branchiopod outgroups. Numbers at nodes correspond to the fossil calibrations given in Table 2.1. Nodes with black circles have ML Bootstrap support values greater than 70 and posterior probabilities greater than 95 from the RAxML and MrBayes analyses respectively. Error bars show the 95% confidence intervals of divergence times. Colour coded squares show the biogeographic regions in which each species is found.

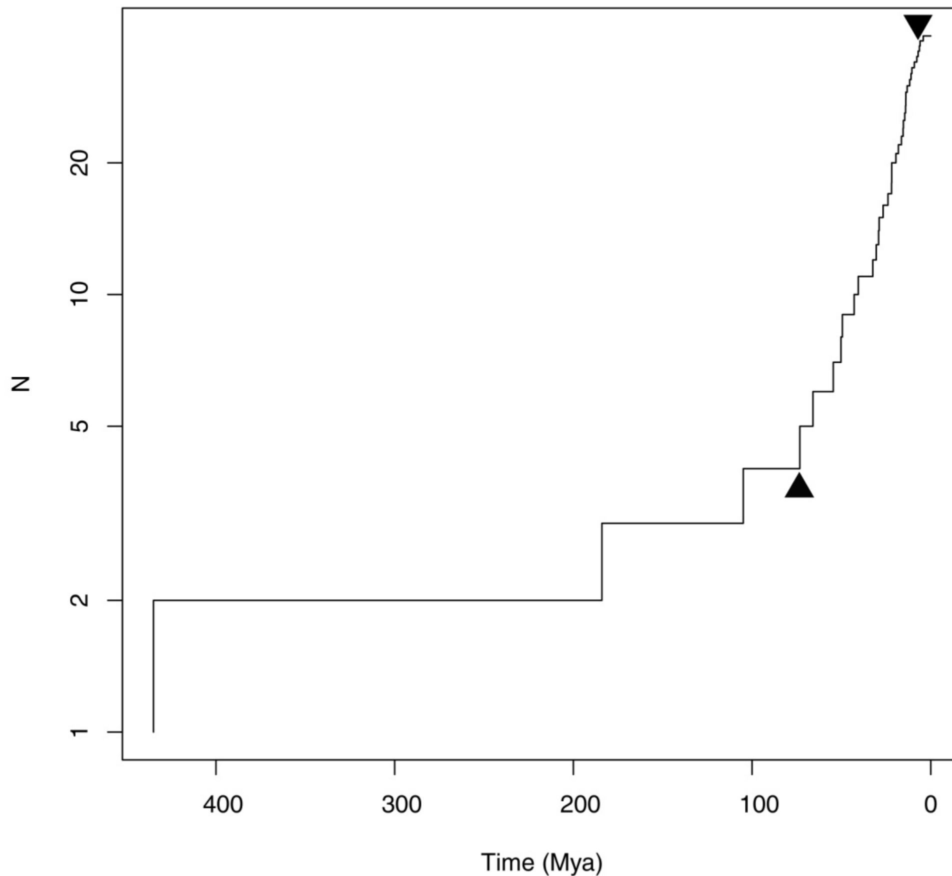


Figure 2.2: Diversification of Notostraca through time. Arrows indicate the timing and direction of shifts in rate of diversification inferred by LASER. *N* is the number of species. The best fit ML model of diversification (Appendix 2, Table 5) identified three distinct rates of diversification during the evolutionary history of Notostraca with an increase in speciation rate 73 Mya followed by a decrease 6 Mya.

The almost synchronous radiation of *Triops* and *Lepidurus* (Figure 2.1) suggests that a common factor may have triggered diversification of the two genera. The diversification of modern birds – widely involved in dispersal in aquatic invertebrates (Green et al. 2005; van Leeuwen et al. 2012) – coincided with the initiation of the notostracan radiation (Pacheco et al. 2011) and may have facilitated the long distance dispersal and subsequent diversification of Notostraca. Indeed, the geographical distribution of extant taxa (Figure 2.1) suggests several instances of intercontinental dispersal. For example, the colonisation of North America from Australia could have resulted from dispersal events during bird migration. Such long distance dispersal and colonisation events might also have been facilitated by the flexible nature of sexual systems found within Notostraca (see Chapter 3). Indeed, the evolution of androdioecy – a sexual system where males and hermaphrodites coexist (Weeks 2012; Zierold et al.

2007; Zierold et al. 2009) – from gonochorism appears to have favoured postglacial recolonization in the species *Triops cancriformis* (Zierold et al. 2009).

The concept of 'living fossils' has been a controversial one as it has often been interpreted to imply a lack of evolutionary change, even against evidence of molecular evolutionary change (Avise et al. 1994; Casane & Laurenti 2013). Our divergence dating analysis has shown that tadpole shrimps can be regarded as 'living fossils' only on the grounds of morphological conservatism, not on the basis of limited diversification or relict status. Instead, throughout their long evolutionary history, notostracans have undergone multiple global radiations and high species turnover. Recent, time calibrated, phylogenetic analysis of other traditional 'living fossils' such as cycads (Nagalingum et al. 2011), nautiloids (Wray et al. 1995) horseshoe crabs (Obst et al. 2012) and monoplacophorans (Kano et al. 2012) have also revealed that extant species are more recently diverged than suggested by fossil data alone. We therefore caution against drawing conclusions on patterns of diversification based on fossil data alone in groups where widespread morphological conservatism may obscure rampant cryptic speciation. Furthermore, our results help clarify the term 'living fossils', putting important questions into focus. Namely, is such morphological conservatism, in the face of evolutionarily recent diversification and radiation, best accounted for by unchanging selection or by developmental genetic constraints?

Chapter 3

High lability of sexual system over 250 million years of evolution in morphologically conservative tadpole shrimps[1]

3.1 Abstract

Sexual system is a key factor affecting the genetic diversity, population structure, genome structure and the evolutionary potential of species. The sexual system androdioecy – where males and hermaphrodites coexist in populations – is extremely rare, yet is found in three crustacean groups, barnacles, a genus of clam shrimps *Eulimnadia*, and in the order Notostraca, the tadpole shrimps. In the ancient crustacean order Notostraca, high morphological conservatism contrasts with a wide diversity of sexual systems, including androdioecy. An understanding of the evolution of sexual systems in this group has been hampered by poor phylogenetic resolution and confounded by the widespread occurrence of cryptic species. Here we use a multigene supermatrix for 30 taxa to produce a comprehensive phylogenetic reconstruction of Notostraca. Based on this phylogenetic reconstruction we use character mapping techniques to investigate the evolution of sexual systems. We also tested the hypothesis that reproductive assurance has driven the evolution of androdioecy in Notostraca. Character mapping analysis showed that sexual system is an extremely flexible trait within Notostraca, with repeated shifts between

---

[1] A modified version of this chapter was published as: Mathers T. C., Hammond R. L., Jenner R. A., Zierold T., Hänfling B. and Gómez A. (2013) High lability of sexual system over 250 million years of evolution in morphologically conservative tadpole shrimps. *BMC Evolutionary Biology* 13:30 doi:10.1186/1471-2148-13-30.

gonochorism and androdioecy, the latter having evolved a minimum of five times. In agreement with the reproductive assurance hypothesis androdioecious notostracans are found at significantly higher latitudes than gonochoric ones indicating that post glacial re-colonisation may have selected for the higher colonisation ability conferred by androdioecy. In contrast to their conserved morphology, sexual system in Notostraca is highly labile and the rare reproductive mode androdioecy has evolved repeatedly within the order. Furthermore, we conclude that this lability of sexual system has been maintained for at least 250 million years and may have contributed to the long-term evolutionary persistence of Notostraca. Our results further our understanding of the evolution of androdioecy, and indicate that reproductive assurance is a recurrent theme involved in the evolution of this sexual system.

3.2 Introduction

Plants and animals have evolved a great diversity of sexual systems that range from the extremes of habitual self-fertilisation to complete outcrossing. Transitions between these sexual systems have long fascinated biologists due to the impacts they have on key parameters such as inbreeding depression, genetic diversity, population structure, genome structure and the evolutionary potential of species (Barrett et al. 2008; Barrett 1998; Charlesworth & Wright 2001; Charnov 1982; Charnov et al. 1976; Darwin 1876; Holsinger 2000). Transitions between sexual systems often present tradeoffs between short and long term selective advantages and can have significant connotations for the long-term viability of species. For example, selection for reproductive assurance and colonisation advantage due to mate limitation during range expansions, or as a result of high population turnover in metapopulations, can drive transitions to self-fertilisation strategies (Baker 1955; Baker 1967; Hesse & Pannell 2011). These transitions occur despite the deleterious effects of self-fertilisation, which include inbreeding depression, reduction in effective recombination rates and reduction in effective population size (Glemin et al. 2006).

Transitions to androdioecy (AD) – a sexual system where males and hermaphrodites co-occur in varying frequencies in populations, with different levels of self-fertilisation and outcrossing – are extremely rare in plants and animals (Charlesworth 1984; Darwin 1877; Pannell 1997; Pannell 2002a; Pannell 2002b). In animals, AD has only been

described in five groups, rhabditid nematodes, the killifish *Kryptolebias marmoratus,* and three crustacean groups, barnacles, a genus of clam shrimps *Eulimnadia*, and in the order Notostraca, the tadpole shrimps (Weeks et al. 2006a; Yusa et al. 2012). AD can evolve either through the invasion of males into hermaphrodite only populations, as in barnacles (Urano et al. 2009; Yusa et al. 2012), or through the replacement of females with hermaphrodites in gonochoric populations (where males and females are found in approximate equality), as in the plants *Mercurialis annua* (Obbard et al. 2006a; Pannell et al. 2004) and *Datisca glomerata* (Listen et al. 1990). As models to describe the evolution and maintenance of AD only predict its evolution under stringent conditions, AD has historically been considered an unstable, transitional sexual system beween gonochorism and hermaphroditism (or vice versa) (Charlesworth 1984; Lloyd 1975; Pannell 1997; Pannell 2002a; Wolf & Takebayashi 2004). This view is born out by the scarcity of AD in nature (Pannell 2002a; Weeks et al. 2006a), although recent research in the branchiopod *Eulimnadia* has revealed an unexpected stability of androdioecy (Weeks et al. 2006b).

Notostraca, or tadpole shrimps, is a small order of branchiopod crustaceans characterised by a high level of morphological stasis. Fossils dating back as far as the Triassic are almost indistinguishable from contemporary species leading them to be referred to as ‘living fossils’ (Barnard 1929; Fryer 1988; Gall & Grauvogel-Stamm 2005; Gore 1986; Longhurst 1955c). In contrast, Notostraca has diverse sexual systems, including gonochorism, self-fertile hermaphroditism and AD, with variation occurring on both an interspecific and intraspecific level (Macdonald et al. 2011; Zierold et al. 2007). Remarkably, AD is found in species from both notostracan genera, *Triops* and *Lepidurus*, suggesting that transitions in reproductive system might have evolved repeatedly in the order. Despite this, the evolutionary history of reproductive systems in Notostraca is unknown due to the lack of a resolved phylogeny (Mantovani et al. 2004; Murugan et al. 2002), and the poor knowledge of the diversity of the group, partly due to the widespread presence of cryptic species (King & Hanner 1998; Korn et al. 2010; Korn & Hundsdoerfer 2006; Korn et al. 2006; Macdonald et al. 2011). Gonochorism has been hypothesized to be the ancestral state in the group, and the evolution of self-fertile hermaphroditism and AD has been linked to reproductive assurance in the context of range expansions, possibly after glacial retreat (Baker 1955;

Bernard 1891; Longhurst 1955a; Zierold et al. 2007), although this has never been explicitly tested.

Here we combine newly generated and GenBank sequence data to assess Notostraca taxonomic diversity, identifying considerable cryptic diversity, and employ a multigene phylogenetic approach to create a well-supported, global phylogeny of Notostraca. Information on sexual system was compiled and Maximum Parsimony (MP) and model-based Maximum Likelihood (ML) character mapping approaches were used on the phylogeny to investigate sexual system evolution across the order. We also tested the hypothesis that reproductive assurance has driven the evolution of self-fertilisation across Notostraca (Baker 1955; Longhurst 1955a). Taxa found at higher latitudes are likely to have experienced bouts of colonisation during post glacial range expansions, which would select for AD / hermaphroditism. We therefore compared the latitudes that AD / hermaphroditic and gonochoric taxa are found using a phylogenetic *t*-test. Our analyses reveal high levels of reproductive lability with frequent transitions occurring to and from androdioecy. Furthermore, this flexibility is conserved across Notostraca, and may have been maintained for at least 250 million years. Additionally, AD / hermaphroditic taxa are found at significantly higher latitudes than gonochoric ones suggesting that colonisation advantage through reproductive assurance is likely to be involved in transitions between sexual systems in Notostraca.

3.3 Materials and Methods

Sampling and Sequencing

We produced de novo sequence data from 12 taxa from the two notostracan genera, *Triops* and *Lepidurus* (Longhurst 1955c). Samples consisted of either sediments containing resting eggs or wild caught individuals preserved in 100% ethanol (Appendix 2, Table 1). Total genomic DNA was extracted from ethanol-preserved individuals using a Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) or directly from individual resting eggs using a modified 'HotSHOT' procedure (Montero-Pau et al. 2008). DNA sequences were generated for three mitochondrial gene fragments, cytochrome oxidase subunit one (COI), 12S rDNA and 16S rDNA, and four nuclear gene fragments; elongation factor 1 alpha, glycogen synthase, RNA polymerase II and 28S rDNA. We used primer pairs known to amplify across Notostraca for the mitochondrial and

ribosomal genes (Folmer et al. 1994; Korn & Hundsdoerfer 2006; Murugan et al. 2002), and designed new primers for the nuclear protein coding genes based on alignments of available sequences from notostracans and other branchiopods using PriFi (Fredslund et al. 2005) (see Appendix 2, Table 2 for primer sequences and optimised reaction conditions). Reactions were carried out in a final volume of 50 µl containing 2 µl of template DNA, 200 µM of each primer, 200 µM of each nucleotide, 0.01 U of *BioTaq* DNA polymerase (Bioline), 1x NH_4 buffer (Bioline) and 2–3 mM $MgCl_2$. Amplified fragments were purified and sequenced for both forward and reverse strands by Macrogen using an ABI 3730*xl* DNA Analyser (Macrogen Inc, Seoul, Korea). Sequences were manually edited using CodonCode Aligner v3.5 (CodonCode Corporation, Dedham, MA) with consensus sequences produced for each forward and reverse pair.

GMYC Model Based Species Delimitation

Available Notostraca COI sequences were downloaded from GenBank (Appendix 2, Table 3) and aligned with our newly generated sequences in MEGA 5 (Tamura et al. 2011) using MUSCLE (Edgar 2004) with default parameters. We applied a generalized mixed Yule coalescent (GMYC) model (Pons et al. 2006) to identify independently evolving clusters in our COI dataset, which correspond to significant taxonomic units (STUs). First, we created an ultrametric phylogeny based on our COI alignment using BEAST v1.6.2 (Drummond & Rambaut 2007). The phylogenetic analysis was run for 600,000 iterations with trees printed every 1,000 iterations and the first 100,000 iterations removed as burnin. A GTR + *Γ* nucleotide substitution model was used with a strict molecular clock with the rate fixed to 1. From this the ultrametric maximum clade credibility consensus tree was constructed. The GMYC analysis was performed in R v2.14.1 (R Development Core Team, 2011) with the package splits v1.0-11 (https://r-forge.r-project.org/projects/splits/). Clusters defined by the GMYC analysis were then assigned an STU I.D. based on the geographic location and species assignment of the accessions they contained. Uncorrected mean genetic distances in COI between STUs were calculated in MEGA 5 (Tamura et al. 2011) with all positions containing gaps or missing data removed.

Alignment and Supermatrix Construction

Single sequences for each gene for each STU identified were selected (where available) for inclusion in our multigene phylogenetic analysis (Appendix 2, Table 4). Sequences generated in this study were used preferentially but, where only GenBank sequences were available, records were checked to confirm that samples were from the same or close geographic location to samples used for STU identification.

The final alignment of each nuclear protein coding gene and the mitochondrial gene COI was carried out in MEGA 5 (Tamura et al. 2011) with MUSCLE (Edgar 2004) using default parameters. The ribosomal genes were aligned based on secondary structure information using RNAsalsa v0.8.1 (Stocsits et al. 2009) with *Apis mellifera* structural data used as a constraint. Weakened constraint settings (S1, S2 and S3 = 0.51) were used to preserve structural information as described by Letsch and Kjer (2011). To confirm that the individual alignments were suitable for concatenation, phylogenetic congruence was tested with Concaterpillar v1.4 (Leigh et al. 2008; Stamatakis 2006) using the GTR model and an α-level cut off of 0.05. No significant phylogenetic incongruence was identified ($p = 0.55$) and so all genes were concatenated using FASconCAT v1.0 (Kuck & Meusemann 2010). The final supermatrix contained 5253 positions with 54% missing data. This number represents the overall missing data, not including indels, in the supermatrix alignment. It reflects the fact that most taxa retrieved from GenBank do not have coverage for all the genes used in this study and we could not obtain sequences for some genes on a few of our samples.

Phylogenetic Analysis

Phylogeny, based on the concatenated supermatrix, was inferred by ML and Bayesian Markov chain Monte Carlo (MCMC) methods. We estimated the ML tree with RAxML using RAxMLHPC-PTHREADS v7.0.4 (Stamatakis 2006), treating each gene as an individual partition. An initial ML search using GTR + Γ was performed with 100 iterations to identify the best scoring ML tree. 1000 Bootstrap replicates were then conducted using GTR +Γ and drawn onto this best scoring ML tree. Bayesian phylogenetic analysis was conducted using BayesPhylogenies v1.0 (Pagel & Meade 2004) with a reversible jump mixture model (Pagel & Meade 2005) using a GTR model of sequence evolution with 4 discrete Γ rate categories. The analysis was run for

10,000,000 iterations with trees printed and saved every 10,000 iterations. Three independent rate matrices were assigned by BayesPhylogenies. Following this analysis, the first 500,000 iterations were removed as burnin, and the remaining 950 trees were used to create a consensus tree in BayesTrees v1.3 (www.evolution.rdg.ac.uk/BayesTrees.html).

Sexual system

Male notostracans are readily identified by the lack of ovisacs, subtle morphological differences in carapace shape, numbers of legless rings and active mating behaviour in live individuals (Akita 1971; Linder 1952; Longhurst 1955c; Macdonald et al. 2011; Obregon-Barboza et al. 2007; Sassaman et al. 1997). Females and hermaphrodites, although identical in external morphology and behaviour, differ histologically by the presence in hermaphrodites of an ovotestis (testicular lobes amongst the ovarian lobes) and by their ability to reproduce in isolation (Bernard 1891; Sassaman 1991).

We compiled data from the literature for sex ratio, histology (i.e. presence/absence of ovotestis) and the inability/ability of females/hermaphrodites to reproduce in isolation. Studies showing inability to reproduce in isolation were only included if reproduction in the presence of males was confirmed, to rule out poor lab rearing conditions or lack of reproductive maturity of individuals. In addition, we estimated sex ratio from available samples for a few populations (Appendix 2, Table 5). Using these data, we assigned populations as either being androdioecious or gonochoric. Androdioecious populations consist of hermaphrodites and males and exhibit skewed sex ratios with hermaphrodites found in greater numbers than males (Pannell 2002a; Pannell 2008; Sassaman et al. 1997; Weeks et al. 2006a). Gonochoric populations consist of males and females and have an approximately equal sex ratio. We did not categorise any population as purely hermaphroditic because this would necessitate showing a complete absence of males. Given that males in androdioecious species can be maintained by metapopulation dynamics (Pannell 1997; Pannell 2000; Pannell 2002a), and can be present in exceedingly low frequencies (e.g. eight males per thousand in *L. apus* and similar proportions in *T. cancriformis* (Engelmann & Hahn 2005; Heidecke & Neumann 1987; Simon 1886), large samples sizes where no males are found would be needed to establish that a population is hermaphroditic (Weeks et

al. 2006b). In view of the sample sizes available to us we decided to conservatively categorise STUs into two sexual systems, gonochoric and androdioecious. In the AD notostracan species *T. newberryi* male proportions never exceed 27% (Macdonald et al. 2011; Sassaman et al. 1997) and in populations of AD *Eulimnadia* male proportions were always significantly lower than 50% male with a mode of ~20% (Weeks et al. 2008). Weeks et al. (2008) did, however, note that upper values for population sex ratio of AD taxa overlapped with the lower values of gonochoric taxa in the 35% - 45% range. We therefore used a conservative population sex ratio cut-off of 30% male to assign an AD sexual system in the absence of additional histological or reproduction in isolation data in order to prevent misclassification due to stochastic variation in natural population sex ratios (Machado et al. 1999; Petrov & Cvetkovic 1997). Populations with a male proportion of 30% - 45% were coded as equivocal and populations with male proportions greater than 45% were coded as gonochoric.

Character Mapping

Sexual system was mapped onto the best scoring ML tree as a discrete character. STUs for which sexual system could not be inferred, or for which data was lacking, were left uncoded for sexual system in our analyses. The *Leptestheria* outgroup used to root the tree for character mapping analyses was also left uncoded for sexual system. MP reconstruction of ancestral states was conducted using Mesquite v2.74 (Maddison & Maddison 2010) with an unordered model. In addition, we used BayesMultistate (Pagel et al. 2004) implemented in BayesTraits v1.0 in an ML framework to evaluate four alternative models of sexual system evolution using likelihood ratio tests (*D*) assuming the result approximates a chi-squared distribution with degrees of freedom equal to the difference in the number of estimated parameters between the models. The simplest model is a one-parameter model, with a single rate of transition between gonochorism and AD and vice versa. The second model is a two-parameter model, where the transition rates from AD to gonochorism and vice versa can vary. The third and fourth models allow only unidirectional changes in sexual system, one from gonochorism to AD only, as in clam shrimps (Weeks et al. 2006b), and the other from AD to gonochorism only. Ancestral character states were reconstructed based on the best fit model using the AddNode function of BayesTraits.

Testing the reproductive assurance hypothesis

We used a proxy for the exposure of STUs to glacial cycles, and therefore presumed range expansions, to test whether the reproductive assurance hypothesis is responsible for sexual system evolution in Notostraca (Bernard 1891; Longhurst 1955a). As STUs found at higher latitudes are more likely to have recently re-colonised following the last glacial maxima than lower latitude ones, we expect AD STUs to be found at higher latitudes than gonochoric ones. The absolute latitude values at which gonochoric and AD STUs are found were compiled using the collection location of each representative STU as an unbiased representation of the latitude at which that lineage is found (Appendix 2, Table 6). We used the program BayesTraits (Pagel et al. 2004) in an ML framework to conduct a *t*-test which accounts for the shared ancestry as implied by our best scoring ML phylogeny (phylogenetic t-test) to determine if latitude significantly differs between gonochoric STUs and ones where AD populations are found (the presence and absence of AD was incorporated using standard contrast or 'dummy' coding). We simultaneously estimated the parameter λ which detects the phylogenetic signal in the data (Freckleton et al. 2002; Pagel 1999b), if λ is close to 1 there is strong phylogenetic signal if it is 0 there is no phylogenetic signal and the model collapses to an ordinary *t*-test.

3.4 Results

Delimitation of significant taxonomic units

Notostraca is known to contain cryptic species complexes (e.g. King & Hanner 1998; Korn & Hundsdoerfer 2006) so we first used a COI–based barcoding approach to identify STUs for inclusion in our phylogeny prior to the multigene analysis. Including available GenBank data and 12 newly generated sequences for this study, 243 Notostraca COI sequences were aligned. We applied a GMYC model to identify independently evolving clusters in our COI dataset, which correspond to STUs. The GMYC model identified 26 STUs (Appendix 2, Figure 1). Uncorrected mean genetic distances in COI between STUs ranged from 2.3% to 24.3%. Four Notostraca lineages did not have COI data available, but are represented by other genes used in our multigene phylogenetic analysis; *T. gadensis, T. cf. granarius (Tunisia), L. bilobatus* and *L. cryptus*. As the species status of these lineages has been confirmed in regional

studies of cryptic diversity in Notostraca (King & Hanner 1998; Korn et al. 2010; Korn & Hundsdoerfer 2006; Rogers 2001) they were included as additional STUs for the multigene phylogenetic analysis. In total we recognise 30 STUs within Notostraca.

Notostracan phylogeny

Phylogeny, based on a concatenated supermatrix of 110 sequences (54 of which were newly generated for this study) from three mitochondrial genes and four nuclear genes for 30 STUs, was inferred by ML and Bayesian Markov-chain Monte Carlo methods. Both methods of phylogenetic reconstruction gave congruent topologies, with most branches having high levels of support (Figure 1). The two recognised notostracan genera, *Triops* and *Lepidurus*, formed highly supported clades. Within *Triops*, four main monophyletic lineages with a strong geographic association (Australian, N American, Palearctic and African/Asian respectively) were highly supported. The analysis supported a close relationship between the *T. australiensis* complex (Australia) and the *T. longicaudatus* complex (North America), and a sister relationship of these with the *T. granarius* complex (Asia / Africa). The *T. cancriformis / mauritanicus* complex (Palearctic) appears as the sister group to the rest of *Triops.* Within *Lepidurus*, *L. lubbocki* (Mediterranean), and *L. apus* sensu stricto (N European) have long branches and are sister species to the rest of *Lepidurus*. Four North American species (King & Hanner 1998; Rogers 2001) *L. packardi*, *L. cryptus*, *L. bilobatus* and *L. lemmoni*, with narrowly endemic, mostly allopatric distributions in western North America, form a well supported group. *L. arcticus,* a circumpolar species from Arctic and Subarctic regions, forms a sister relationship with a clade containing *L. couesii*, which forms a widely distributed species complex.

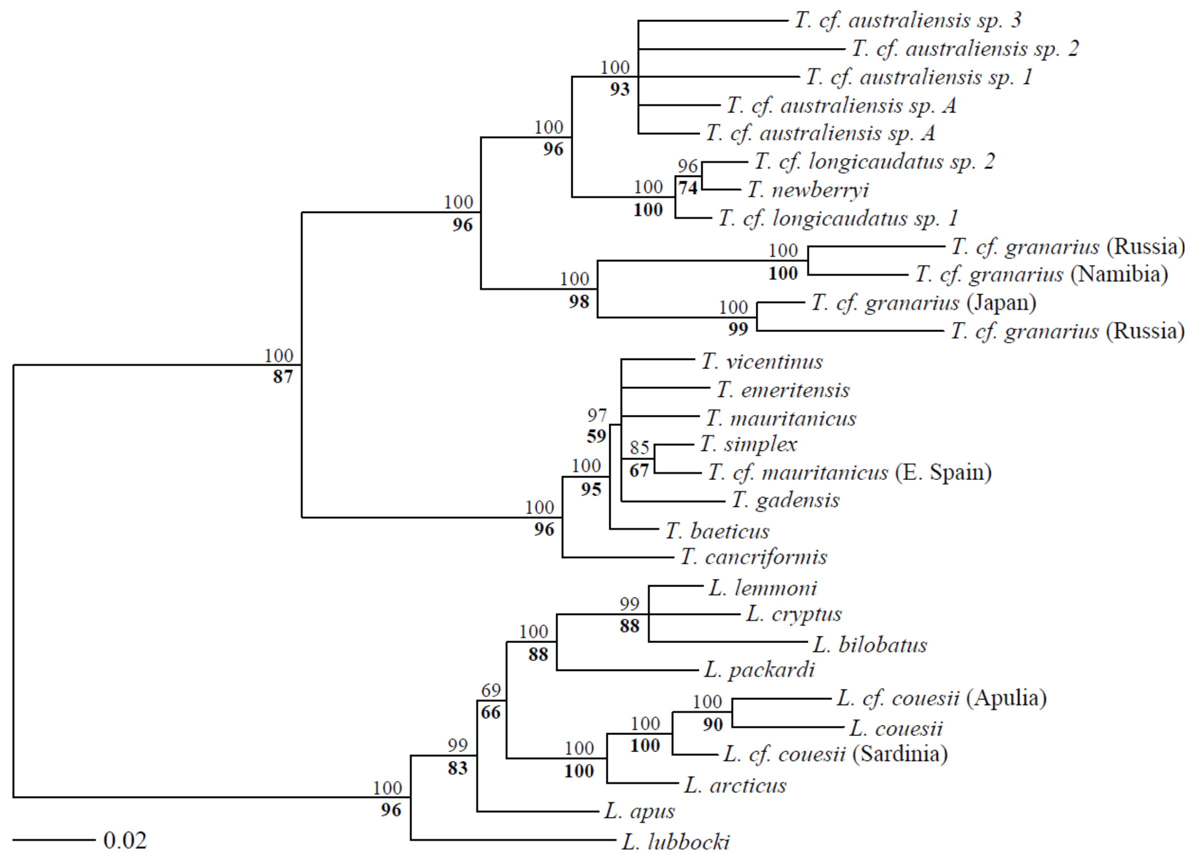


Figure 3.1: Phylogenetic relationships in Notostraca based on a multigene supermatrix. Tree topology shows the best scoring ML tree from the RAxML analysis. Values above nodes show Bayesian posterior probabilities and values below nodes, in bold, show bootstrap support (1000 replicates). Branches with bootstrap support lower than 50 are collapsed. The *Leptestheria* outgroup was removed after rooting.

Sexual system assignment

Our literature review identified sexual system data for 22 STUs (18 of these having at least one barcoded population - either COI, 12S or 16S) and is summarised in Table 3.1. In the few cases where an STU does not have a barcode sequence for the same population for which sexual system data is derived (*L. arcticus, L. couesii, L. lemmoni, L. packardi*) these species have been well studied, no further cryptic diversity has been identified, and the populations used for sexual system inference fall inside known species ranges (see Hessen et al. 2004; King & Hanner 1998; Rogers 2001). We found two polymorphic STUs, *T. cancriformis* and *T. cf. longicaudatus sp.2*, which include both androdioecious and gonochoric populations, and four androdioecious STUs; *T. newberryi*, *T. cf. australiensis sp. B, L. apus (sensu stricto)* and *L. arcticus* (Table 3.1 and Appendix 2, Table 5). Sex ratio of populations assigned AD ranged from 0% to 27% males. In addition, either histological data or reproduction in isolation data, or both,

confirmed the presence of anatomical hermaphrodites. Fourteen taxa were considered gonochoric on the basis of histology, lack of reproduction in isolation and sex ratio. The quality and quantity of information on sexual system was heterogeneous amongst gonochoric taxa, with actual sex ratios tending to be underreported, and histological studies lacking for many taxa. Eight taxa could not be assigned to a sexual system due to an absence of data or equivocal sex ratio.

Table 3.1: Notostraca sexual system information. Sex ratio (percent male), ability of ovisac bearing individuals to reproduce in isolation and the presence or absence of ovotestis in ovisac bearing individuals is shown. See Appendix 2 Table 5 for detailed information.

STU	Sex ratio	Reproduction in isolation	Ovotestis	Sexual system	References
T. baeticus	> 45			Gonochoric	(Korn et al. 2010)
T. cancriformis	0 – 53*	Yes**	Yes**	Gonochoric / AD	(Bernard 1891; Longhurst 1954; Longhurst 1955c; Zierold et al. 2007; Zierold et al. 2009)
T. cf. australiensis sp. 1				No data	
T. cf. australiensis sp. 2				No data	
T. cf. australiensis sp. 3				No data	
T. cf. australiensis sp. A	> 45		No	Gonochoric	(Murugan et al. 2009)
T. cf. australiensis sp. B	< 30		Yes	AD	(Murugan et al. 2009)
T. emeritensis	>45			Gonochoric	(Korn et al. 2010)
T. gadensis	36			Equivocal	(Korn et al. 2010)
T. cf. granarius (Japan)	> 45	No		Gonochoric	(Longhurst 1954; Longhurst 1955c; Mitsumoto & Yahata 2006; SunoUchi et al. 1997)
T. cf. granarius (Namibia)	Even***	No		Gonochoric	(Korn & Hundsdoerfer 2006; Longhurst 1954; Longhurst 1955c)
T. cf. granarius (Tunisia)	Even***	No		Gonochoric	(Korn & Hundsdoerfer 2006; Longhurst 1954; Longhurst 1955c)
T. cf. granarius (Russia)				No data	
T. cf. longicaudatus sp. 1	> 45			Gonochoric	(Macdonald et al. 2011)
T. cf. longicaudatus sp.2	0 – 68*		Yes**	Gonochoric / AD	(Garcia-Velazco et al. 2009; Macdonald et al. 2011; Sassaman et

					al. 1997)
T. cf. mauritanicus (E Spain)				No data	
T. mauritanicus				No data	
T. newberryi	< 30	Yes		AD	(Macdonald et al. 2011; Sassaman 1991; Sassaman et al. 1997)
T. simplex	> 45			Gonochoric	(Korn et al. 2006)
T. vicentinus	> 45			Gonochoric	(Korn et al. 2010; Machado et al. 1999)
L. apus	< 45		Yes	AD	(Bernard 1891; Bernard 1896; Engelmann & Hahn 2005; Longhurst 1954; Longhurst 1955c; Simon 1886; Stephan 2008)
L. arcticus	<45		Yes	AD	(Bernard 1891; Bernard 1896; Bushnell & Byron 1979; Linder 1952; Longhurst 1954; Longhurst 1955c; Wojtasik & Brylka-Wolk 2010)
L. bilobatus	35			Equivocal	
L. cf. couesii (Apulia)				No data	
L. cf. couesii (Sardinia)	>45			Gonochoric	(Margraf & Maass 1982)
L. couesii (Canada)	>45			Gonochoric	(Packard 1875)
L. cryptus				No data	
L. lemmoni	>45			Gonochoric	(Linder 1952; Lynch 1966)
L. lubbocki	>45		No	Gonochoric	(Kuller & Gasith 1996; Scanabissi & Mondini 2002)
L. packardi	>45			Gonochoric	(Ahl 1991)

* Depending on population. ** Only in populations where sex ratio is < 30%. *** Reported gonochoric with even sex ratio, exact numbers not given.

Sexual system evolution

MP reconstruction of ancestral character states infers that gonochorism is the ancestral state of Notostraca (Figure 3.2). Furthermore, AD appears to have multiple origins in Notostraca, having evolved three times in *Triops* and twice in *Lepidurus*. Sexual system is highly flexible across Notostraca and varies even between closely related species (*T. cf. australiensis A* vs. *T. cf. australiensis B*; *T. newberryi* vs. *T. cf. longicaudatus sp.1*) or shows intraspecific variation (*T. cancriformis; T. cf. longicaudatus sp.2*). Model based ML methods showed that a two-parameter model, which allows distinct transition rates for AD to gonochorism and from gonochorism to AD, was a significantly better fit for the data than a model where both transitions have an equal rate (Table 3.2), or models where transitions were restricted to one direction, either from gonochorism to AD or AD to gonochorism, indicating that in Notostraca changes in sexual system could be bidirectional. Overall, the ML model suggests that transition rates between sexual systems were high and in particular, transitions from AD to gonochorism were more than three times higher than transitions from gonochorism to AD. This result is in striking contrast to the MP results, which suggested repeated evolution in the opposite direction, to AD from gonochorism. This indicates that, once evolved, AD may be unstable and likely to revert back to gonochorism. The high rates of change across the order meant that, unlike for the MP analysis where a minimum number of transitions is inferred, ancestral sexual systems for all nodes were equally likely to be either gonochoric or AD. Virtually identical results were achieved using an ultrametric phylogeny constructed in BEAST v1.7.4 (Drummond et al. 2012) with a lognormal relaxed molecular clock both from the full dataset and with a reduced dataset containing only the mitochondrial genes COI, 12S and 16S (see Appendix 3).

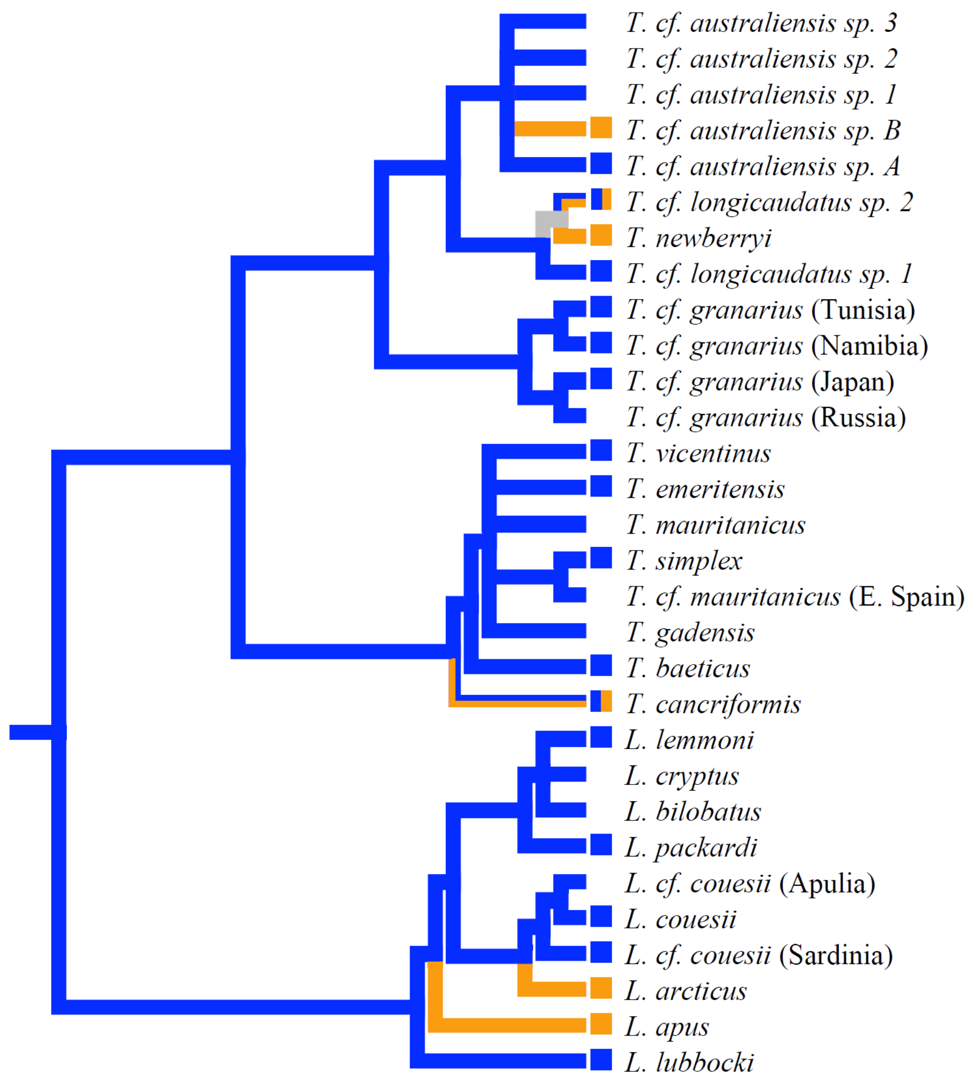


Figure 3.2: Maximum parsimony ancestral character state reconstruction of sexual systems in Notostraca. Sexual system is mapped onto the best scoring ML tree, and indicated by the colour of the square that precedes the taxon names. Blue = gonochoric, Orange = Androdioecy. STUs without squares preceding the taxon name do not have sexual system information or sexual system is equivocal. Bi-coloured squares indicate that both sexual systems are found. Branches are coloured according to MP reconstruction of ancestral sexual systems based on an unordered model with grey branches denoting an equivocal ancestral state assignment. The *Leptesheria* outgroup was removed after rooting and left uncoded for sexual system.

Table 3.2: Comparison of models of sexual system evolution in Notostraca. Models of transitions in sexual system across the best scoring ML estimation of Notostraca phylogeny compared using the ML implementation of BayesMultistate. lnL = log-likelihood of model, q_{GA} = transition rate from gonochorism to androdioecy, q_{AG} = transition rate from androdioecy to gonochorism, p = p-value for D with 1 degree of freedom comparing the restricted models to the unrestricted model.

Model	**lnL**	q_{GA}	q_{AG}	p
Unrestricted	-7.229	23.729	110.640	-
Equal rates	-10.262	5.730	5.730	0.0138
AD to gonochorism only	-16.855	0.000	6.673	< 0.0001
Gonochorism to AD only	-12.0864	1.755	0.000	0.0018

Testing the reproductive assurance hypothesis

The mean latitude of STUs where AD is found was significantly higher than for gonochoric ones (phylogenetic t-test, $t = 2.922$, $p = 0.010$, d.f. = 17) with an average latitude of 49.329 compared to 37.256. ML value for λ was estimated to be 0.508 and its inclusion significantly improved the likelihood of the model (D, $p = 0.028$, d.f. = 1) indicating that the latitude of STUs has moderate phylogenetic signal.

3.5 Discussion

Sexual system evolution

Our analyses reveal that sexual system is a highly labile trait within Notostraca. The rare sexual system AD has evolved multiple times in both notostracan genera, with MP indicating at least five independent origins across the whole group (Figure 3.2). This lability is also supported by the model based ML analysis which infers high transition rates between sexual systems. Unlike the MP analysis, the ML analysis did not resolve the ancestral sexual system for any nodes. This suggests that MP may give an oversimplified reconstruction of the evolutionary history of sexual system in Notostraca and highlights the importance of incorporating branch lengths in ancestral state reconstructions (Pagel 1999a). The ML analysis also contrasts with MP by inferring that transitions between gonochorism and AD can occur in both directions as is the case for pedunculate barnacles (Yusa et al. 2012), rather than in a unidirectional

manner as implied by MP. As there are no known biological constraints to transitions between sexual systems in either direction, there is no reason to discredit the model of sexual system evolution inferred by either method. For example, a transition from AD to gonochorism could occur as a result of selection for outcrossing over selfing, followed by the loss of testis lobes in hermaphrodites. In conclusion, although it is not possible to infer the history of sexual system change in Notostraca with current data or methods of ancestral state reconstruction, our results do display a consistent pattern of high lability of sexual system across Notostraca.

Although flexibility in sexual system has previously been shown in *T. cancriformis*, where a recent transition from gonochorism to androdioecy has occurred (Zierold et al. 2007; Zierold et al. 2009), our results demonstrate that reproductive flexibility is a general feature of Notostraca as a whole. Given the age of the order – well-preserved notostracan fossils date back to the Carboniferous (Guthörl 1934) and *Triops* and *Lepidurus* are known from the Permian and Triassic (Barnard 1929; Gall & Grauvogel-Stamm 2005; Gand et al. 1997), dating the split in the two genera to at least ~250 million years ago – it can be inferred that reproductive lability has been maintained for at least 250 million years within tadpole shrimps and may have aided their long term evolutionary persistence. Such lability of sexual system in Notostraca contrasts strongly with the dynamics of sexual system evolution in the clam shrimps of the genus *Eulimnadia*, the other branchiopod crustacean where AD is found. In *Eulimnadia*, AD evolved once and has persisted for at least 24 million years, passing through many speciation events (Weeks et al. 2009; Weeks et al. 2005; Weeks et al. 2006b; Zucker et al. 1997). The contrast in sexual system evolution between both taxa is striking considering that tadpole shrimps and clam shrimps inhabit similar ecological niches (Longhurst 1955c; Marcus & Weeks 1997), in some cases occupying the same pool (MacKay et al. 1990; Simovich 1998), and have similar life histories, producing long-lived dormant cysts that survive during adverse periods and also provide the dispersal stage (Hamer & Appleton 1991; Takahashi 1977; Weeks 1990). They also have a similar genetic mechanism of sex determination where males are recessive to hermaphrodites in AD populations (Sassaman 1989; Sassaman & Weeks 1993; Weeks et al. 2010). Finally, in both groups, hermaphrodites are morphologically derived from obligatory outcrossing females, and can either self-fertilise or outcross with males but, unlike in

plants or barnacles, are unable fertilise each other (Longhurst 1955c; Zucker et al. 1997). Within Branchiopoda therefore, superficially similar ecological backgrounds and genetic mechanisms of sex determination have resulted in two very different, but equally unusual, evolutionary outcomes for sexual system; stable and conserved in *Eulimnadia*, and labile and dynamic in Notostraca. In contrast, the flexibility of sexual system in notostracans resembles that of barnacles, where AD has evolved on several occasions in response to mate limitation, although in this system AD has evolved from a hermaphroditic ancestor rather than a gonochoric one (Høeg 1995; Yusa et al. 2012). Modelling has explained the persistence of AD in *Eulimnadia* and highlights the importance of the presence of males for limiting inbreeding depression and that a proportion of progeny produced by selfing – those that are homogametic – have reduced fitness due to the expression of sex-linked genetic load (Otto et al. 1993; Pannell 2008). In Notostraca, however, these key parameters, along with other factors that contribute to the evolution and maintenance of AD, such as male-hermaphrodite encounter rates and fecundity, have not been investigated. Further research is therefore required to ascertain whether the dynamics of sexual system evolution in Notostraca necessitate the re-evaluation of current models for the evolution of AD in animals. A lack of phylogenetic signal regarding sexual systems in Notostraca suggests that switches in sexual system occur in response to ecological factors. Notostraca share similarities in life history with many plants (Baker 1955), particularly those that exist in highly dynamic metapopulations, where colonisation of new habitats is a key feature of survival causing selection for reproductive assurance (Obbard et al. 2006b; Pannell & Barrett 1998; Pannell & Dorken 2006). In many cases selection for reproductive assurance has caused a breakdown in plant self-incompatibility systems and the evolution of self-fertile hermaphrodites, which make optimal pioneer genotypes (Igic et al. 2008). If reproductive assurance drives the evolution of self-fertile hermaphroditism, and hence AD, in Notostraca as has been hypothesised (Bernard 1891; Longhurst 1955a), we would expect AD STUs to occur in areas which have recently become available for colonisation. These predictions appear to be met in *T. cancriformis* where hermaphroditic and androdioecious populations occur in previously glaciated areas, whereas known gonochoric populations occur in what were previously unglaciated refugia (Zierold et al. 2007). Furthermore, our results support the reproductive assurance hypothesis for the whole of Notostraca as our

phylogenetically informed analysis reveals that AD notostracan STUs occur at significantly higher latitudes than gonochoric ones. These tests indicate that AD STUs occur in areas where habitat disruption through glacial cycles is more likely, providing further support for the role of colonisation advantage and reproductive assurance in a metapopulation as drivers for the evolution of AD.

Highly fluctuating population densities which result in mate limitation could also provide an advantage to lineages containing hermaphrodites through reproductive assurance (Pannell 1997; Pannell & Barrett 1998). In the case of Branchiopoda, the role of low population densities and mate limitation in the evolution of sexual systems is still poorly understood (Pannell 2008). Ephemeral ponds are often very small (Blaustein & Schwartz 2001) with strong inter-annual fluctuations in branchiopod population density (Weeks et al. 2006b), which could provide a context in which self-fertilising hermaphrodites would enjoy an increased fitness relative to females, driving the evolution of AD. Further research on Notostraca species distribution, genetic diversity, metapopulation dynamics and phylogeography, will help to understand the underlying factors behind changes in sexual systems in this group.

Taxonomic implications of notostracan phylogeny

Our results resolve the phylogenetic relationships of Notostraca, supporting the previously disputed (Mantovani et al. 2004; Murugan et al. 2002) monophyly of *Triops*. Although the four main lineages in *Triops* coincide with the four species recognised by Longhurst (1955c) in the last revision of Notostraca, our analyses support previous work revealing that each of these lineages is made up of cryptic species. Each main lineage has maintained a remarkably stable, mostly allopatric geographic distribution (only Japan has representatives of the *T. cancriformis*, *T. cf. granarius* and *T. cf. longicaudatus* complexes, and N Africa shares both *T. mauritanicus* and *T. cf. granarius*). In *Lepidurus*, *L. lubbocki* was found to be the sister species to the rest of the genus and as suggested by Mantovani et al. (2009) full species status is warranted. Further cryptic diversity was also identified in the *L. couesii* complex and given that the type locality for *L. couesii* is in N America (Rogers 2001) and the level of divergence with Apulian (S. Italy) and Sardinian lineages, we propose these latter lineages are new species. Overall, although highlighting the need for further research into Notostraca

taxonomic diversity, our phylogeny represents an excellent framework for the study of evolutionary processes within the order.

Conclusions

Our analyses show that sexual systems are extremely flexible in Notostraca, with repeated switches between gonochorism and androdioecy possibly driven by postglacial range expansions. This unconstrained, labile pattern strongly contrasts with the single origin of androdioecy in the genus *Eulimnadia*, despite the similarity of their habitats and life histories, and resembles the pattern found in barnacles. Despite the idiosyncratic evolution in these crustacean taxa, reproductive assurance in the face of fluctuating population sizes, habitat turnover or climate changes, appears to be a recurrent theme in the evolution of androdioecy. Flexibility in sexual system evolution has been maintained throughout the evolutionary history of Notostraca (over 250 my), and given the extreme morphological and life history conservatism in the group, could have facilitated their evolutionary persistence.

Chapter 4

RAD sequencing reveals chromosomal ZW sex determination in androdioecious and gonochoric populations of *Triops cancriformis*

4.1 Abstract

Genetically determined gonochorism (separate sexes) is expected to lead to selection for linkage between sex-associated loci and, ultimately, to the evolution of non-recombining, differentiated sex chromosomes. Species with variable sexual systems – such as the European tadpole shrimp, *Triops cancriformis* (which includes gonochoric, androdioecious and hermaphroditic populations) - are therefore excellent models to test the impact of transitions between sexual systems on the evolution of sex chromosomes. In order to identify sex-linked markers in *T. cancriformis*, understand its mechanisms of sex determination, and assess differences in patterns of sex-linkage between gonochoric and androdioecious populations, we used restriction site associated DNA sequencing (RAD-seq) of 47 individually bar-coded males, females and hermaphrodites from one gonochoric and one androdioecious population. We analyzed over 22 Gb of paired-end sequences and identified and scored over 1,000 high coverage novel genomic markers. A combination of presence/absence of loci, SNP segregation patterns and locus coverage depth were used to identify 63 putatively sex-linked RAD loci. We show that in *Triops cancriformis*, sex is genetically determined with a ZW chromosomal sex determination system being conserved across gonochoric and androdioecious populations but also find evidence of structural differences between the sex chromosomes of the two populations, likely reflecting different patterns of sex antagonistic selection.

4.2 Introduction

Sex chromosomes have evolved independently numerous times in plants and animals and represent an extraordinary case of evolutionary convergence with many shared features, indicating the presence of similar evolutionary forces acting across diverse lineages (Charlesworth 1996; Ellegren 2011; Fraser et al. 2004; Fraser & Heitman 2005). The main chromosomal sex determination systems are those with male heterogamety (XY), as in therian mammals and *Drosophila melanogaster*, and those with female heterogamety (ZW) such as those found in birds and some reptiles, although rarer systems have been described such as the XO XY system found in the rodent *Microtus oregoni* (Charlesworth & Dempsey 2001) and the XO XX system found in the nematode *Caenorhabditis elegans* (reviewed by Meyer 2000).

Sex chromosomes are derived from ordinary autosomes and their evolution is a highly dynamic process (Fridolfsson et al. 1998; Graves 2008; Lahn & Page 1999; Pease & Hahn 2012; Skaletsky et al. 2003). The evolution of heteromorphic sex chromosomes, where X and Y (or W and Z) are morphologically distinct, is driven by a loss of recombination. Selection favours a cessation of recombination between sex-determining and sexually antagonistic loci, during the evolution of separate sexes from hermaphrodites, or during the evolution of genetic sex determination from environmental sex determination (Bachtrog et al. 2011; Bergero & Charlesworth 2009; Charlesworth et al. 2005). This lack of recombination causes "degeneration" of the chromosome specific to the heterogametic sex through a variety of processes, ultimately leading to morphologically distinct pairs of chromosomes (reviewed by Bachtrog 2013). Sex chromosomes are not always heteromorphic, however, with homomorphic sex chromosomes having evolved in many groups and for a variety of reasons (Stöck et al. 2011). In many cases sex chromosomes are homomorphic due to a recent origin, meaning there has been insufficient time for degeneration, as in the incipient sex chromosome system found in the strawberry, *Fragaria virginiana* (Spigler et al. 2008; Spigler et al. 2010). Alternatively, occasional recombination between sex chromosomes, as in tree frogs (Guerrero et al. 2012; Stöck et al. 2011), or the evolution of sex biased gene expression which alleviates sexually antagonistic selection, as in emus (Vicoso et al. 2013), has reduced genetic degeneration of the Y or

W sex chromosome preventing morphological divergence. Studying these young, and or, unusual sex chromosome systems in non-model organisms has shed light on a range of processes that take place during sex chromosome evolution. Sexual system and reproductive mode may also affect the course of sex chromosome evolution and, although the evolution of separate sexes from hermaphrodites has been implicated in the initial stages of sex chromosome evolution, sex chromosomes in labile sexual systems have been little studied (Pires-daSilva 2007).

Notostracans, or tadpole shrimps, are globally distributed branchiopod crustaceans that have shown a remarkably stable morphology for over 300 million years, despite at least two bouts of diversification resulting in around 40 extant species (Chapter 2). In contrast to their extreme morphological conservatism, tadpole shrimp show diverse sexual systems including gonochorism (dioecy), hermaphroditism and androdioecy (Sassaman 1991). Androdioecy (AD), the occurrence of self-fertile hermaphrodites and males (Pannell 2002a; Pannell 2002b; Weeks et al. 2006a), is particularly interesting as a potential transitional stage in the evolution of dioecy from hermaphroditism, or vice versa (Barrett 2010; Weeks 2012). Although AD has been described only in a few animal groups (Weeks et al. 2006a), it is found in three genera of branchiopod crustaceans, the spinicaudatan *Eulimnadia* (Sassaman & Weeks 1993) and the notostracans *Triops* and *Lepidurus* (See Chapter 3 for a review of AD in tadpole shrimps). AD is the main mode of reproduction in *Eulimnadia*, probably having evolved once (Weeks et al. 2009; Weeks et al. 2006b). In contrast, AD is likely to have evolved from dioecy multiple times during the evolutionary history of Notostraca (Chapter 3). This labile nature of sexual system within Notostraca, with multiple transitions between separate and combined sexes, makes this group an excellent model to understand the genomic impacts of such transitions.

The European tadpole shrimp, *Triops cancriformis*, has particularly attractive features to investigate the genomic basis of transitions in sexual system as it is amongst the few species in which three sexual systems (AD, hermaphroditism and the putatively ancestral state, gonochorism) occur in different populations (Sassaman 1991; Zierold et al. 2007). In addition, phylogeographic analysis revealed that European populations are closely related and sexual system variation has evolved recently (Zierold et al.

2007). These unique attributes therefore allow a direct comparison of the genetic details of each sexual system without confounding phylogenetic effects.

The mechanisms of sex determination have been little studied in branchiopod crustaceans. In two of the androdioecious species, *Eulimnadia texana* and *Triops newberryi*, sex determination is genetic with maleness a recessive character to hermaphroditism, suggesting a ZW sex chromosome system (Otto et al. 1993; Sassaman 1989; Sassaman 1991; Sassaman & Weeks 1993). In these species, two types of hermaphrodites are found, amphigenic (assumed ZW, which when selfed produce one quarter male and three quarters hermaphrodite offspring) and monogenics (assumed WW), which, when selfed, produce exclusively monogenic hermaphrodite offspring. Sex determination is also ZW in the gonochoric anostracan *Artemia franciscana* (De Vos et al. 2013). In contrast, in the cyclical parthenogen *Daphnia* sex is environmentally determined (Kato et al. 2011). In *T. cancriformis*, the paucity of available microsatellite loci and other genomic resources has hampered the identification of sex-linked markers (Cesari et al. 2004; Zierold et al. 2009) and chromosomes are small with no evidence of heteromorphism (Ombretta et al. 2005; Trentini 1976).

Here, we use restriction site associated DNA sequencing (RAD-seq) to investigate the mechanism of sex determination in gonochoric and androdioecious populations of *Triops cancriformis*. RAD-seq is a powerful and affordable tool to identify and score thousands of genome-wide single nucleotide polymorphisms (SNPs) in many individuals, even for non-model organisms (Baird et al. 2008; Davey & Blaxter 2010; Davey et al. 2011; Narum et al. 2013). The suitability of RAD-seq for characterising sex chromosomes has been demonstrated in frogs (Bewick et al. 2013) and zebrafish (Anderson et al. 2012) and additionally, has been recently used to characterize the supergenes determining social organization in fire ants (Wang et al. 2013), a genomic system with similar features to sex-chromosomes.

We scored individually bar-coded males, females and hermaphrodites from one androdioecious and one gonochoric population and used presence/absence of loci, SNP segregation patterns and read coverage to identify putatively sex-linked markers. This work confirms for the first time the occurrence of genetic sex determination in

this species with a ZW sex chromosome system. The conservation of some Z and W-linked markers for both populations indicates a common evolutionary origin for their sex determination system, but population differences in sex-linkage patterns highlights the occurrence of substantial rearrangements in the genomic architecture of both sex chromosomes.

4.3 Materials and Methods

Samples, rearing and individual sexing

Sediment samples containing *T. cancriformis* diapausing cysts were obtained from Espolla temporary pond in Spain (ESP) and Königswartha pond 12 in Germany (KOE), which harbor gonochoric and androdioecious populations respectively (for details see Zierold et al. (2007) and Zierold et al. (2009)). Sediment subsamples were scattered at the bottom of tanks and purified water added to encourage cyst hatching. Tanks were kept under 24 h illumination at 20°C and the resulting hatchlings were reared on *Triops* food (obtained from www.triops.es). Mature individuals were fixed in 100% ethanol and sexed prior to use in the RAD-seq protocol. Individuals not bearing ovisacs were considered males, whereas individuals bearing ovisacs were regarded as females or hermaphrodites, depending on their population of origin (see Chapter 3 for details). Overall, 47 individuals were selected for RAD sequencing: 14 males and 12 females from ESP, and 15 hermaphrodites and 6 males from KOE (Table 4.1).

Table 4.1: Sample information for all individuals included in the four RAD libraries. In total 12 barcodes were used and split across the 4 RAD libraries.

Sample I.D.	Adapter barcode	RAD library I.D.	Collection location	Sex
ESP_F10	TCAGTGCT	Lib_2	Espolla, Spain	Female
ESP_F11	TCTCGCTC	Lib_2	Espolla, Spain	Female
ESP_F13	TGACATAC	Lib_2	Espolla, Spain	Female
ESP_F14	TGTGACTG	Lib_2	Espolla, Spain	Female
ESP_F15	ACACTGAC	Lib_3	Espolla, Spain	Female
ESP_F17	ACGTCTAC	Lib_3	Espolla, Spain	Female
ESP_F19	AGAGTCGA	Lib_3	Espolla, Spain	Female
ESP_F20	TACGCTCG	Lib_2	Espolla, Spain	Female
ESP_F2	GCTACAGC	Lib_2	Espolla, Spain	Female
ESP_F4	GTCACTCA	Lib_2	Espolla, Spain	Female
ESP_F5	GTGTACTG	Lib_2	Espolla, Spain	Female
ESP_F8	TAGCGACG	Lib_2	Espolla, Spain	Female

ESP_M10	CATGTCGT	Lib_3	Espolla, Spain	Male
ESP_M11	CGATCAGC	Lib_3	Espolla, Spain	Male
ESP_M1	GATCGTGA	Lib_4	Espolla, Spain	Male
ESP_M12	CGTACTCG	Lib_3	Espolla, Spain	Male
ESP_M13	CTAGCTCT	Lib_3	Espolla, Spain	Male
ESP_M14	CTGATGCT	Lib_3	Espolla, Spain	Male
ESP_M15	GACTACGA	Lib_4	Espolla, Spain	Male
ESP_M16	GTCACTCA	Lib_4	Espolla, Spain	Male
ESP_M2	AGTCTGCT	Lib_3	Espolla, Spain	Male
ESP_M3	GCATGTGC	Lib_4	Espolla, Spain	Male
ESP_M5	GCTACAGC	Lib_4	Espolla, Spain	Male
ESP_M7	ATCGCAGC	Lib_3	Espolla, Spain	Male
ESP_M8	ATGCTGTA	Lib_3	Espolla, Spain	Male
ESP_M9	CACATGAC	Lib_3	Espolla, Spain	Male
KOE_12_H10	ACGTCTAC	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H18	AGAGTCGA	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H19	AGTCTGCT	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H20	ATCGCAGC	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H21	ATGCTGTA	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H2	ACACTGAC	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H23	CACATGAC	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H24	CATGTCGT	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H25	CGATCAGC	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H26	CGTACTCG	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H27	CTAGCTCT	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H28	GACTACGA	Lib_2	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H29	GATCGTGA	Lib_2	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H30	GCATGTGC	Lib_2	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_H4	CTGATGCT	Lib_1	Pond 12, Königswartha, Germany	Hermaphrodite
KOE_12_M10	TGACATAC	Lib_4	Pond 12, Königswartha, Germany	Male
KOE_12_M1	GTGTACTG	Lib_4	Pond 12, Königswartha, Germany	Male
KOE_12_M3	TACGCTCG	Lib_4	Pond 12, Königswartha, Germany	Male
KOE_12_M4	TAGCGACG	Lib_4	Pond 12, Königswartha, Germany	Male
KOE_12_M5	TCAGTGCT	Lib_4	Pond 12, Königswartha, Germany	Male
KOE_12_M6	TCTCGCTC	Lib_4	Pond 12, Königswartha, Germany	Male

Sample preparation and DNA extraction

Total genomic DNA was extracted using the DNAeasy Blood and Tissue Kit (Qiagen) including a digestion step with RNase A. DNA was extracted from ~1cm long tail sections of *T. cancriformis* individuals. Digestive tracts were dissected out and the gut cavity flushed with 100% ethanol to minimise co-amplification of bacteria or algae. Two separate elutions of 100 µl were combined and concentrated by evaporation to

80 µl in a vacudryer (40 minutes at 30°C). DNA quality was assessed on 0.8% agarose gels and precise quantification carried out using fluorimetry (Quant-iT PicoGreen Kit, Invitrogen). Only samples with high molecular weight DNA and concentrations of a least 18 ng μl^{-1} were included in the RAD analysis.

RAD library preparation and sequencing

Four paired-end RAD libraries were prepared by the NERC/NBAF facility at the University of Edinburgh (The GenePool) following the protocol of Baird et al. (2008) with some modifications (Ogden et al. 2013). Briefly, for each of the 47 individuals 1 µg total genomic DNA was digested using the restriction enzyme *Sbf*1 followed by ligation of a bar-coded P1 RAD adapter (see Table 4.1 for barcode sequences and sample details). The samples were then split into four libraries before sonic shearing, size selection (300 bp – 700 bp) and ligation of non-model P2 RAD adaptors. Libraries were then PCR amplified, quantified and sequenced in separate flow cells on an Illumina HiSeq 2000 platform with 100 base paired-end chemistry (v1). In total, two sequencing runs were carried out and the raw reads concatenated in silico for each library.

Generation of RAD loci and SNP calling

Sequencing quality for the first and second end reads of each RAD library was assessed with FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc) (Appendix 4, Figures 1 and 2). RADtools (Baxter et al. 2011) was used to de-multiplex the raw first and second end reads and to discard any sequences not in the expected RAD format (i.e. those not containing the expected Sbf1 restriction site or one of the specified barcodes). Sequences were trimmed to 70 bp after removal of the barcode sequence and filtered based on quality with reads containing any position with a quality score less than 20 discarded. Fuzzy matching, which allows up to 1 bp deviation from the expected restriction site and barcode sequence, was used to account for sequencing error (Baxter et al. 2011). Files were outputted in fastq format to retain per base quality scores for later mapping of reads.

Generation of candidate RAD alleles and loci and SNP calling was performed using components of the Stacks pipeline v0.99993 (Catchen et al. 2011). Firstly, PCR duplicates from the amplification stage of the library preparation were removed using

the Stacks script *clone_filter.pl* which identifies groups of sequences with an identical second end read and removes all but one copy. This step is necessary as the random shearing stage of RAD library preparation means that very few exactly matching second end reads are expected for each locus by chance, therefore reads with an identical second end i.e. perfectly overlapping with no overhang, are likely to be due to duplication during PCR. Removal of PCR duplicates is important as the Stacks SNP calling model relies on read counts of alternative alleles and PCR duplicates do not represent independent samples from the pool of genomic DNA (Hohenlohe et al. 2013). Additionally, removal of PCR duplicates reduces the effects of GC bias during PCR amplification (Davey et al. 2012). For each individual, Ustacks was then used to cluster first end reads into candidate alleles and loci based on sequence similarity and call SNPs based on a maximum likelihood diploid genotyping model (Hohenlohe et al. 2010). Alleles with less than 15 identical reads were discarded and up to 6 mismatches between alleles allowed when calling candidate RAD loci. Lumberjack stacks – candidate RAD loci with coverage 2 standard deviations above the average – were removed with the '-t' option of Ustacks as they likely represent repetitive elements (Catchen et al. 2013b; Catchen et al. 2011). A catalogue of *T. cancriformis* RAD loci found across both populations was then constructed using Cstacks, also allowing up to 6 mismatches between RAD loci from different individuals. Consensus sequences for each RAD locus were extracted from the catalogue in fasta format using a custom awk script. Sstacks was then used to identify which locus / haplotype combinations were found in each individual and the Stacks 'populations' program used to extract haplotypes for every locus in the catalogue for each individual and generate population level summary statistics.

Coverage analysis

To obtain per individual coverage information for each candidate RAD locus, raw first end reads from each individual were mapped back to the set of consensus sequences from the catalogue using Stampy (Lunter & Goodson 2011) with default settings. SAMtools (Li et al. 2009) was then used to generate a pileup from which coverage information (per locus, per individual) was extracted using a custom awk script. This generated a spreadsheet that could be manipulated in Excel to filter loci based on

expectations of presence / absence in particular groups of individuals and to identify patterns of coverage across loci.

As depth of coverage between individually bar-coded samples in RAD libraries is often highly heterogeneous (e.g. Richards et al. 2013; Wagner et al. 2013) normalisation of read counts is essential for coverage differences between individuals to be comparable. To allow for variation in coverage between individuals we normalised per locus coverage for each individual relative to the median coverage of putatively autosomal loci for that individual. Median autosomal coverage was estimated for each individual using loci found in all 47 individuals with a minimum of 20x coverage. This excludes population specific loci, and potentially sex specific loci, but ensures a robust sample of loci are used to estimate autosomal coverage for *T. cancriformis* loci as opposed to low coverage contaminant loci.

Confirmation of orthology between sex-linked loci

Stacks assembles RAD alleles and loci based on sequence similarity (allowing a user specified number of mismatches between reads of the same allele and locus) and expects all loci to be of identical length (Catchen et al. 2011). Indels can therefore cause splitting of alleles into separate loci. Alternatively, divergent but orthologous alleles may be split into different loci if the level of divergence exceeds the mismatch parameter specified when building putative RAD loci. To check for orthology between putative RAD loci we attempted to align all the first end consensus sequences from our catalogue of RAD loci using CodonCode Aligner v4.1.1 (CodonCode Corporation). Sex-linked loci that align to another locus in the catalogue were then identified.

4.4 Results

Sequencing

A total of 114,304,783 read pairs were obtained from the four RAD libraries totalling 22.9 Gb of sequence (Table 4.2). After trimming, removal of poor quality reads, removal of reads with ambiguous barcodes and filtering of PCR duplicates we retained 28,986,269 read pairs 70 bp long totaling 4.06 Gb of sequence. Coverage varied between individuals from 246,930 (KOE_12_H2) to 1,390,021 reads (ESP_M11) (Table 4.3).

Table 4.2: Summary of sequencing results and quality control for the 4 RAD libraries.

Library	Number of read pairs	Not in RAD format	No Sbf1 site	No matching barcode	Quality score below 20	PCR duplicates*	Retained read pairs
Lib_1	25664353	6831566	693342	18358	5374735	5922130	6824222
Lib_2	36108312	9864384	1166594	9839	7266784	8714967	9085744
Lib_3	23500808	6274717	808301	15027	5043105	4334274	7025384
Lib_4	29031310	7796305	980922	9620	5865785	8327759	6050919
Total	114304783	30766972	3649159	52844	23550409	27299130	**28986269**
% Reads		26.92	3.19	0.05	20.60	23.88	**25.36**

* PCR duplicates refers to the number of duplicates removed from the cleaned reads, not the number of duplicates in the unprocessed data.

Table 4.3: Sequencing results for all 47 individuals included in the RAD analysis. Median coverage is shown for 800 loci found in all individuals with a minimum of 20x coverage that are assumed to be autosomal.

Sample I.D.	Library	Number of reads (clones included)	Number of reads (clones removed)	% PCR clones	Median coverage of 800 assumed autosomal loci (Inter quartile range)
ESP_F10	Lib_2	1017151	571231	56.16	327 (284 – 369)
ESP_F11	Lib_2	1536925	1023316	66.58	307 (259 – 348)
ESP_F13	Lib_2	999051	601611	60.22	278 (242 – 317)
ESP_F14	Lib_2	825697	491439	59.52	260 (220 – 297)
ESP_F15	Lib_3	619632	402848	65.01	239 (208 – 268)
ESP_F17	Lib_3	856391	567920	66.32	286 (246 – 322)
ESP_F19	Lib_3	575059	374487	65.12	232 (198 – 264)
ESP_F2	Lib_2	1510484	804893	53.29	430 (369 – 493)
ESP_F20	Lib_2	1960923	954832	48.69	527 (460 – 607)
ESP_F4	Lib_2	688107	398325	57.89	247 (213 – 281)
ESP_F5	Lib_2	2513056	1103002	43.89	640 (557 – 747)
ESP_F8	Lib_2	2726913	1214613	44.54	658 (571 – 755)
ESP_M1	Lib_4	855669	380230	44.44	224 (186 – 263)
ESP_M10	Lib_3	621032	390134	62.82	233 (202 -258)
ESP_M11	Lib_3	2239455	1390021	62.07	469 (410 – 531)
ESP_M12	Lib_3	1042976	651546	62.47	322 (280 – 357)
ESP_M13	Lib_3	1365719	749766	54.90	457 (362 – 521)
ESP_M14	Lib_3	754455	459783	60.94	278 (246 – 310)
ESP_M15	Lib_4	1047355	457514	43.68	246 (206 – 291)
ESP_M16	Lib_4	1483267	611595	41.23	353 (296 – 412)
ESP_M2	Lib_3	911196	561635	61.64	323 (285 – 363)
ESP_M3	Lib_4	827404	396570	47.93	186 (155 – 220)
ESP_M5	Lib_4	1054979	448650	42.53	281 (235 – 333)

ESP_M7	Lib_3	835119	522363	62.55	308 (262 – 342)
ESP_M8	Lib_3	941464	562073	59.70	346 (306 – 381)
ESP_M9	Lib_3	597160	392808	65.78	204 (176 – 228)
KOE_12_H10	Lib_1	928315	490450	52.83	277 (222 – 334)
KOE_12_H18	Lib_1	894009	474034	53.02	259 (210 – 306)
KOE_12_H19	Lib_1	1457456	795620	54.59	334 (276 – 411)
KOE_12_H2	Lib_1	418394	246930	59.02	139 (111 – 172)
KOE_12_H20	Lib_1	968154	500571	51.70	278 (224 – 342)
KOE_12_H21	Lib_1	1419720	748391	52.71	356 (288 – 424)
KOE_12_H23	Lib_1	1059799	536921	50.66	289 (235 – 348)
KOE_12_H24	Lib_1	782487	404452	51.69	240 (190 – 289)
KOE_12_H25	Lib_1	1318478	671620	50.94	348 (281 – 419)
KOE_12_H26	Lib_1	970011	497629	51.30	286 (231 – 344)
KOE_12_H27	Lib_1	1222727	704986	57.66	263 (212 – 319)
KOE_12_H28	Lib_2	626983	382946	61.08	197 (171 – 224)
KOE_12_H29	Lib_2	2501148	1068521	42.72	617 (542 – 699)
KOE_12_H30	Lib_2	894273	471015	52.67	298 (262 – 335)
KOE_12_H4	Lib_1	1306802	752618	57.59	287 (235 – 350)
KOE_12_M1	Lib_4	1106998	481117	43.46	273 (231 – 317)
KOE_12_M10	Lib_4	1337745	542403	40.55	335 (286 – 388)
KOE_12_M3	Lib_4	1983935	789837	39.81	454 (383 – 527)
KOE_12_M4	Lib_4	1261096	537857	42.65	313 (265 – 362)
KOE_12_M5	Lib_4	1440629	611423	42.44	343 (292 – 399)
KOE_12_M6	Lib_4	1979601	793723	40.10	456 (394 – 526)

Assembly of RAD loci and SNP calling

Stacks recovered 20,902 candidate RAD loci, from here on in referred to as “the catalogue”. However, many of these loci are found only in low coverage and in single individuals, suggesting they may represent contamination from non-*Triops* genomic DNA as we would expect most RAD loci to be found in the majority of individuals given sufficient coverage. In contrast, loci found in all individuals have consistently high coverage, accounting for 53% of the total raw reads mapped back to the catalogue (Figure 4.1). Specifically, only 800 RAD loci were found in all individuals and sequenced to a minimum coverage depth of 20x. These 800 loci have a median depth of coverage of 299x. An additional 907 loci are recovered when filtering for loci found in all individuals (min. 20x coverage) from at least one of the two populations – combined, these 1707 loci account for 83% off the total number of first end reads mapped back to the catalogue (seen in the increase in coverage between 21 and 30 in Figure 4.1). A subset of loci were found to be population specific with 205 loci found only in the ESP

population (loci found in all ESP individuals with a minimum coverage of 20x and never in KOE individuals) and 114 loci found only in KOE. In total 1221 loci were found in all individuals from the ESP population (minimum 20x coverage) and 1288 loci were found in all individuals from the KOE population. As this filtering process potentially removes sex chromosome specific loci, subsequent searches for sex-linked loci were carried out using the complete catalogue of 20,902 loci with specific filtering criteria, specified in the sections below.

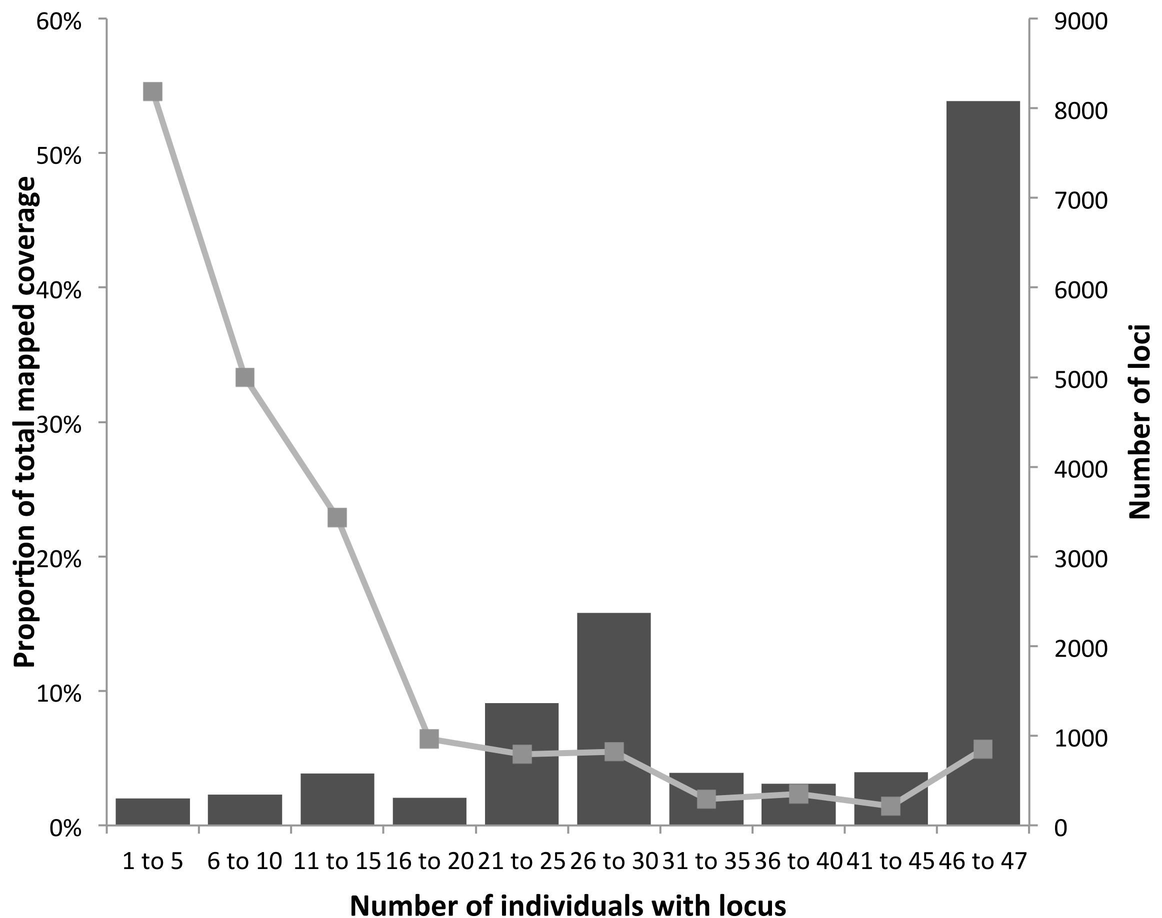


Figure 4.1: Distribution of raw library coverage (4 libraries pooled) amongst 20,902 loci in the Stacks catalogue of *T. cancriformis* RAD loci. We compared the number of individuals each candidate RAD locus was found in (x axis) to the proportion of the total mapped coverage those loci correspond to (primary y axis, dark grey bars), and the number of loci (secondary y axis, grey squares).

Population level summary stats were generated with the Stacks program 'populations'. We included loci with a Stacks genotype call in all individuals of at least one of the two populations with a minimum coverage of 20x, generating a dataset of 1559 loci. Population genetic statistics for the two populations therefore potentially exclude sex chromosome specific loci. F_{ST} between the two populations was estimated to be 0.71,

in agreement with a previous study using microsatellites which estimated F_{ST} between ESP and KOE to 0.72 (Zierold et al. 2009), and indicating substantial genetic differentiation between the two populations. The level of polymorphism within the two populations also varied with 351 loci in ESP containing at least one SNP, compared to just 122 in KOE.

Identification of candidate sex-linked RAD loci

There are two types of polymorphic RAD loci: presence / absence loci and loci that contain one or more single nucleotide polymorphism (SNP) within the sequenced read (Davey & Blaxter 2010). Fully sex-linked presence / absence patterns of RAD loci can be caused by one of several factors: (1) the lack of a homolog on the corresponding sex chromosome, (2) a mutation in the restriction enzyme cut site of one of the sex chromosomes, (3) divergent evolution and lack of recombination between homologous loci leading to sequence divergence in excess of the maximum specified for the Stacks assembly causing loci to be erroneously split and (4) the presence of indels causing Stacks to erroneously split loci despite otherwise high sequence similarity.

We used presence / absence of RAD loci, segregation of SNP markers and locus coverage depth to identify candidate sex-linked RAD loci and to test the expectations of a ZW chromosomal sex determination system by making comparisons between females and males from ESP and hermaphrodites and males from KOE (summarised in Table 4.4). Sex-linked markers were identified separately in ESP (gonochoric) and KOE (androdioecious) before being compared across the two populations in order to detect potential differences in either the sex determination system or sex chromosome structure associated with the evolution of androdioecy. We also compared the loci amongst themselves to identify loci that had been incorrectly split by Stacks due to indels, or divergence over the 6 nucleotide threshold.

Table 4.4: Predictions of RAD locus presence-absence, locus coverage and SNP segregation patterns for fully sex linked RAD loci given a ZW chromosomal sex determination system**.**

Population	**Espolla (ESP)**		**Königswartha (KOE)**		
Sexual system	gonochoric		androdioecious		
Sex	female	male	hermaphrodite		male
			monogenic	amphigenic	
Genetic model	ZW	ZZ	WW	ZW	ZZ
W-linked loci – pattern	present	absent	present	present	absent
W-linked loci – coverage*	50%	-	100%	50%	-
Z-linked loci – pattern	present	present	absent	present	present
Z-linked loci – coverage*	50%	100%	-	50%	100%

*Coverage relative to autosomal loci found in all individuals in both ESP and KOE.

1) W specific RAD loci and identification of monogenic hermaphrodites (WW)

If the sex determination system in *T. cancriformis* is of the type ZW, where females are the heterogametic sex, we expect a subset of markers to be W specific. In ESP these loci would be present in all females and absent in all males. We therefore searched for loci in the catalogue that had a minimum of 20x coverage in all females and no more than 5x coverage in all males. A non-zero coverage value was used because 4 loci had a few reads in some individuals, possibly due to small quantities of DNA cross contamination during DNA extraction or sample barcoding. These criteria recovered 20 candidate W specific RAD loci in ESP, indicating that they are located on a non-recombining region of a W sex chromosome. Filtering the opposite way, for loci with no more than 5x coverage in all females but present in all males with a minimum coverage of 20x, as would be expected for an XY chromosomal sex determination system, recovered no RAD loci. Using the same criteria to find W specific loci in KOE hermaphrodites (present in all hermaphrodites (min. 20x coverage) but no more than 5x coverage in all males) recovered only two loci with no reads in any males, both of which are also W specific in ESP. The remaining 18 loci that are W specific in ESP were also found in KOE but were present in both males and hermaphrodites indicating that

in the androdioecious population these loci are not completely sex-linked and are either found on autosomes or in a pseudoautosomal region of the sex chromosomes.

A further expectation of candidate W linked loci found in females from the gonochoric population is that they should have approximately half the coverage of autosomal loci due to their hemizygous state. The 20 candidate W specific loci show some variability in normailsed coverage (Figure 4.2), as expected by known variability in coverage across RAD loci (Davey et al. 2012). However, they are all distributed around 0.5 meeting expectations for hemizygous loci.

In KOE predicted coverage patterns for W specific loci are less straightforward due to the expectation of the presence of both monogenic (WW) and amphigenic (ZW) hermaphrodites as have been identified in other androdioecious branchiopod taxa such as *Triops newberryi* (Sassaman 1991) and *Eulimnadia texana* (Sassaman & Weeks 1993; Weeks et al. 2010). We therefore used patterns of coverage in the two candidate W specific loci found in KOE to identify ZW and WW hermaphrodites. WW hermaphrodites are expected to have coverage equivalent to autosomal loci in both candidate W specific loci, whereas ZW hermaphrodites are expected to have approximately half autosomal coverage in both loci. Plotting normalised coverage for the two candidate W specific loci against each other for all 15 hermaphrodites resulted in two distinct clusters of points, one around 0.5x median autosomal coverage in both loci and a second around 1x median autosomal coverage, confirming the presence of both WW and ZW hermaphrodites in our dataset. In total we identified 11 ZW hermaphrodites and 4 WW hermaphrodites (Figure 4.3).

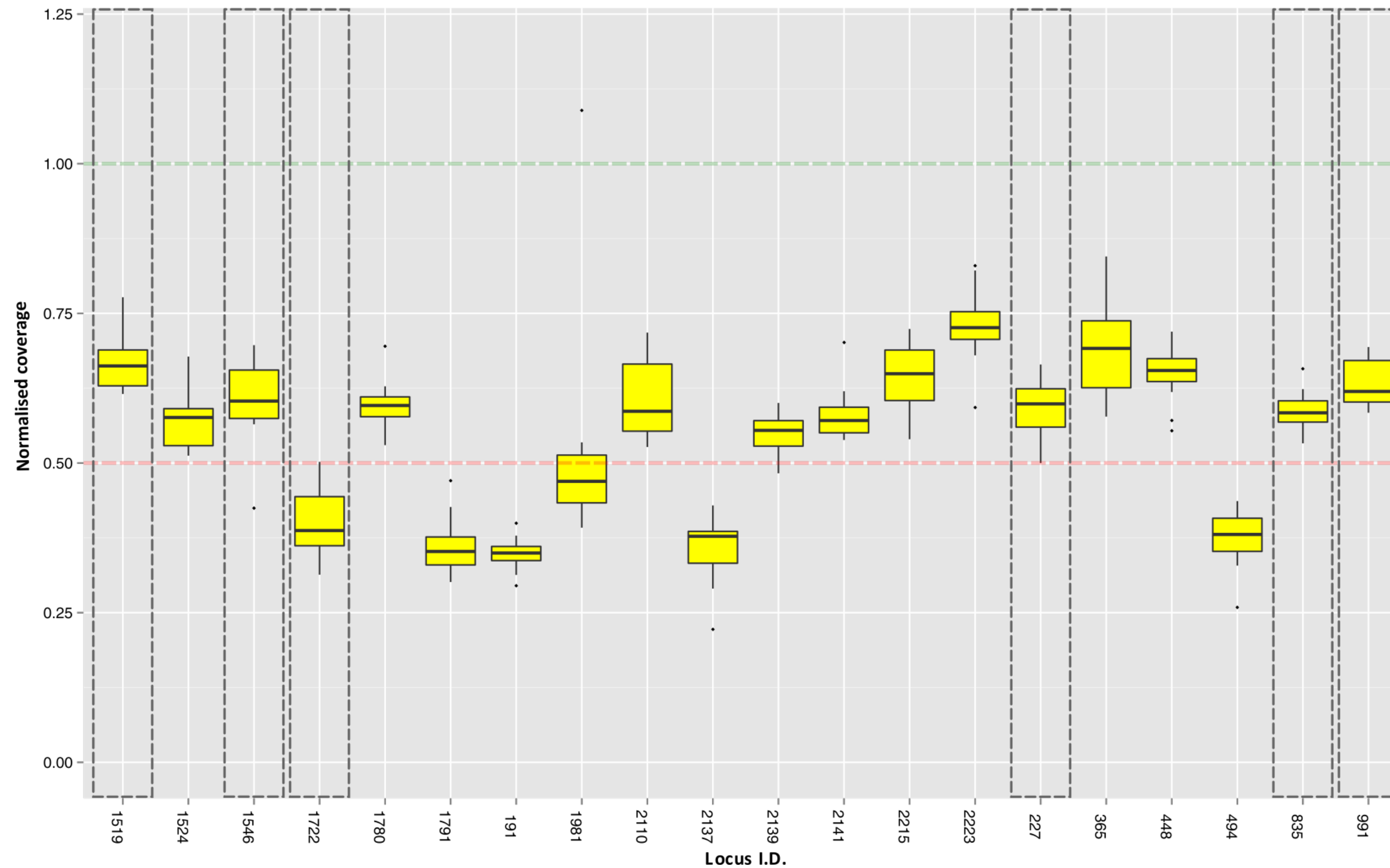


Figure 4.2: 20 W specific loci in ESP females. Box plots show the average normalised coverage of the 20 identified W specific loci in ESP females. Green dotted line shows expected coverage for autosomal loci, red dotted line shows expected coverage for hemizygous loci. Loci surrounded by grey dotted lines have ESP Z specific gametologs (Table 4.5). All loci were identified based on patterns of presence / absence.

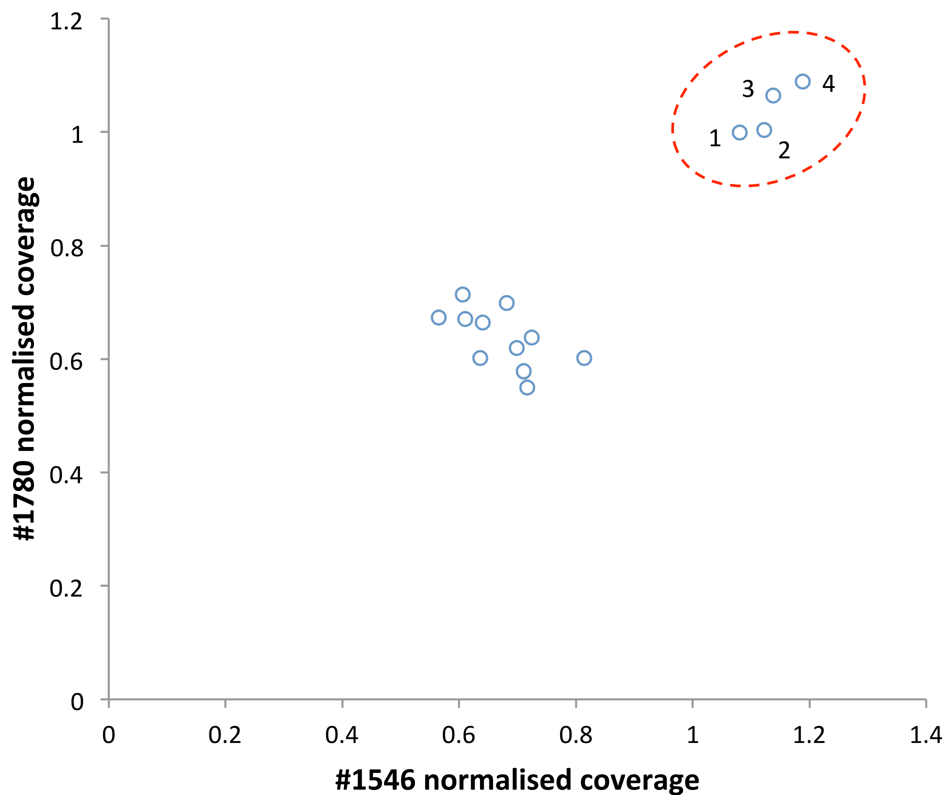


Figure 4.3: Identifying KOE monogenic (WW) hermaphrodites. Normalised coverage of #1780 against #1546, the two W specific loci found in KOE, for the 15 KOE hermaphrodites. Two clouds can be seen corresponding to either amphigenic (ZW) hermaphrodites which have approximately half autosomal coverage in both loci as expected for hemizygous loci, and monogenic (WW) hermaphrodites (circled in red) which have equivalent coverage to autosomal loci. In total 4 monogenic (WW) hermaphrodites are identified: 1 – KOE_12_H29, 2 – KOE_12_H25, 3 – KOE_12_H10, 4 – KOE_12_H21.

2) Z specific RAD markers

In addition to W specific loci, given a ZW chromosomal sex determination system, we might expect a subset of markers to be Z specific. In KOE, identification of these loci was relatively straightforward due to the presence of monogenic hermaphrodites (WW) (identified in the previous section) in which Z specific loci are expected to be absent. Filtering the catalogue for loci present in all males (ZZ) and all amphigenic hermaphrodites (ZW), with a minimum of 20x coverage, but absent in all monogenic hermaphrodites (WW) (no more than 5x coverage) recovered 11 candidate Z specific loci. These loci have approximately 0.5 mean normalised coverage in amphigenic

hermaphrodites (ZW) compared to 1 in males (ZZ), further confirming their Z specific status in KOE (Figure 4.4).

Just two of the 11 candidate Z specific loci in KOE are found in ESP and only one of them, #1913, shows the pattern of coverage expected for a Z specific locus with approximately half mean normalised coverage in females (WZ) compared to males (ZZ) (data not shown). The other locus that is Z specific in KOE (#4909), and also found in ESP, has patterns of presence / absence incompatible with sex linkage (data not shown).

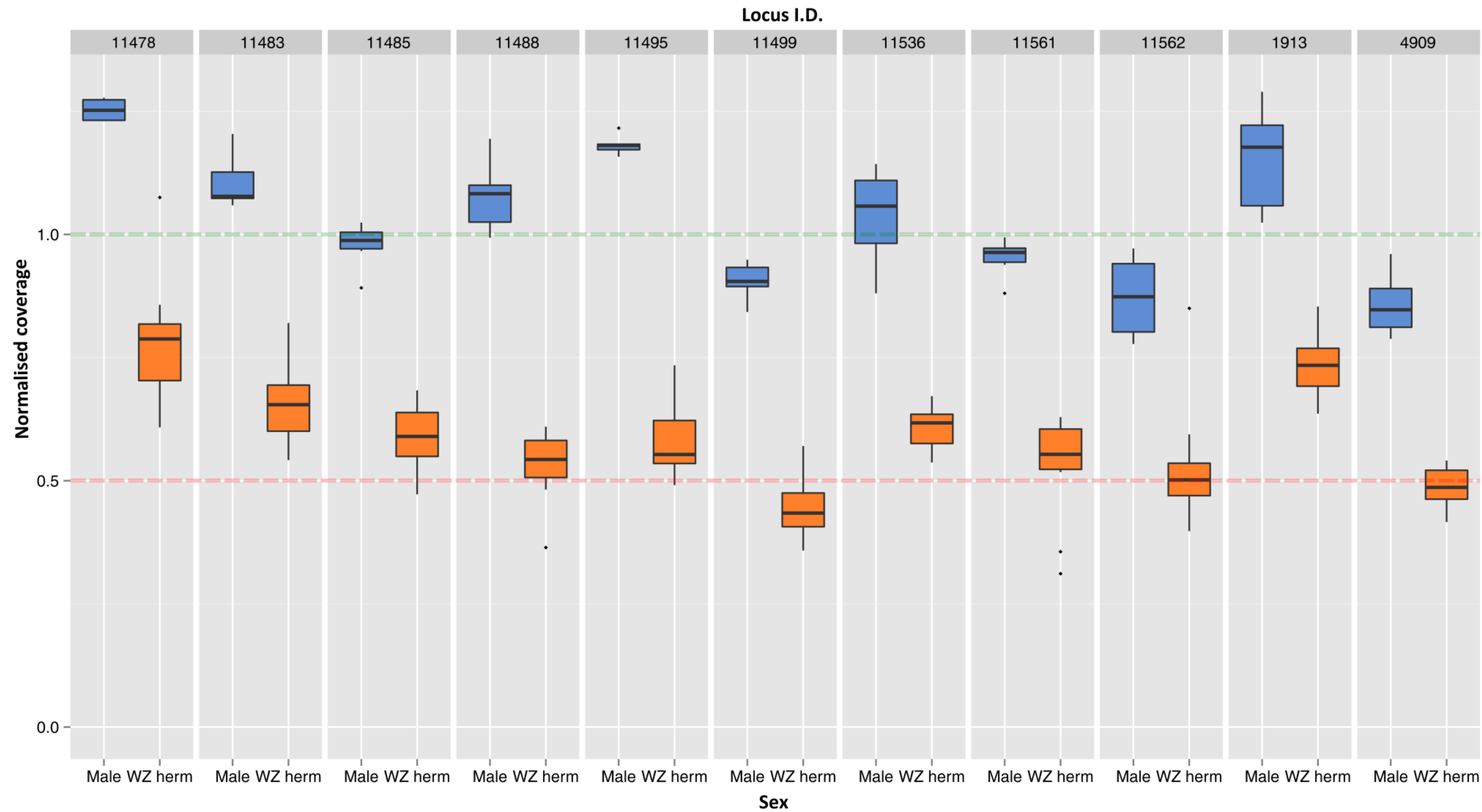


Figure 4.4: Male vs. amphigenic hermaphrodite normalised coverage for 11 candidate Z specific loci in KOE. Box plots show normalised coverage for 11 Z linked markers in KOE males (ZZ) (blue) and amphigenic hermaphrodites (ZW) (orange). The green dotted line shows expected coverage for autosomal loci and the red dotted line shows expected coverage for hemizygous loci. These markers are absent in monogenic (WW) KOE hermaphrodites. #1913 is also found in ESP and shows similar patterns of coverage with females (ZW) having approximately half coverage vs. males (ZZ).

In the gonochoric population (ESP), patterns of presence / absence are uninformative to identify further Z specific loci as both males (ZZ) and females (ZW) are expected to carry the Z chromosome. However, candidate Z specific loci can be identified by comparing mean per locus read depth between males and females, as these loci are expected to have approximately half mean normalised coverage in ESP females (ZW) compared to ESP males (ZZ). To identify candidate Z specific loci in ESP, a subset of 1180 loci were selected that are found in all ESP individuals (minimum of 20x coverage and mean normalised coverage in either males or females no greater than 3, to exclude loci likely to be in repetitive or duplicated regions). Plotting mean normalised coverage of females vs. males for each of these loci (Figure 4.5) shows that the majority of loci have approximately equal coverage in males and females as is expected for autosomal loci. A subset, however, appear to deviate from this 1:1 ratio, more closely fitting the 0.5:1 ratio (female:male) expected for Z specific loci.

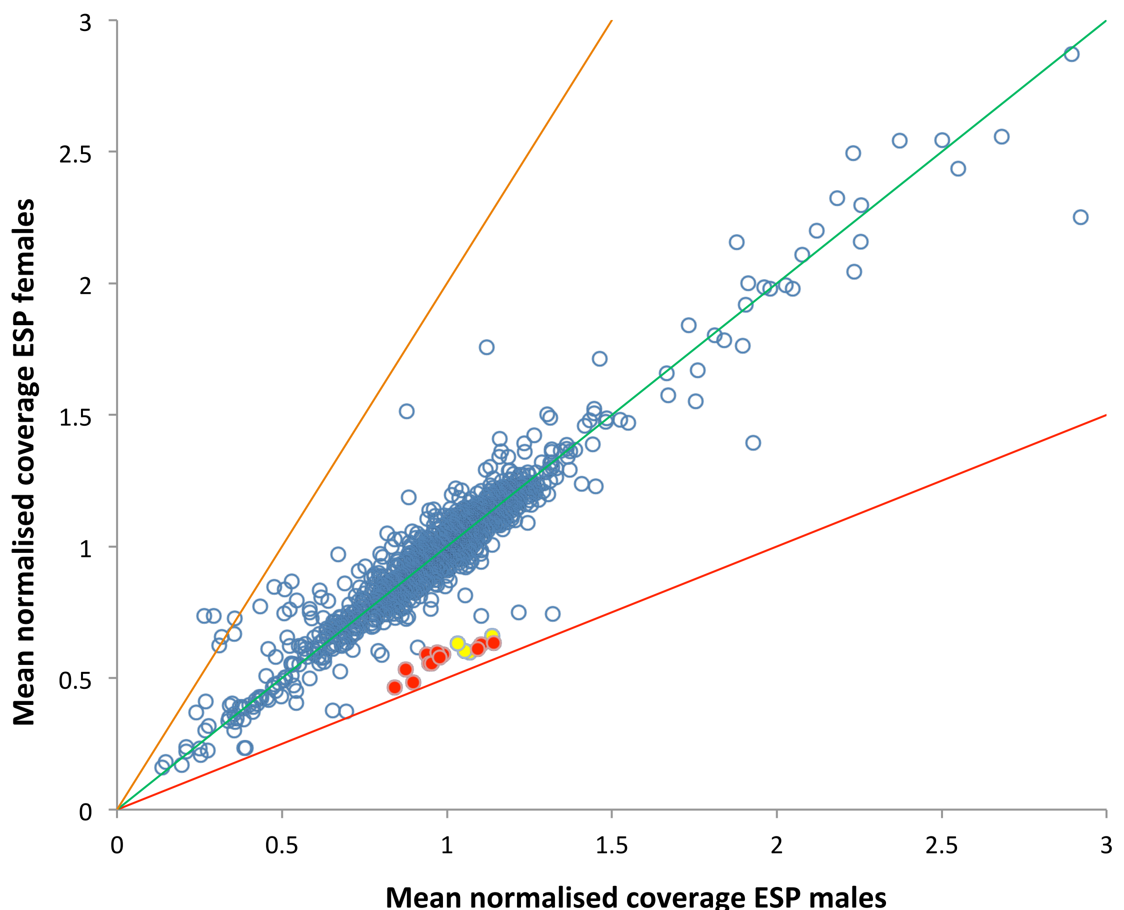


Figure 4.5: Ratio of normalised coverage in females compared to males for 1180 RAD loci found in all ESP individuals with a minimum of 20x coverage. Loci with greater than 3 times median autosomal coverage in either sex are excluded. The green line shows a 1:1 ratio of mean female to male normalised

(Figure 4.5 cont.) coverage expected for autosomal RAD loci, the orange line shows a 1.5:1 ratio, and the red line shows 0.5:1 ratio expected for Z specific RAD loci. Red points show 14 RAD loci identified as Z specific based on read depth patterns alone (see text), yellow points show 4 RAD loci identified as Z specific based on read depth patterns that also align to a W specific partner (see CodonCode analysis).

To identify putative Z specific RAD loci in the subset of 1180 loci found in all ESP individuals we first identified the loci that have significantly different mean normalised coverage in females compared to males by calculating the Student t statistic for each comparison in Microsoft Excel, recovering 50 loci (alpha = 0.05/1180 (Bonferoni corrected for 1180 tests), 24 d.f.) (Appendix 4, Figure 3). These loci were then filtered based on expectations of a ZW sex determination system using criteria inferred from W and Z specific RAD loci identified based on patterns of presence / absence. Mean normalised coverage of W specific loci in ESP (identified by patterns of presence / absence) deviates from the expected value of 0.5x autosomal coverage by a maximum of 0.23 (Figure 4.2). Therefore, as a conservative cut off, we only considered loci to be putatively Z specific in ESP if they had mean normalised coverage between 0.27 and 0.73 in females and between 0.77 and 1.23 in males. This subset of loci was then filtered again to remove RAD loci that have a significant difference in coverage, but not so great as would be expected to be caused by sex chromosome specificity. We removed loci with a difference in mean normalised coverage between males and females (males – females) less than 0.34, a value that corresponds to the minimum difference found between amphigenic hermaphrodites (WZ) and males (ZZ) for Z specific RAD loci in KOE (Figure 4.4). In total we recovered 16 candidate Z specific RAD loci in ESP (Figure 4.6), including #1913 that is also Z specific in KOE and 4 loci (including #1913) that have W specific gametologs in ESP females (Table 4.5).

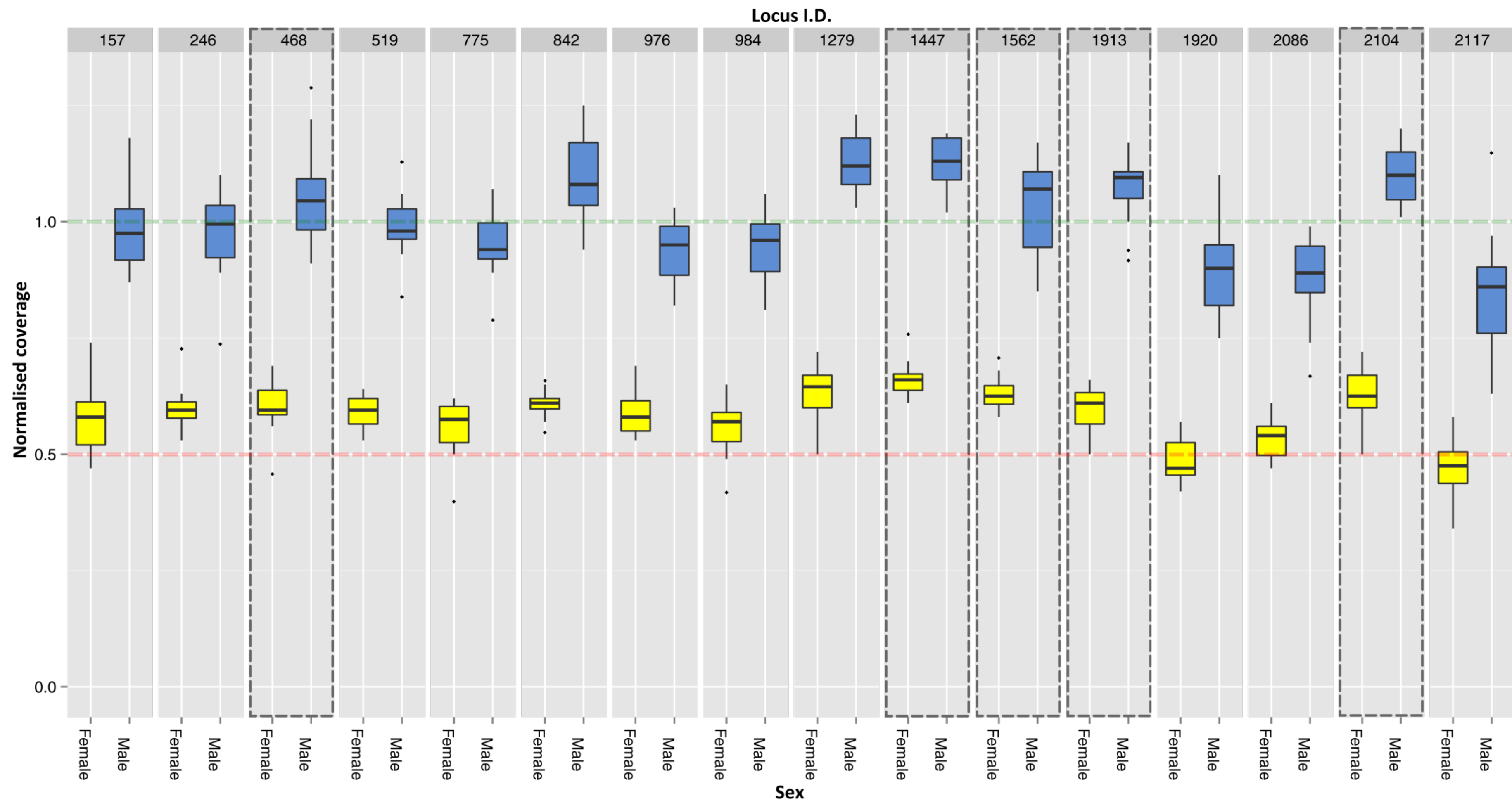


Figure 4.6: Female vs. male normalised coverage for 16 candidate Z specific RAD loci in ESP.

(Figure 4.6 cont.) Box plots show normalised coverage for 16 candidate Z specific RAD loci in ESP females (WZ) (yellow) and males (ZZ) (blue). Loci were identified using read depth patterns across 1180 loci found in all ESP individuals (see main text). The green dotted line shows expected coverage for autosomal RAD loci and the red dotted line shows expected coverage for hemizygous RAD loci. #1913 is also Z specific in KOE (Figure 4.4). Loci surrounded by grey dotted lines have W specific gametologs in the catalogue (Table 4.5).

3) Sex specific RAD alleles

We searched for polymorphic loci with alleles that segregate with the expectations for a ZW chromosomal sex determination system in each population within the catalogue. In ESP we expect fully sex-linked polymorphic RAD loci to be heterozygote in all females (ZW) and homozygote in all males (ZZ), furthermore we expect the W allele to be unique to females allowing us to phase W and Z haplotypes for these loci. Filtering for loci heterozygote in all females and homozygote in all males recovered 21 fully sex-linked loci in ESP, all of which have a female specific allele (Table 4.6). All 21 of these loci are found in KOE but only one, #317, is sex-linked. Significantly, in all cases the ESP W specific haplotype is found in KOE – 17 of the loci in KOE are fixed for the ESP W haplotype; 2 are polymorphic but do not show patterns of sex linkage and have the ESP W haplotype along with additional haplotypes not found in ESP; 1 is polymorphic and has both the ESP W and Z haplotypes but is not sex-linked in KOE; and finally one of the loci is sex-linked and has both the ESP W and Z haplotypes with the phase preserved. Filtering for RAD loci always heterozygote in males and always homozygote in females, as would be expected for a chromosomal sex determination system with male heterogamety (X/Y), recovered no candidate sex-linked RAD loci.

Filtering for sex-linked SNP loci in KOE - loci heterozygous in all amphigenic hermaphrodites (ZW) and homozygous in monogenic hermaphrodites (WW) and males (ZZ) - recovered two additional loci (Table 5). In both cases W and Z haplotypes segregated as expected with the W specific haplotype found only in hermaphrodites. Both loci were also found in ESP, one was monomorphic (#2099), the other W specific, with the KOE and ESP W haplotypes matching exactly (#1981). Overall 4 RAD loci with sex-linked alleles were found in KOE. Coverage across RAD loci identified as having fully sex-linked alleles was consistently high in all individuals with median coverage across loci ranging from 133x to 412x.

Identification of orthologs to sex-linked RAD loci

Assembly of loci in the catalogue produced 608 contigs that consisted of 1427 RAD loci identified by Stacks. Searching for contigs containing at least one sex-linked catalogue locus recovered 11 locus pairs (Table 4.5). Pairings between loci found in the same population, but specific to opposing sex chromosomes, can be considered gametologs – homologous loci that have diverged due to a lack of recombination between sex chromosomes (García-Moreno & Mindell 2000). In ESP we identified five W specific loci with matching putative Z specific gametologs identified based on read depth analysis. One of these 5 locus pairs, #1546 and #1913, also correspond to W and Z gametologs in KOE. A further three loci that are W specific in ESP have a match in the catalogue that does not appear to correspond to a Z specific gametolog. Two of these loci, #1524 and #2137, match loci found only in KOE and are not sex-linked, whilst locus 835 matches the ESP specific locus #834 which has coverage approximately twice that of autosomal loci in all individuals (Appendix 4, Figure 4), suggesting that #834 is a non sex-linked duplicated locus that has been further duplicated into the W chromosome identified as locus #835. None of the KOE W or Z specific loci have any matches in the catalogue other than those already mentioned that are also sex-linked in ESP (#1546 and #1780). Additionally, none of the sex-linked SNP loci have orthologous loci in the catalogue.

Table 4.5: Pairs of RAD loci in the catalogue that align to each other with at least one partner showing patterns of sex linkage. RAD locus pairs that can be considered gametologs in at least one of the two populations are shaded, light grey shading indicates pairs of loci that are gametologs in only one population, dark grey shading indicates pairs of loci that are gametologs in both populations. NSL = not sex-linked, - = absent, > 6 bp refers to loci split by Stacks due to having more than the specified 6 mismatches.

	Chromosome linkage			Chromosome linkage		
Locus I.D.	**ESP**	**KOE**	**Match to catalogue**	**ESP**	**KOE**	**Reason for splitting locus**
#227	W	NSL	#2104	Z	-	Indel
#835	W	NSL	#834	NSL	-	Indel
#991	W	NSL	#1562	Z	-	Indel
#1519	W	NSL	#1447	Z	-	Indel
#1524	W	NSL	#11281	-	NSL	Indel
#1546	W	W	#1913	Z	Z	Indel
#1722	W	NSL	#468	Z	-	Indel
#1780	W	W	#11536	-	Z	Indel
#2137	W	NSL	#11512	-	NSL	Indel
#842	Z	-	#11119	-	NSL	> 6 bp
#984	Z	-	#10990	-	NSL	Indel

Summary of sex-linked RAD loci and comparisons between ESP and KOE

In total we identified 69 putatively sex-linked RAD loci within the catalogue which correspond to 63 genomic markers when orthologous / gametologous loci are combined (Table 4.6; consensus sequences for each catalogue locus are given in Appendix 5). In the gonochoric population (ESP) we recovered 20 W specific RAD loci and 16 Z specific loci – five of which align to a gametolog on the W chromosome (Table 4.5). These loci also show coverage patterns consistent with being sex chromosome specific (Figures 4.2 and 4.6). In addition we recovered 21 RAD loci in ESP with completely sex-linked alleles. In the androdioecious population (KOE) we recovered two W specific RAD loci, both of which are also W specific in ESP, and 11 Z specific RAD loci one of which is also sex-linked in ESP (#1913). Two W/Z pairs are gametologs. Again, coverage patterns for these loci are consistent with them being sex chromosome specific (Figures 4.3 and 4.4). Fewer loci with sex-linked alleles were found in KOE (three) than in ESP (21). One of these loci is W specific in ESP (#1981), whilst another has sex specific alleles in ESP (#317). Overall we recovered 57 putative sex-linked RAD loci in ESP and 16 in KOE of which 4 loci are sex-linked in both

populations. Significantly, all putative W specific alleles from ESP (presence /absence and sex-linked alleles) are also found KOE with many not being completely sex-linked.

Table 4.6: Summary of fully sex-linked RAD loci found in gonochoric (ESP) and androdioecious (KOE) populations of *T. cancriformis*. Loci were identified through a combination of patterns of presence / absence, read depth analysis and segregation patterns of alleles. Where a sex-linked locus aligns to another locus in the catalogue, locus I.D. is combined and the sex linkage of each RAD locus given in the 'sex linkage' column (detailed information given in Table 4.5). Loci that are sex-linked in both populations are shaded grey. Note that some loci are found in both ESP and KOE but only show patterns of sex-linkage in one of the two populations. SL = sex-linked.

Locus I.D.	Population found	Sex linkage
#191	ESP and KOE	ESP W
#227-#2104	ESP and KOE	ESP W (#227), ESP Z (#2104)
#365	ESP and KOE	ESP W
#448	ESP and KOE	ESP W
#494	ESP and KOE	ESP W
#835-#834	ESP and KOE	ESP W
#991-#1562	ESP and KOE	ESP W (#991), ESP Z (#1562)
#1519-#1447	ESP and KOE	ESP W (#1519), ESP Z (#1447)
#1524-#11281	ESP and KOE	ESP W (#1542)
#1546-#1913	ESP and KOE	ESP W (#1546), KOE W (#1546), ESP Z (#1913), KOE Z (#1913)
#1722-#468	ESP and KOE	ESP W (1722), ESP Z (468)
#1780-#11536	ESP and KOE	ESP W (#1780), KOE W (1780), KOE Z (#11536)
#1791	ESP and KOE	ESP W, KOE SL Allele
#1981	ESP and KOE	ESP W, KOE SL Allele
#2110	ESP and KOE	ESP W
#2137-#11512	ESP and KOE	ESP W
#2139	ESP and KOE	ESP W
#2141	ESP and KOE	ESP W
#2215	ESP and KOE	ESP W
#2223	ESP and KOE	ESP W
#157	ESP	ESP Z
#246	ESP	ESP Z
#519	ESP	ESP Z
#689	ESP and KOE	ESP Z
#775	ESP	ESP Z
#842-#11119	ESP	ESP Z
#976	ESP and KOE	ESP Z
#984-#10990	ESP	ESP Z
#1279	ESP	ESP Z
#1920	ESP	ESP Z
#2086	ESP	ESP Z
#2117	ESP	ESP Z
#4909	ESP and KOE	KOE Z
#11478	KOE	KOE Z

#11483	KOE	KOE Z
#11485	KOE	KOE Z
#11488	KOE	KOE Z
#11495	KOE	KOE Z
#11499	KOE	KOE Z
#11561	KOE	KOE Z
#11562	KOE	KOE Z
#11	ESP and KOE	ESP SL Allele
#75	ESP and KOE	ESP SL Allele
#143	ESP and KOE	ESP SL Allele
#317	ESP and KOE	ESP SL Allele, KOE SL Allele
#321	ESP and KOE	ESP SL Allele
#368	ESP and KOE	ESP SL Allele
#513	ESP and KOE	ESP SL Allele
#550	ESP and KOE	ESP SL Allele
#592	ESP and KOE	ESP SL Allele
#623	ESP and KOE	ESP SL Allele
#732	ESP and KOE	ESP SL Allele
#765	ESP and KOE	ESP SL Allele
#779	ESP and KOE	ESP SL Allele
#781	ESP and KOE	ESP SL Allele
#943	ESP and KOE	ESP SL Allele
#1021	ESP and KOE	ESP SL Allele
#1073	ESP and KOE	ESP SL Allele
#1346	ESP and KOE	ESP SL Allele
#1369	ESP and KOE	ESP SL Allele
#1398	ESP and KOE	ESP SL Allele
#1953	ESP and KOE	ESP SL Allele
#2099	ESP and KOE	KOE SL Allele

4.5 Discussion

We identified over 1000 high coverage, novel genomic markers for *Triops cancriformis*. Through a combination of patterns of presence / absence, locus coverage depth analysis and SNP segregation patterns our data confirm chromosomal sex determination in *T. cancriformis* for the first time. In both gonochoric (ESP) and androdioecious (KOE) populations we identified sex-linked RAD loci with patterns of marker segregation matching the expectations of a ZW chromosomal sex determination system. Males from ESP and KOE are homogametic (ZZ), females from ESP heterogametic (ZW) and hermaphrodites from KOE are of two genetic types – monogenic WW hermaphrodites and amphigenic ZW hermaphrodites (Figure 4.3). This is akin to the sex determination system found for the androdioecious American

tadpole shrimp species *Triops newberryi* (Sassaman 1991), in which maleness is recessive to hermaphroditism and two types of hermaphrodites are present, suggesting that male recessivity might be conserved in the genus *Triops*. A ZW sex determination system is also found in several other branchiopod taxa (reviewed by Legrand et al. 1987) such as the clam shrimp *Eulimnadia texana* (Sassaman & Weeks 1993; Weeks et al. 2010) and the brine shrimp *Artemia franciscana* (De Vos et al. 2013).

Limitations of the dataset

Previous studies using RAD-seq to identify sex-linked markers and characterise sex chromosomes have relied on mapping RAD sequences to a reference genome (Anderson et al. 2012; Bewick et al. 2013). As no such reference genome is available for *T. cancriformis*, we carried out a de novo assembly of RAD loci and identification of fully sex-linked markers by association (see Etter et al. 2011a for a discussion of the relative merits of aligning to a reference vs. de novo assembly of RAD data). The identification of sex-linked markers relied on multiple lines of support, except for putative Z specific RAD loci in ESP, which were identified based only on patterns of coverage. In all cases, the markers identified here will require validation by PCR in additional individuals and assumptions of linkage between makers, and hence their synteny to putative sex chromosomes, confirmed through linkage mapping or genome sequencing.

The true number of *T. cancriformis* RAD loci in our dataset is uncertain. The distribution of sequence coverage across the catalogue is uneven, with many loci found in low coverage and in only one, or a few, individuals. In contrast, loci found in all individuals in the dataset, or in all individuals from each population, have consistently high coverage and account for the majority of sequenced reads, despite corresponding to a small fraction of loci in the catalogue (Figure 4.1 and Table 4.3). This pattern differs from typical coverage variation found in RAD-seq datasets (Davey et al. 2012) and suggests there is considerable low coverage contamination within our libraries. To mitigate this we applied stringent filtering criteria when identifying sex-linked loci that employed coverage cut-offs and presence in all individuals of either both populations, one of the populations, or within a sex, depending on the type of

makers we were looking for. Additionally, it should be noted that coverage across RAD loci that are highly likely to be from to *T. cancriformis* (i.e. those found in all individuals or all members of a population or sex) have an order of magnitude higher coverage than RAD loci sequenced in other studies, which typically achieve around 50x coverage (e.g. Baird et al. 2008; O'Quin et al. 2013; Richards et al. 2013; Sharma et al. 2012) enabling us to make inferences based on coverage patterns with greater confidence. With these strengths and limitations in mind we discuss our results in the following sections.

Structure of *Triops cancriformis* sex chromosomes

Our data suggest there are structural differences between both the W and Z sex chromosomes found in the gonochoric and androdioecious populations analysed. In ESP, 20 W specific RAD loci were identified, all of which are found in KOE but only two are W specific, with the remainder not being completely sex-linked in our dataset. The pattern is the same for RAD loci in ESP with completely sex-linked alleles – 21 completely sex-linked loci are found, all of which are present in KOE but only one remains sex-linked. Two scenarios could explain the difference in patterns of sex linkage between ESP and KOE at these loci: a large region of the KOE W sex chromosome could have been translocated to an autosomal location; alternatively, the same region of the W chromosome was translocated to the Z chromosome, creating or adding to a pseudoautosomal region (PAR) or a region where recombination occurs only occasionally. The lack of a genome sequence or genetic map means that it is not yet possible to disentangle these two hypotheses, although translocations between sex chromosomes are known in humans, where a large region of the X chromosome has been transposed onto the male specific Y, creating a highly similar region (Skaletsky et al. 2003). Rearrangements of sex chromosomes are also known in other species such as the frog *Rana rugosa*, where both XY and ZW chromosomal sex determination systems are found in separate populations (Miura et al. 1998; Nishioka et al. 1994; Ogata et al. 2007). Regardless of the mechanism, both scenarios suggest substantial rearrangement of the sex chromosomes found in the gonochoric and androdioecious populations of *T. cancriformis.* further highlighting the dynamic nature of sex chromosomes (Bachtrog 2006).

Levels of recombination between the sex chromosomes may also vary between the gonochoric and androdioecious populations sampled. Restricted recombination of sex chromosomes is expected to evolve in systems that have genetic sex determination and sexually dimorphic males and females due to sexually antagonistic selection (Charlesworth & Charlesworth 1980; Jordan & Charlesworth 2012; Mank & Ellegren 2009). Restricted recombination is often achieved through chromosomal rearrangements such as translocations, inversions or transpositions (Bachtrog 2013; Bergero & Charlesworth 2009; Charlesworth et al. 2005). In old sex chromosome systems this has occurred in successive events creating evolutionary strata such as those found in the sex chromosomes of humans (Lahn & Page 1999; Skaletsky et al. 2003), mice (Sandstedt & Tucker 2004), chickens (Handley et al. 2004; Nam & Ellegren 2008) and the plant *Silene latifolia* (Bergero et al. 2007; Bergero et al. 2013). In the case of *T. cancriformis*, sexual system has most likely undergone a transition from gonochorism to androdioecy (Zierold et al. 2007), as in *Eulimnadia* (Weeks et al. 2010; Weeks et al. 2006b), which will be expected to reduce sexually antagonistic selection as the Z chromosome is predicted to spend a greater proportion of time in hermaphrodites, where male function is also necessary, than males (Box 1). Although some sexually antoginistic selection is expected to occur in hermaphroditic plants and animals (Abbott 2011; Bedhomme et al. 2009), it is unclear what effect this transition will have on sex chromosome evolution. Our data show there are many more fully sex-linked markers in ESP than KOE indicating that the non-recombining region is much larger in the gonochoric than the androdioecious population. This is consistent with a reduction in the suppression of recombination in the androdioecious population which could be due to reduced sexually antagonistic selection, suggesting that the transition in sexual system has had major influence on the course of sex chromosome evolution in *T. cancriformis*.

Box 1: Sexually antagonistic selection in gonochoric and androdioecious populations of *Triops cancriformis*

In a gonochoric population with a 50:50 sex ratio, as found in ESP (Zierold et al. 2009), the W chromosome is always found in females (ZW), so there is potential for selection for accumulation of mutations that specifically benefit females and that are deleterious in males. The situation is reversed for the Z chromosome but less pronounced, as it is found in males (ZZ) 2/3s of the time and in females (WZ) 1/3 of the time giving a slight advantage for the accumulation of male biased variants.

In androdioecious populations where monogenic (WW) and amphigenic (ZW) hermaphrodites make up the bulk of the population, the dynamics of sexually antagonistic selection change considerably. The W chromosome is always found in hermaphrodites so is not expected to accumulate mutations that harm either its male or female function, although selection could favour the accumulation of mutations beneficial to production of either male or female gametes, but that are harmful to male specific factors such as mate searching, mating behaviour or mating efficiency. More significantly, however, are possible changes in sexually antagonistic selection for the Z chromosome. Male, amphigenic hermaphrodite and monogenic hermaphrodite proportions in KOE (based on the male proportion sampled by Zierold et al. (2009) and the proportions of monogenic and amphigenic hermaphrodites found in this study) are given below:

10% males (ZZ)
23% monogenic hermaphrodites (WW)
67% amphigenic hermaphrodites (ZW)

The Z chromosome would therefore be present in hermaphrodites 78% of the time and in males only 22% of the time, meaning little or no sexually antagonistic selection is expected.

Evolution of androdioecy in *Triops cancriformis*

If we assume that a translocation took place between the W and the Z in KOE, homogenizing them, our results indicate that the ESP and KOE W sex chromosomes are highly similar; in all cases ESP W specific haplotypes are found in KOE, some of which are sex-linked. These data strongly suggest that hermaphrodites have evolved from females rather than males, based on the conservation of the W chromosome between females and hermaphrodites. It also suggests that the transition between gonochorism and androdioecy has been relatively recent, as proposed by Zierold et al. (2007). In contrast, the Z chromosome shows significant differences between the two

populations with many unique Z specific RAD loci found in both KOE and ESP. This raises the possibility that whilst hermaphrodites may have originated from females, genetically similar to those found in ESP, males in the androdioecious population could be of a different origin.

The evolution of hermaphrodites from females, rather than males, is predicted by theory. Weeks et al. have argued that in sexually dimorphic species it is simpler to evolve a self-fertilising hermaphrodite from a female rather than a male due to developmental constraints (Weeks 2009; Weeks 2012; Weeks et al. 2006b). In females, only a mutation causing sperm production in the ovaries, to produce an ovotestis, is required for the evolution self-fertile hermaphrodites. Simple ovotestis are found throughout androdioecious and hermaphroditic species of *Triops* (Akita 1971; Garcia-Velazco et al. 2009; Longhurst 1955b; Murugan et al. 2009; Scanabissi et al. 2005), including *T. cancriformis*, suggesting this pathway to hermaphroditism could be conserved across the genus. The low number of shared sex-linked RAD loci between the two populations suggests that these loci may be tightly linked to the master sex determining locus, assuming this is also conserved. The conserved sex-linked markers, given that they are amongst the few that are sex-linked in KOE, must also be linked to the mutation determining hermaphroditism in androdioecious populations. Further characterisation of these regions of the sex chromosomes, and investigation of these markers in related androdioecious taxa, could shed light on the genetic changes involved in the evolution of hermaphroditism in *Triops* and indeed, whether this mechanism is conserved across the genus.

Conclusions

Our data confirm, for the first time, the presence of a ZW chromosomal sex determination system in the European tadpole shrimp *Triops cancriformis* that is conserved in both gonochoric and androdioecious populations. Although similar patterns of segregation of completely sex-linked RAD loci and alleles were found in ESP and KOE, there were considerable differences between the two populations in the number of sex-linked loci identified, with many more sex-linked markers found in ESP than KOE. These data suggest that chromosomal rearrangements have occurred, possibly due to changes in sexually antagonistic selection between gonochoric and

androdioecious sexual systems. Additionally, we identified high similarity between the W chromosomes, but not the Z chromosomes, found in ESP and KOE. It appears that hermaphrodites in *T.cancriformis* have evolved from a female progenitor as predicted by theory (Weeks 2009; Weeks 2012; Weeks et al. 2006b). The presence of differentiated sex chromosomes in *T. cancriformis* alongside diverse sexual systems (Zierold et al. 2007; Zierold et al. 2009) makes the species an excellent model for the study of sex chromosome evolution. Future work should extend the study of sex chromosomes across Notostraca, where multiple transitions in sexual system have occurred (Chapter 3), to bring further insights into the genomic effects of labile sexual systems.

Chapter 5

General discussion and concluding remarks

In the first chapter of this thesis I set out four key obstacles hindering the advancement of the study of sexual system evolution in tadpole shrimps. These were: (1) the lack of a resolved phylogeny of the order, (2) a lack of knowledge on the extent of sexual system variation in tadpole shrimps and its evolutionary dynamics, (3) limited genomic resources and (4) a limited understanding of the genetic basis of sex determination in the order. The research presented here does much to alleviate these issues and establishes tadpole shrimps as a model to study the evolution of sexual systems and sex chromosomes in animals. In this final chapter I will summarise the main findings of the thesis and suggest some directions for future work.

5.1 Tadpole shrimp phylogeny

A key output of this thesis is the estimation of the first well resolved phylogeny of Notostraca that includes all known species. The last formal review of the taxonomy of the order was conducted by Longhurst over 50 years ago (Longhurst 1955c) and it is now clear that an up to date revision of the order is required to formally name the numerous cryptic species that have been discovered in this study and elsewhere (King & Hanner 1998; Korn et al. 2010; Korn & Hundsdoerfer 2006; Macdonald et al. 2011; Mantovani et al. 2009; Rogers 2001; Vanschoenwinkel et al. 2012). Resolving notostracan phylogeny has made it possible to infer the timing and tempo of tadpole shrimp diversification (Chapter 2) and to investigate the evolutionary dynamics of the diverse sexual systems found in the order (Chapter 3). In the future this phylogeny will provide a framework to explore tadpole shrimp genomes.

5.2 The timing and tempo of tadpole shrimp diversification

In Chapter 2 I investigate the timing and tempo of speciation in tadpole shrimps and find, contrary to expectations from the fossil record, that Notostraca have undergone at least two bouts of diversification resulting in around 40 extant species. Tadpole shrimps have often been cited as examples of 'living fossils' – relicts of once diverse groups – (e.g. Fryer 1988; Mantovani et al. 2004; SunoUchi et al. 1997) due to the high level of morphological conservatism found in the group and the similarity of extant species to fossils found as far back as the Early Devonian (Fayers & Trewin 2002). The pattern and timing of diversification of tadpole shrimps identified in Chapter 2 is, however, incompatible with the designation of 'living fossil' status, except on the grounds of morphological similarity. Instead, extant tadpole shrimp species belong to a group that has undergone considerable diversification over the last 60 million years. This pattern of cryptic diversification has also been highlighted in other living fossil groups such as cycads (Nagalingum et al. 2011), nautiloids (Wray et al. 1995) horseshoe crabs (Obst et al. 2012) and monoplacophorans (Kano et al. 2012) suggesting that many so called 'living fossil' fossils may not hold up to the scrutiny of modern molecular phylogenetic methods.

5.3 Labile sexual systems

In Chapter 3 I investigate the dynamics of sexual system evolution in tadpole shrimps. I demonstrate that sexual system is extremely labile across the order with multiple transitions having occurred between gonochorism and androdioecy. Support was also found for the reproductive assurance hypothesis, also known as Baker's Law (Baker 1955; Baker 1967), as an explanation for the evolution of self-fertile hermaphroditism. Androdioecious taxa were found at significantly higher latitudes than gonochoric ones, where post glacial re-colonisation may have caused selection for reproductive assurance. It appears that reproductive assurance plays a key role in evolutionary transitions from separate to combined sexes in both plants and animals (Foxe et al. 2009; Hesse & Pannell 2011; Holsinger 2000; Jarne & Auld 2006; Kalisz et al. 2004; Obbard et al. 2006b; Pannell 2000; Pannell 2002a; Pannell & Barrett 1998; Weeks et al. 2009). The labile nature of sexual system found in Notostraca makes tadpole shrimps

an excellent system to study the genomic consequences of transitions in sexual system.

5.4 Sex determination and sex chromosome evolution

In Chapter 4 I used restriction site associated DNA sequencing (RAD-seq) (Baird et al. 2008) to generate thousands of genome wide markers to investigate the genetics of sex determination in gonochoric and androdioecious populations of the European tadpole shrimp *Triops cancriformis*, a species with mixed sexual systems. Using patterns of RAD marker segregation and locus coverage depth across individually bar-coded females, males and hermaphrodites I confirmed the presence of a ZW chromosomal sex determination system that is conserved in both the gonochoric and androdioecious populations sampled. Although some completely sex-linked markers were shared, there was evidence of substantial rearrangements in the structure of both the W and Z chromosomes between the two populations, likely due to a change in sexually antagonistic selection caused by the transition in sexual system between the two populations. The analysis also shed light on the origins of hermaphrodites in the androdioecious population with the distribution of RAD haplotypes suggesting that hermaphrodites are likely to be derived from females, based on the high similarity of the W chromosomes found in the two populations.

Although the RAD sequencing approach used here has been successful in identifying many sex-linked markers, it has not yet been used to its full potential and further analysis of the dataset is likely to provide a more nuanced view of the structure and evolution of sex chromosomes in tadpole shrimp. In the present study, we identified completely sex-linked RAD loci. However, population linkage patterns could be used to identify loci in partial linkage with sex-linked regions. Additionally, paired-end sequence data produced for this study could be used to generate contigs up to 600bp in length, associated with the first end RAD markers (Etter et al. 2011b; Willing et al. 2011). The contigs associated with sex-linked RAD loci could then be used in BLAST searches to identify candidate sex-linked genes. Alternatively, the high sequencing coverage of the RAD libraries generated for this study means the paired-end contigs could be used to identify additional polymorphic RAD markers, as was done to investigate genetic diversity in the Eurasian beaver (Senn et al. 2013). Comparing the

sequence evolution of gametologous RAD loci and their paired-end contigs from the W and Z sex chromosomes of each population could also resolve evolutionary relationships between the W and Z chromosomes, as has been done for tree frogs (Stöck et al. 2011).

5.5 Concluding remarks

The research presented within this thesis provides an excellent basis for the future use of tadpole shrimps as a model to study sexual system evolution and its various, far reaching effects. In particular, the system appears to be ideal to investigate the influence of transitions between sexual systems on the progress of sex chromosome evolution, combining two major themes of research in evolutionary biology – the forces driving sex chromosome evolution (Bachtrog 2006; Bachtrog et al. 2011; Bergero & Charlesworth 2009; Charlesworth 1991; Charlesworth et al. 2005; Ellegren 2011; Rice 1987b), and the evolution of sexual systems (Barrett 2002; Barrett 2010; Charlesworth 2006; Charlesworth & Charlesworth 1979; Charlesworth & Wright 2001; Charnov et al. 1976; Eppley & Jesson 2008; Pannell 2009). Tadpole shrimps have undergone multiple transitions between separate and combined sexes (Chapter 3) over an established timeframe (Chapter 2), providing numerous comparisons for targeted comparative genomic studies. Investment in the further development of genomic resources across the notostracan phylogeny could therefore elevate tadpole shrimps from interesting quirks of natural history to very useful tools in the study of sex chromosome evolution.

References

Abbott JK. 2011. Intra-locus sexual conflict and sexually antagonistic genetic variation in hermaphroditic animals. *Proceedings of the Royal Society B: Biological Sciences* 278:161-169.

Ahl J. 1991. Factors affecting contributions of the tadpole shrimp, *Lepidurus packardi*, to its oversummering egg reserves. *Hydrobiologia* 212:137-143.

Akita M. 1971. On the reproduction of *Triops longicaudatus* (LECONTE). *Doubutsugaku zasshi* 80:242-250.

Anderson JL, Marí AR, Braasch I, Amores A, Hohenlohe P, Batzel P, and Postlethwait JH. 2012. Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. *Plos One* 7:e40701.

Atkinson QD, Meade A, Venditti C, Greenhill SJ, and Pagel M. 2008. Languages Evolve in Punctuational Bursts. *Science* 319:588.

Avise JC, Nelson WS, and Sugita H. 1994. A speciational history of "living fossils": Molecular evolutionary patterns in horseshoe crabs. *Evolution* 48:1986-2001.

Bachtrog D. 2006. A dynamic view of sex chromosome evolution. *Current Opinion in Genetics & Development* 16:578-585.

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nature Reviews Genetics* 14:113-124.

Bachtrog D, and Charlesworth B. 2002. Reduced adaptation of a non-recombining neo-Y chromosome. *Nature* 416:323-326.

Bachtrog D, Kirkpatrick M, Mank JE, McDaniel SF, Pires JC, Rice W, and Valenzuela N. 2011. Are all sex chromosomes created equal? *Trends in Genetics* 27:350-357.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, and Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *Plos One* 3:e3376.

Baker HG. 1955. Self-compatibility and establishment after 'long-distance' dispersal. *Evolution* 9:347-349.

Baker HG. 1967. Support for Baker's Law-as a rule. *Evolution* 21:853-856.

Barnard KH. 1929. A revision of the South African Branchiopoda Phyllopoda. *Annals of the South African Museum* 29:181 - 272.

Barrett LG, Thrall PH, Burdon JJ, and Linde CC. 2008. Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends in Ecology & Evolution* 23:678-685.

Barrett SC. 2002. The evolution of plant sexual diversity. *Nature Reviews Genetics* 3:274-284.

Barrett SCH. 1998. The evolution of mating strategies in flowering plants. *Trends in Plant Science* 3:335-341.

Barrett SCH. 2010. Understanding plant reproductive diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:99-109.

Bawa KS. 1980. Evolution of dioecy in flowering plants. *Annual Review of Ecology and Systematics* 11:15-39.

Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG, Jiggins CD, and Blaxter ML. 2011. Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *Plos One* 6:e19315.

Bedhomme S, Bernasconi G, Koene JM, Lankinen Å, Arathi HS, Michiels NK, and Anthes N. 2009. How does breeding system variation modulate sexual antagonism? *Biology Letters* 5:717-720.

Bergero R, and Charlesworth D. 2009. The evolution of restricted recombination in sex chromosomes. *Trends in Ecology & Evolution* 24:94-102.

Bergero R, Forrest A, Kamau E, and Charlesworth D. 2007. Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics* 175:1945-1954.

Bergero R, Qiu S, Forrest A, Borthwick H, and Charlesworth D. 2013. Expansion of the pseudoautosomal region and ongoing recombination suppression in the *Silene latifolia* sex chromosomes. *Genetics*.

Bernard H. 1891. Hermaphroditism of the Apodidae. *Nature* 43:343-344.

Bernard HM. 1896. XL.—Hermaphroditism among the Apodidæ. *Journal of Natural History Series 6* 17:296-309.

Bewick AJ, Chain FJ, Zimmerman LB, Sesay A, Gilchrist MJ, Owens ND, Seifertova E, Krylov V, Macha J, and Tlapakova T. 2013. A large pseudoautosomal region on

the sex chromosomes of the frog *Silurana tropicalis*. *Genome Biology and Evolution* 5:1087-1098.

Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, and Purvis A. 2007. The delayed rise of present-day mammals. *Nature* 446:507-512.

Blaustein L, and Schwartz SS. 2001. Why study ecology in temporary pools? *Israel Journal of Zoology* 47:303-312.

Brandley MC, Wang Y, Guo X, Montes de Oca AN, Fería-Ortíz M, Hikida T, and Ota H. 2011. Accommodating heterogenous rates of evolution in molecular divergence dating methods: an example using intercontinental dispersal of Plestiodon (Eumeces) lizards. *Systematic Biology* 60:3-15.

Brendonck L. 1996. Diapause, quiescence, hatching requirements: What we can learn from large freshwater branchiopods (Crustacea:Branchiopoda:Anostraca, Notostraca, Conchostraca). *Hydrobiologia* 320:85-97.

Brendonck L, Rogers DC, Olesen J, Weeks S, and Hoeh WR. 2008. Global diversity of large branchiopods (Crustacea : Branchiopoda) in freshwater. *Hydrobiologia* 595:167-176.

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

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

Bromham L, Rambaut A, Fortey R, Cooper A, and Penny D. 1998. Testing the Cambrian explosion hypothesis by using a molecular dating technique. *Proceedings of the National Academy of Sciences* 95:12386-12389.

Burbrink FT, and Pyron RA. 2010. How does ecological opportunity influence rates of speciation, extinction, and morphological fiversification in New World ratsnakes (tribe Lampropeltini)? *Evolution* 64:934-943.

Bushnell JH, and Byron ER. 1979. Morphological variability and distribution of aquatic invertebrates (principally Crustacea) from the Cumberland Peninsula and Frobisher Bay regions, Baffin Island, NWT, Canada. *Arctic and Alpine Research* 11:159-177.

Cardinal S, and Danforth BN. 2011. The antiquity and evolutionary history of social behavior in bees. *Plos One* 6:e21086.

Casane D, and Laurenti P. 2013. Why coelacanths are not 'living fossils'. *BioEssays* 35:332-338.

Catchen J, Bassham S, Wilson T, Currey M, O'Brien C, Yeates Q, and Cresko WA. 2013a. The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Molecular Ecology* 22:2864-2883.

Catchen J, Hohenlohe PA, Bassham S, Amores A, and Cresko WA. 2013b. Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22:3124-3140.

Catchen JM, Amores A, Hohenlohe P, Cresko W, and Postlethwait JH. 2011. Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics* 1:171-182.

Cesari M, Mularoni L, Scanabissi F, and Mantovani B. 2004. Characterization of dinucleotide microsatellite loci in the living fossil tadpole shrimp *Triops cancriformis* (Crustacea Branchiopoda Notostrace). *Molecular Ecology Notes* 4:733-735.

Charlesworth B. 1991. The evolution of sex chromosomes. *Science* 251:1030-1033.

Charlesworth B. 1996. The evolution of chromosomal sex determination and dosage compensation. *Current Biology* 6:149-162.

Charlesworth B, and Charlesworth D. 1978. A model for the evolution of dioecy and gynodioecy. *American Naturalist*:975-997.

Charlesworth B, and Charlesworth D. 1997. Rapid fixation of deleterious alleles can be caused by Muller's ratchet. *Genetics Research* 70:63-73.

Charlesworth B, and Dempsey ND. 2001. A model of the evolution of the unusual sex chromosome system of *Microtus oregoni*. *Heredity* 86:387-394.

Charlesworth D. 1984. Androdioecy and the evolution of dioecy. *Biological Journal of the Linnean Society* 22:333-348.

Charlesworth D. 2006. Evolution of plant breeding systems. *Current Biology* 16:R726-R735.

Charlesworth D, and Charlesworth B. 1979. The evolutionary genetics of sexual systems in flowering plants. *Proceedings of the Royal Society of London Series B, Biological Sciences* 205:513-530.

Charlesworth D, and Charlesworth B. 1980. Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genet Res* 35:205-214.

Charlesworth D, Charlesworth B, and Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95:118-128.

Charlesworth D, and Wright SI. 2001. Breeding systems and genome evolution. *Current Opinion in Genetics & Development* 11:685-690.

Charnov EL. 1982. *The theory of sex allocation*. Princeton, NJ: Princeton University Press.

Charnov EL, Bull JJ, and Smith JM. 1976. Why be an hermaphrodite? *Nature* 263:125-126.

Chong Z, Ruan J, and Wu C-I. 2012. Rainbow: an integrated tool for efficient clustering and assembling RAD-seq reads. *Bioinformatics* 28:2732-2737.

Clutton-Brock T. 2007. Sexual selection in males and females. *Science* 318:1882-1885.

Colbourne JK, and Hebert PDN. 1996. The systematics of North American *Daphnia* (Crustacea: Anomopoda): A molecular phylogenetic approach. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 351:349-360.

Dacks J, and Roger AJ. 1999. The first sexual lineage and the relevance of facultative sex. *Journal of Molecular Evolution* 48:779-783.

Darwin C. 1859. *On the origin of species by means of natural selection*. London: Murray.

Darwin C. 1877. *The different forms of flowers on plants of the same species*. London: J. Murray.

Darwin CR. 1876. *The effects of cross and self fertilisation in the vegetable kingdom*. London: John Murray.

Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin AV, Hughes DS, Ferguson LC, and Martin SH. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487:94.

Davey JW, and Blaxter ML. 2010. RADSeq: next-generation population genetics. *Briefings in Functional Genomics* 9:416-423.

Davey JW, Cezard T, Fuentes-Utrilla P, Eland C, Gharbi K, and Blaxter ML. 2012. Special features of RAD sequencing data: implications for genotyping. *Molecular Ecology* 22:3151-3164.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, and Blaxter ML. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* 12:499-510.

De Vos S, Bossier P, Van Stappen G, Vercauteren I, Sorgeloos P, and Vuylsteke M. 2013. A first AFLP-based genetic linkage map for brine shrimp *Artemia franciscana* and Its application in mapping the sex locus. *Plos One* 8:e57585.

dos Reis M, Inoue J, Hasegawa M, Asher RJ, Donoghue PCJ, and Yang Z. 2012. Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proceedings of the Royal Society B: Biological Sciences* 279:3491-3500.

Drummond A, and Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7.

Drummond AJ, Ho SYW, Phillips MJ, and Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4:e88.

Drummond AJ, Suchard MA, Xie D, and Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29:1969-1973.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792-1797.

Ekblom R, and Galindo J. 2010. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107:1-15.

Eldredge N, and Gould SJ. 1972. Punctuated equilibria: an alternative to phyletic gradualism. *Models in paleobiology* 82:115.

Ellegren H. 2011. Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nature Reviews Genetics* 12:157-166.

Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, and Holzapfel CM. 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences* 107:16196-16200.

Engelmann M, and Hahn T. 2005. Vorkommen von *Lepidurus apus*, *Triops cancriformis*, *Eubranchipus (Siphonophanes) grubei*, *Tanymastix stagnalis* und *Branchipus*

schaefferi in Deutschland und Österreich (Crustacea: Notostraca und Anostraca). *Faunistische Abhandlungen* 25:3-67.

Eppley SM, and Jesson LK. 2008. Moving to mate: the evolution of separate and combined sexes in multicellular organisms. *Journal of Evolutionary Biology* 21:727-736.

Etter P, Bassham S, Hohenlohe P, Johnson E, and Cresko W. 2011a. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In: Orgogozo V, and Rockman MV, eds. *Molecular Methods for Evolutionary Genetics*: Humana Press, 157-178.

Etter PD, Preston JL, Bassham S, Cresko WA, and Johnson EA. 2011b. Local *de novo* assembly of RAD paired-end contigs using short sequencing reads. *Plos One* 6:e18561.

Fayers SR, and Trewin NH. 2002. A new crustacean from the Early Devonian Rhynie chert, Aberdeenshire, Scotland. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 93:355-382.

Fedonkin MA, Simonetta A, and Ivantsov AY. 2007. New data on *Kimberella*, the Vendian mollusc-like organism (White Sea region, Russia): palaeoecological and evolutionary implications. *Geological Society, London, Special Publications* 286:157-179.

Felsenstein J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125:1-15.

Fisher RA. 1931. The evolution of dominance. *Biological Reviews* 6:345-368.

Folmer O, Black M, Hoeh W, Lutz R, and Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3:294-299.

Foxe JP, Slotte T, Stahl EA, Neuffer B, Hurka H, and Wright SI. 2009. Recent speciation associated with the evolution of selfing in Capsella. *Proceedings of the National Academy of Sciences of the United States of America* 106:5241-5245.

Fraser JA, Diezmann S, Subaran RL, Allen A, Lengeler KB, Dietrich FS, and Heitman J. 2004. Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *PLoS Biology* 2:e384.

Fraser JA, and Heitman J. 2005. Chromosomal sex-determining regions in animals, plants and fungi. *Current Opinion in Genetics & Development* 15:645-651.

Freckleton R, Harvey P, and Pagel M. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist* 160:712-726.

Fredslund J, Schauser L, Madsen LH, Sandal N, and Stougaard J. 2005. PriFi: using a multiple alignment of related sequences to find primers for amplification of homologs. *Nucleic Acids Research* 33:W516-W520.

Fridolfsson A-K, Cheng H, Copeland NG, Jenkins NA, Liu H-C, Raudsepp T, Woodage T, Chowdhary B, Halverson J, and Ellegren H. 1998. Evolution of the avian sex chromosomes from an ancestral pair of autosomes. *Proceedings of the National Academy of Sciences* 95:8147-8152.

Fryer G. 1988. Studies on the functional-morphology and biology of the Notostraca (Crustacea, Branchiopoda). *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 321:27-142.

Fryer G. 1996. Diapause, a potent force in the evolution of freshwater crustaceans. *Hydrobiologia* 320:1-14.

Gall J-C, and Grauvogel-Stamm L. 2005. The early Middle Triassic 'Grès àVoltzia' Formation of eastern France: a model of environmental refugium. *Paléontologie Générale* 4:569 560 584.

Gand G, Garric J, and Lapeyrie J. 1997. Biocénoses à triopsidés (Crustacea, Branchiopoda) du Permien du bassin de Lodève (France). *Geobios* 30:673-700.

García-Moreno J, and Mindell DP. 2000. Rooting a phylogeny with homologous genes on opposite sex chromosomes (gametologs): a case study using avian CHD. *Molecular Biology and Evolution* 17:1826-1832.

Garcia-Velazco H, Obregon-Barboza H, Rodriguez-Jaramillo C, and Maeda-Martinez AM. 2009. Reproduction of the tadpole shrimp *Triops* (Notostraca) in Mexican waters. *Current Science* 96:91-97.

Garrouste R, Nel A, and Gand G. 2009. New fossil arthropods (Notostraca and Insecta: Syntonopterida) in the Continental Middle Permian of Provence (Bas-Argens Basin, France). *Comptes Rendus Palevol* 8:49-57.

Ghiselin MT. 1969. The evolution of hermaphroditism among animals. *The Quarterly Review of Biology* 44:189-208.

Glémin S. 2007. Mating systems and the efficacy of selection at the molecular level. *Genetics* 177:905-916.

Glemin S, Bazin E, and Charlesworth D. 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proceedings of the Royal Society B-Biological Sciences* 273:3011-3019.

Goldberg EE, and Igić B. 2008. On phylogenetic tests of irreversible evolution. *Evolution* 62:2727-2741.

Gore PJW. 1986. Triassic notostracans in the Newark supergroup, Culpeper Basin, Northern Virginia. *Journal of Paleontology* 60:1086-1096.

Gould SJ. 1970. Dollo on Dollo's law: irreversibility and the status of evolutionary laws. *Journal of the History of Biology* 3:189-212.

Graves JAM. 2006. Sex chromosome specialization and degeneration in mammals. *Cell* 124:901-914.

Graves JAM. 2008. Weird animal genomes and the evolution of vertebrate sex and sex chromosomes. *Annual Review of Genetics* 42:565-586.

Green AJ, Sánchez MI, Amat F, Figuerola J, Hontoria F, Ruiz O, and Hortas F. 2005. Dispersal of invasive and native brine shrimps *Artemia* (Anostraca) via waterbirds. *Limnology and Oceanography* 50:737-742.

Guerrero RF, Kirkpatrick M, and Perrin N. 2012. Cryptic recombination in the ever-young sex chromosomes of Hylid frogs. *Journal of Evolutionary Biology* 25:1947-1954.

Guthörl P. 1934. Die arthropoden aus dem Carbon und Perm des Saar-Nahe-Pfalz-gebietes. *Preussischen Geologischen Landesanstalt* 164:1 - 219.

Hamer M, and Appleton C. 1991. Life history adaptations of phyllopods in response to predators, vegetation, and habitat duration in north-eastern Natal. *Hydrobiologia* 212:105-116.

Handley L-JL, Ceplitis H, and Ellegren H. 2004. Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics* 167:367-376.

Harvey THP, Vélez MI, and Butterfield NJ. 2012. Exceptionally preserved crustaceans from western Canada reveal a cryptic Cambrian radiation. *Proceedings of the National Academy of Sciences* 109:1589-1594.

Hasegawa M, Kishino H, and Yano T-a. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22:160-174.

Haughton S. 1924. The fauna and stratigraphy of the Stormberg Series. *Annals of the South African Museum* 12:323-497.

Hecht BC, Campbell NR, Holecek DE, and Narum SR. 2013. Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. *Molecular Ecology* 22:3061-3076.

Hegna TA, and Dong R. 2010. Two new "Notostracans", *Chenops* gen. nov and *Jeholops* gen. nov (Crustacea: Branchiopoda: ?Notostraca) from the Yixian Formation, Northeastern China. *Acta Geologica Sinica-English Edition* 84:886-894.

Heidecke D, and Neumann V. 1987. Zur Verbreitung und Ökologie von *Triops cancriformis* BOSC und *Lepidurus apus* L. in der DDR. *Hercynia NF* 24:166-173.

Hesse E, and Pannell JR. 2011. Density dependent pollen limitation and reproductive assurance in a wind pollinated herb with contrasting sexual systems. *Journal of Ecology* 99:1531-1539.

Hessen DO, Rueness EK, and Stabell M. 2004. Circumpolar analysis of morphological and genetic diversity in the Notostracan *Lepidurus arcticus*. *Hydrobiologia* 519:73-84.

Ho SYW, and Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology* 58:367-380.

Høeg J. 1995. Sex and the single cirripede: a phylogenetic perspective. *Crustacean issues* 10.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, and Cresko WA. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet* 6:e1000862.

Hohenlohe PA, Day MD, Amish SJ, Miller MR, Kamps-Hughes N, Boyer MC, Muhlfeld CC, Allendorf FW, Johnson EA, and Luikart G. 2013. Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology* 22:3002-3013.

Holsinger KE. 2000. Reproductive systems and evolution in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America* 97:7037-7042.

Huelsenbeck JP, Larget B, and Swofford D. 2000. A compound Poisson process for relaxing the molecular clock. *Genetics* 154:1879-1892.

Hug LA, and Roger AJ. 2007. The impact of fossils and taxon sampling on ancient molecular dating analyses. *Molecular Biology and Evolution* 24:1889-1897.

Igic B, Lande R, and Kohn JR. 2008. Loss of self-incompatibility and its evolutionary consequences. *International Journal of Plant Sciences* 169:93-104.

Inoue J, Donoghue PCJ, and Yang Z. 2010. The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. *Systematic Biology* 59:74-89.

Jarne P, and Auld JR. 2006. Animals mix it up too: the distribution of self-fertilistaion among hermaphrodtitic animals. *Evolution* 60:1816-1824.

Jarne P, and Charlesworth D. 1993. The evolution of the selfing rate in functionally hermaphrodite plants and animals. *Annual Review of Ecology and Systematics*:441-466.

Jordan CY, and Charlesworth D. 2012. The potential for sexually antagonistic polymorphism in differnt genome regions. *Evolution* 66:505-516.

Joyce WG, Parham JF, Lyson TR, Warnock RC, and Donoghue PC. 2013. A divergence dating analysis of turtles using fossil calibrations: an example of best practices. *Journal of Paleontology* 87:612-634.

Kalisz S, Vogler DW, and Hanley KM. 2004. Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. *Nature* 430:884-887.

Kano Y, Kimura S, Kimura T, and Warén A. 2012. Living Monoplacophora: morphological conservatism or recent diversification? *Zoologica Scripta* 41:471-488.

Kato Y, Kobayashi K, Watanabe H, and Iguchi T. 2011. Environmental sex determination in the branchiopod crustacean *Daphnia magna*: deep conservation of a *Doublesex* gene in the sex-determining pathway. *PLoS Genet* 7:e1001345.

Keller I, Wagner C, Greuter L, Mwaiko S, Selz O, Sivasundar A, Wittwer S, and Seehausen O. 2012. Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Molecular Ecology* 22:2848-2863.

Kimura M. 1968. Evolutionary rate at the molecular level. *Nature* 217:624-626.

Kimura M. 1969. The rate of molecular evolution considered from the standpoint of population genetics. *Proceedings of the National Academy of Sciences* 63:1181-1188.

Kimura M. 1984. *The neutral theory of molecular evolution*: Cambridge University Press.

Kimura M, and Ohta T. 1971. On the rate of molecular evolution. *Journal of Molecular Evolution* 1:1-17.

King JL, and Hanner R. 1998. Cryptic species in a "Living Fossil" lineage: taxonomic and phylogenetic relationships within the genus *Lepidurus* (Crustacea : Notostraca) in North America. *Molecular Phylogenetics and Evolution* 10:23-36.

Kishino H, Thorne JL, and Bruno WJ. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Molecular Biology and Evolution* 18:352-361.

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

Korn M, Green AJ, Machado M, Garcia-de-Lomas J, Cristo M, da Fonseca LC, Frisch D, Perez-Bote JL, and Hundsdoerfer AK. 2010. Phylogeny, molecular ecology and taxonomy of southern Iberian lineages of *Triops mauritanicus* (Crustacea: Notostraca). *Organisms Diversity and Evolution* 10:409-440.

Korn M, and Hundsdoerfer AK. 2006. Evidence for cryptic species in the tadpole shrimp *Triops granarius* (Lucas, 1864) (Crustacea : Notostraca). *Zootaxa* 1257:57-68.

Korn M, Marrone F, Perez-Bote JL, Machado M, Cristo M, da Fonseca LC, and Hundsdoerfer AK. 2006. Sister species within the *Triops cancriformis* lineage (Crustacea, Notostraca). *Zoologica Scripta* 35:301-322.

Kotov AA, and Taylor DJ. 2011. Mesozoic fossils (> 145 Mya) suggest the antiquity of the subgenera of *Daphnia* and their coevolution with chaoborid predators. *BMC Evolutionary Biology* 11:129.

Kuck P, and Meusemann K. 2010. FASconCAT: convenient handling of data matrices. *Molecular Phylogenetics and Evolution* 56:1115-1118.

Kuller Z, and Gasith A. 1996. Comparison of the hatching process of the tadpole shrimps *Triops cancriformis* and *Lepidurus apus lubbocki* (Notostraca) and its relation to their distribution in rain-pools in Israel. *Hydrobiologia* 335:147-157.

Kumar S. 2005. Molecular clocks: four decades of evolution. *Nature Reviews Genetics* 6:654-662.

Kumar S, Filipski A, Swarna V, Walker A, and Hedges SB. 2005. Placing confidence limits on the molecular age of the human–chimpanzee divergence. *Proceedings of the National Academy of Sciences of the United States of America* 102:18842-18847.

Lahn BT, and Page DC. 1999. Four evolutionary strata on the human X chromosome. *Science* 286:964-967.

Lanfear R, Calcott B, Ho SYW, and Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* 29:1695-1701.

Lanfear R, Welch JJ, and Bromham L. 2010. Watching the clock: studying variation in rates of molecular evolution between species. *Trends in Ecology & Evolution* 25:495-503.

Legrand J, LEGRAND ‐ HAMELIN E, and Juchault P. 1987. Sex determination in Crustacea. *Biological Reviews* 62:439-470.

Lehtonen J, Jennions MD, and Kokko H. 2012. The many costs of sex. *Trends in Ecology & Evolution* 27:172-178.

Leigh JW, Susko E, Baumgartner M, and Roger AJ. 2008. Testing congruence in phylogenomic analysis. *Systematic Biology* 57:104-115.

Letsch HO, and Kjer KM. 2011. Potential pitfalls of modelling ribosomal RNA data in phylogenetic tree reconstruction: evidence from case studies in the Metazoa. *BMC Evolutionary Biology* 11.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Subgroup GPDP. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078-2079.

Linder F. 1952. Contributions to the morphology and the taxonomy of the Branchiopoda Notostraca, with special reference to the North American species. *Proceedings of the United States National Museum* 102:1-69.

Listen A, Rieseberg LH, and Elias TS. 1990. Functional androdioecy in the flowering plant *Datisca glomerata*. *Nature* 343:641-642.

Livernois A, Graves J, and Waters P. 2011. The origin and evolution of vertebrate sex chromosomes and dosage compensation. *Heredity* 108:50-58.

Lloyd DG. 1975. The maintenance of gynodioecy and androdioecy in angiosperms. *Genetica* 45:325-339.

Longhurst AR. 1954. Reproduction in Notostraca (Crustacea). *Nature* 173:781-782.

Longhurst AR. 1955a. Evolution in the Notostraca. *Evolution* 9:84-86.

Longhurst AR. 1955b. The reproduction and cytology of the Notostraca (Crustacea, Phyllopoda). *Proceedings of the Zoological Society of London* 125:671-680.

Longhurst AR. 1955c. A review of the Notostraca. *Bulletin of the British Museum (Natural History)* 3:1-57.

Lunter G, and Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research* 21:936-939.

Lynch JE. 1966. *Lepidurus lemmoni* Holmes: a redescription with notes on variation and distribution. *Transactions of the American Microscopical Society* 85:181-192.

Macdonald KS, Sallenave R, and Cowley DE. 2011. Morphological and genetic variation in *Triops* (Branchiopoda: Notostraca) from ephemeral waters of the Northern Chihuahuan desert of North America. *Journal of Crustacean Biology* 31:468-484.

Machado M, Cristo M, Reis J, and Cancela L. 1999. Biological data on *Triops cancriformis mauritanicus* (Ghigi, 1921) and *Cyzicus grubei* (Simon, 1886)—crustacea, branchiopoda—in SW Portugal temporary ponds. *Limnetica* 16:1-7.

MacKay WP, Loring SJ, Frost TM, and Whitford WG. 1990. Population dynamics of a playa community in the Chihuahuan Desert. *The Southwestern Naturalist* 35:393-402.

Maddison WP, Donoghue MJ, and Maddison DR. 1984. Outgroup analysis and parsimony. *Systematic Biology* 33:83-103.

Maddison WP, and Maddison DR. 2010. Mesquite: a modular system for evolutionary analysis. 2.74 ed.

Mank JE, and Ellegren H. 2009. Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. *Evolution* 63:1464-1472.

Mantovani B, Cesari M, Luchetti A, and Scanabissi F. 2008. Mitochondrial and nuclear DNA variability in the living fossil *Triops cancriformis* (Bosc, 1801) (Crustacea, Branchiopoda, Notostraca). *Heredity* 100:496-505.

Mantovani B, Cesari M, and Scanabissi F. 2004. Molecular taxonomy and phylogeny of the 'living fossil' lineages *Triops* and *Lepidurus* (Branchiopoda : Notostraca). *Zoologica Scripta* 33:367-374.

Mantovani B, Cesari M, and Scanabissi F. 2009. Molecular taxonomy and phylogeny of Italian *Lepidurus* taxa (Branchiopoda: Notostraca). *Italian Journal of Zoology* 76:358-365.

Marcus V, and Weeks S. 1997. The effects of pond duration on the life history traits of an ephemeral pond crustacean, *Eulimnadia texana*. *Hydrobiologia* 359:213-221.

Mardis ER. 2008. The impact of next-generation sequencing technology on genetics. *Trends in Genetics* 24:133-141.

Margraf J, and Maass B. 1982. Zur Ökologie der temporären Süßwasserflachseen des Tafelberges „Giara di Gesturi" auf Sardinien. *Spixiana* 5:69-99.

Mathers TC, Hammond RL, Jenner RA, Zierold T, Hänfling B, and Gómez A. 2013. High lability of sexual system over 250 million years of evolution in morphologically conservative tadpole shrimps. *BMC Evolutionary Biology* 13:30.

Maynard Smith J. 1978. *The evolution of sex* Cambridge [Eng.] Cambridge University Press.

Meusemann K, von Reumont BM, Simon S, Roeding F, Strauss S, Kuck P, Ebersberger I, Walzl M, Pass G, Breuers S, Achter V, von Haeseler A, Burmester T, Hadrys H, Wagele JW, and Misof B. 2010. A phylogenomic approach to resolve the arthropod tree of life. *Molecular Biology and Evolution* 27:2451-2464.

Meyer BJ. 2000. Sex in the worm: counting and compensating X-chromosome dose. *Trends in Genetics* 16:247-253.

Mitsumoto H, and Yahata K. 2006. Evidence of cross-fertilization in a gonochoric population of the tadpole shrimp *Triops numidicus* (Crustacea : Branchiopoda : Notostraca). *Zoological Science* 23:1109-1113.

Miura I, Ohtani H, Nakamura M, Ichikawa Y, and Saitoh K. 1998. The origin and differentiation of the heteromorphic sex chromosomes Z, W, X, and Y in the frog *Rana rugosa*, inferred from the sequences of a sex-linked gene, ADP/ATP translocase. *Molecular Biology and Evolution* 15:1612-1619.

Montero-Pau J, Gomez A, and Munoz J. 2008. Application of an inexpensive and high-throughput genomic DNA extraction method for the molecular ecology of

zooplanktonic diapausing eggs. *Limnology and Oceanography-Methods* 6:218-222.

Morgan G. 1998. Emile Zuckerkandl, Linus Pauling, and the molecular evolutionary clock, 1959–1965. *Journal of the History of Biology* 31:155-178.

Morris SC. 1998. Early metazoan evolution: reconciling paleontology and molecular biology. *American Zoologist* 38:867-877.

Murugan G, Maeda-Martinez AM, Obregon-Barboza H, and Hernandez-Saavedra NY. 2002. Molecular characterization of the tadpole shrimp *Triops* (Branchiopoda : Notostraca) from the Baja California Peninsula, Mexico: New insights on species diversity and phylogeny of the genus. *Hydrobiologia* 486:101-113.

Murugan G, Obregon-Barboza H, Maeda-Martinez AM, and Timms BV. 2009. Co-occurrence of two tadpole shrimp, *Triops cf. australiensis* (Branchiopoda : Notostraca), lineages in middle Paroo, north-western New South Wales, with the first record of *Triops* hermaphrodites for the Australian continent. *Australian Journal of Zoology* 57:77-84.

Nagalingum NS, Marshall CR, Quental TB, Rai HS, Little DP, and Mathews S. 2011. Recent synchronous radiation of a living fossil. *Science* 334:796-799.

Nam K, and Ellegren H. 2008. The chicken (*Gallus gallus*) Z chromosome contains at least three nonlinear evolutionary strata. *Genetics* 180:1131-1136.

Narum SR, Buerkle CA, Davey JW, Miller MR, and Hohenlohe PA. 2013. Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology* 22:2841-2847.

Nishioka M, Hanada H, Miura I, and Ryuzaki M. 1994. Four kinds of sex chromosomes in Rana rugosa. *Scientific report of the Laboratory for Amphibian Biology* 13:1-34.

O'Quin KE, Yoshizawa M, Doshi P, and Jeffery WR. 2013. Quantitative genetic analysis of retinal degeneration in the blind cavefish *Astyanax mexicanus*. *Plos One* 8:e57281.

Obbard DJ, Harris SA, Buggs RJA, and Pannell JR. 2006a. Hybridization, polyploidy, and the evolution of sexual systems in *Mercurialis* (Euphorbiaceae). *Evolution* 60:1801-1815.

Obbard DJ, Harris SA, and Pannell JR. 2006b. Sexual systems and population genetic structure in an annual plant: testing the metapopulation model. *American Naturalist* 167:354-366.

Obregon-Barboza H, Maeda-Martinez AM, Murugan G, Timms BV, Grygier MJ, Rogers DC, Rodriguez-Almaraz G, and Dumont HJ. 2007. Morphology and systematic significance of the mystax, a hitherto undescribed cephalic structure of males in certain Notostraca (Branchiopoda). *Journal of Crustacean Biology* 27:18-23.

Obst M, Faurby S, Bussarawit S, and Funch P. 2012. Molecular phylogeny of extant horseshoe crabs (Xiphosura, Limulidae) indicates Paleogene diversification of Asian species. *Molecular Phylogenetics and Evolution* 62:21-26.

Ogata M, Hasegawa Y, Ohtani H, Mineyama M, and Miura I. 2007. The ZZ/ZW sex-determining mechanism originated twice and independently during evolution of the frog, *Rana rugosa*. *Heredity* 100:92-99.

Ogden R, Gharbi K, Mugue N, Martinsohn J, Senn H, Davey JW, Pourkazemi M, McEwing R, Eland C, Vidotto M, Sergeev A, and Congiu L. 2013. Sturgeon conservation genomics: SNP discovery and validation using RAD sequencing. *Molecular Ecology* 28:2732-2737.

Ombretta M, Cesari M, Eder E, Scanabissi F, and Mantovani B. 2005. Chromosomes in sexual populations of Notostracan and Con-chostracan taxa (Crustacea, Branchiopoda). *Caryologia* 58:164-170.

Otto SP, Sassaman C, and Feldman MW. 1993. Evolution of sex determination in the conchostracan shrimp *Eulimnadia texana*. *The American Naturalist* 141:329-337.

Pacheco MA, Battistuzzi FU, Lentino M, Aguilar R, Kumar S, and Escalante AA. 2011. Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. *Molecular Biology and Evolution* 28:1927-1942.

Packard A. 1875. New phyllopod crustaceans. *American Naturalist* 9:311-312.

Pagel M. 1994. Detecting correlated evolution on phylogenies: a general mehod for the coparative analysis of discrete characters. *Proceedings of the Royal Society of London Series B: Biological Sciences* 255:37-45.

Pagel M. 1999a. Inferring the historical patterns of biological evolution. *Nature* 401:877-884.

Pagel M. 1999b. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology* 48:612-622.

Pagel M, and Meade A. 2004. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Systematic Biology* 53:571-581.

Pagel M, and Meade A. 2005. Mixture models in phylogenetic inference. In: Gascuel O, ed. *Mathematics of evolution and phylogeny*: Oxford University Press, New York, 121-139.

Pagel M, Meade A, and Barker D. 2004. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology* 53:673-684.

Palumbi S, Martin A, Romano S, McMillan W, Stice L, and Grabowski G. 1991. The simple fool's guide to PCR, Version 2. . *Special Publication of Department of Zoology, University of Hawaii, Honolulu, Hawaii, USA*:28.

Pannell J. 1997. The maintenance of gynodioecy and androdioecy in a metapopulation. *Evolution* 51:10-20.

Pannell JR. 2000. A hypothesis for the evolution of androdioecy: the joint influence of reproductive assurance and local mate competition in a metapopulation. *Evolutionary Ecology* 14:195-211.

Pannell JR. 2002a. The evolution and maintenance of androdioecy. *Annual Review of Ecology and Systematics* 33:397-425.

Pannell JR. 2002b. What is functional androdioecy? *Functional Ecology* 16:862-865.

Pannell JR. 2008. Consequences of inbreeding depression due to sex-linked loci for the maintenance of males and outcrossing in branchiopod crustaceans. *Genetical Research* 90:73-84.

Pannell JR. 2009. On the problems of a closed marriage: celebrating Darwin 200. *Biology Letters* 5:332-335.

Pannell JR, and Barrett SCH. 1998. Baker's law revisited: reproductive assurance in a metapopulation. *Evolution* 52:657-668.

Pannell JR, and Dorken ME. 2006. Colonisation as a common denominator in plant metapopulations and range expansions: effects on genetic diversity and sexual systems. *Landscape Ecology* 21:837-848.

Pannell JR, Obbard DJ, and Buggs RJA. 2004. Polyploidy and the sexual system: what can we learn from *Mercurialis annua*? *Biological Journal of the Linnean Society* 82:547-560.

Parham JF, Donoghue PCJ, Bell CJ, Calway TD, Head JJ, Holroyd PA, Inoue JG, Irmis RB, Joyce WG, Ksepka DT, Patané JSL, Smith ND, Tarver JE, van Tuinen M, Yang Z, Angielczyk KD, Greenwood JM, Hipsley CA, Jacobs L, Makovicky PJ, Müller J, Smith KT, Theodor JM, Warnock RCM, and Benton MJ. 2012. Best practices for justifying fossil calibrations. *Systematic Biology* 61:346-359.

Pease JB, and Hahn MW. 2012. Sex chromosomes evolved from independent ancestral linkage groups in winged insects. *Molecular Biology and Evolution* 29:1645-1653.

Peterson KJ, and Butterfield NJ. 2005. Origin of the Eumetazoa: testing ecological predictions of molecular clocks against the Proterozoic fossil record. *Proceedings of the National Academy of Sciences of the United States of America* 102:9547-9552.

Petrov B, and Cvetkovic DM. 1997. Community structure of branchiopods (Anostraca, Notostraca and Conchostraca) in the Banat province in Yugoslavia. *Hydrobiologia* 359:23-28.

Phillimore AB, and Price TD. 2008. Density-dependent cladogenesis in birds. *PLoS Biology* 6:e71.

Pires-daSilva A. 2007. Evolution of the control of sexual identity in nematodes. *Seminars in Cell & Developmental Biology* 18:362-370.

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

Pulquério MJF, and Nichols RA. 2007. Dates from the molecular clock: how wrong can we be? *Trends in Ecology & Evolution* 22:180-184.

Rabosky DL. 2006. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evolutionary Bioinformatics Online* 2:273-276.

Rambaut A, and Drummond A. 2007. Tracer v1. 4 [upated to v1. 5], Available from http://beast.bio.ed.ac.uk/Tracer.

Ramesh MA, Malik S-B, and Logsdon Jr JM. 2005. A phylogenomic inventory of meiotic genes: evidence for sex in Giardia and an early eukaryotic origin of meiosis. *Current Biology* 15:185-191.

Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, and Cunningham CW. 2010. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* 463:1079-U1098.

Rehm P, Borner J, Meusemann K, von Reumont BM, Simon S, Hadrys H, Misof B, and Burmester T. 2011. Dating the arthropod tree based on large-scale transcriptome data. *Molecular Phylogenetics and Evolution* 61:880-887.

Reitzel AM, Herrera S, Layden MJ, Martindale MQ, and Shank TM. 2013. Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology* 22:2953-2970.

Reznick DN, and Ricklefs RE. 2009. Darwin's bridge between microevolution and macroevolution. *Nature* 457:837-842.

Rhodes FHT. 1983. Gradualism, punctuated equilibrium and the Origin of Species. *Nature* 305:269-272.

Rice W. 1994. Degeneration of a nonrecombining chromosome. *Science* 263:230-232.

Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution*:735-742.

Rice WR. 1987a. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution* 41:911-914.

Rice WR. 1987b. Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* 116:161-167.

Richards PM, Liu MM, Lowe N, Davey JW, Blaxter ML, and Davison A. 2013. RAD-Seq derived markers flank the shell colour and banding loci of the Cepaea nemoralis supergene. *Molecular Ecology* 22:3077-3089.

Rogers DC. 2001. Revision of the nearctic *Lepidurus* (Notostraca). *Journal of Crustacean Biology* 21:991-1006.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, and Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian

phylogenetic inference and model choice across a large model space. *Systematic Biology* 61:539-542.

Ross M, and Weir B. 1976. Maintenance of males and females in hermaphrodite populations and the evolution of dioecy. *Evolution*:425-441.

Sakai AK, Weller SG, Chen M-L, Chou S-Y, and Tasanont C. 1997. Evolution of gynodioecy and maintenance of females: the role of inbreeding depression, outcrossing rates, and resource allocation in Schiedea adamantis (Caryophyllaceae). *Evolution*:724-736.

Sandstedt SA, and Tucker PK. 2004. Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. *Genome Research* 14:267-272.

Sassaman C. 1989. Sex-ratio and sex determination in Western populations of *Triops* (Notostraca). *American Zoologist* 29:A66-A66.

Sassaman C. 1991. Sex-ratio variation in female-biased populations of notostracans. *Hydrobiologia* 212:169-179.

Sassaman C, Simovich MA, and Fugate M. 1997. Reproductive isolation and genetic differentiation in North American species of *Triops* (Crustacea : Branchiopoda : Notostraca). *Hydrobiologia* 359:125-147.

Sassaman C, and Weeks SC. 1993. The genetic mechanism of sex determination in the Chonchostracan shrimp *Eulimnadia texana*. *American Naturalist* 141:314-328.

Saunders JF. 1980. A rediscription of *Lepidurus bilobatus* (Crustacea, Notostraca). *Transactions of the American Microscopical Society* 99:179-186.

Scanabissi F, Eder E, and Cesari M. 2005. Male occurrence in Austrian populations of *Triops cancriformis* (Branchiopoda, Notostraca) and ultrastructural observations of the male gonad. *Invertebrate Biology* 124:57-65.

Scanabissi F, and Mondini C. 2002. A survey of the reproductive biology in Italian branchiopods. *Hydrobiologia* 486:263-272.

Scotese CR. 2001. *Atlas of earth history*: University of Texas at Arlington. Department of Geology. PALEOMAP Project.

Senn H, Ogden R, Cezard T, Gharbi K, Iqbal Z, Johnson E, Kamps‐Hughes N, Rosell F, and McEwing R. 2013. Reference‐free SNP discovery for the Eurasian beaver from restriction site–associated DNA paired‐end data. *Molecular Ecology*.

Sharma R, Goossens B, Kun-Rodrigues C, Teixeira T, Othman N, Boone JQ, Jue NK, Obergfell C, O'Neill RJ, and Chikhi L. 2012. Two different high throughput sequencing approaches identify thousands of *de novo* genomic markers for the genetically depleted bornean elephant. *Plos One* 7:e49533.

Simon E. 1886. Étude sur les Crustacés du sous-ordre des Phyllopodes. *Annales de la Société entomologique de France* 6:393-460.

Simovich MA. 1998. Crustacean biodiversity and endemism in California's ephemeral wetlands. In: Witham CW, Bauder ET, Belk D, Ferren WR, and Ornduff R, eds. *Ecology, Conservation, and Management of Vernal Pool Ecosystems - Proceedings from a 1996 Conference*. Sacramento: California Native Plant Society, 107-118.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, and Page DC. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423:825-U822.

Soon WW, Hariharan M, and Snyder MP. 2013. High-throughput sequencing for biology and medicine. *Molecular Systems Biology* 9.

Spigler R, Lewers K, Main D, and Ashman T. 2008. Genetic mapping of sex determination in a wild strawberry, *Fragaria virginiana*, reveals earliest form of sex chromosome. *Heredity* 101:507-517.

Spigler RB, Lewers KS, Johnson AL, and Ashman T-L. 2010. Comparative mapping reveals autosomal origin of sex chromosome in octoploid *Fragaria virginiana*. *Journal of Heredity* 101:S107-S117.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.

Stephan S. 2008. Mannchen von *Lepidurus apus* (Notostraca, Crustacea) in der Ruhstadter Elbtalaue. *Abhandlungen und Berichte fur Naturkunde* 31:47-57.

Stöck M, Horn A, Grossen C, Lindtke D, Sermier R, Betto-Colliard C, Dufresnes C, Bonjour E, Dumas Z, Luquet E, Maddalena T, Sousa HC, Martinez-Solano I, and

Perrin N. 2011. Ever-young sex chromosomes in European tree frogs. *PLoS Biology* 9:e1001062.

Stocsits RR, Letsch H, Hertel J, Misof B, and Stadler PF. 2009. Accurate and efficient reconstruction of deep phylogenies from structured RNAs. *Nucleic Acids Research* 37:6184-6193.

SunoUchi N, Sasaki F, Chiba S, and Kawata M. 1997. Morphological stasis and phylogenetic relationships in tadpole shrimps, *Triops* (Crustacea: Notostraca). *Biological Journal of the Linnean Society* 61:439-457.

Szathmáry E, and Maynard Smith J. 1995. The major evolutionary transitions. *Nature* 374:227-232.

Takahashi F. 1977. Pioneer life of the tadpole shrimps, *Triops ssp.* (Notostraca: Triopsidae). *Applied Entomology and Zoology* 12:104-117.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28:2731-2739.

Thorne JL, and Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51:689-702.

Thorne JL, Kishino H, and Painter IS. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15:1647-1657.

Trentini M. 1976. Karyologic observations on *Triops Cancriformis* (Phyllopoda, Notostraca). *Bolletino di Zoologia* 43:321-322.

Trusheim F. 1938. Triopsiden (Crust. Phyll.) aus dem Keuper Frankens. *Paläontologische Zeitschrift* 19:198-216.

Urano S, Yamaguchi S, Yamato S, Takahashi S, and Yusa Y. 2009. Evolution of dwarf males and a variety of sexual modes in barnacles: an ESS approach. *Evolutionary Ecology Research* 11:713-729.

van Doorn GS, and Kirkpatrick M. 2010. Transitions between male and female heterogamety caused by sex-antagonistic selection. *Genetics* 186:629-645.

van Leeuwen CHA, van der Velde G, van Groenendael JM, and Klaassen M. 2012. Gut travellers: internal dispersal of aquatic organisms by waterfowl. *Journal of Biogeography* 39:2031-2040.

Vanschoenwinkel B, Pinceel T, Vanhove MPM, Denis C, Jocque M, Timms BV, and Brendonck L. 2012. Toward a global phylogeny of the “living fossil" crustacean order of the Notostraca. *Plos One* 7:e34998.

Venditti C, Meade A, and Pagel M. 2011. Multiple routes to mammalian diversity. *Nature* 479:393-396.

Vicoso B, Kaiser VB, and Bachtrog D. 2013. Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. *Proceedings of the National Academy of Sciences* 110:6453-6458.

von Reumont BM, Jenner RA, Wills MA, Dell’Ampio E, Pass G, Ebersberger I, Meyer B, Koenemann S, Iliffe TM, Stamatakis A, Niehuis O, Meusemann K, and Misof B. 2012. Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. *Molecular Biology and Evolution* 29:1031-1045.

Vyskot B, and Hobza R. 2004. Gender in plants: sex chromosomes are emerging from the fog. *Trends in Genetics* 20:432-438.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, and Seehausen O. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology* 22:787-798.

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

Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang Y-C, Shoemaker D, and Keller L. 2013. A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* 493:664-668.

Warnock RCM, Yang Z, and Donoghue PCJ. 2012. Exploring uncertainty in the calibration of the molecular clock. *Biology Letters* 8:156-159.

Webb CJ. 1999. Empirical studies: evolution and maintenance of dimorphic breeding systems. *Gender and sexual dimorphism in flowering plants*: Springer, 61-95.

Weeks SC. 1990. Life-history variation under varying degrees of intraspecific competition in the tadpole shrimp *Triops longicaudatus* (LeConte). *Journal of Crustacean Biology* 10:498-503.

Weeks SC. 2009. Can males successfully invade hermaphroditic populations of clam shrimp (*Eulimnadia texana*)? *Current Science* 96:98-102.

Weeks SC. 2012. The role of androdioecy and gynodioecy in mediating evolutionary transitions between dioecy and hermaphroditism in the Animalia. *Evolution* 66:3670-3686.

Weeks SC, Benvenuto C, and Reed SK. 2006a. When males and hermaphrodites coexist: a review of androdioecy in animals. *Integrative and Comparative Biology* 46:449-464.

Weeks SC, Benvenuto C, Sanderson TF, and Duff RJ. 2010. Sex chromosome evolution in the clam shrimp, *Eulimnadia texana*. *Journal of Evolutionary Biology* 23:1100-1106.

Weeks SC, Chapman EG, Rogers DC, Senyo DM, and Hoeh WR. 2009. Evolutionary transitions among dioecy, androdioecy and hermaphroditism in limnadiid clam shrimp (Branchiopoda: Spinicaudata). *Journal of Evolutionary Biology* 22:1781-1799.

Weeks SC, Posgai RT, Cesari M, and Scanabissi F. 2005. Androdioecy inferred in the clam shrimp *Eulimnadia agassizii* (Spinicaudata : Limnadiidae). *Journal of Crustacean Biology* 25:323-328.

Weeks SC, Sanderson TF, Reed SK, Zofkova M, Knott B, Balaraman U, Pereira G, Senyo DM, and Hoeh WR. 2006b. Ancient androdioecy in the freshwater crustacean *Eulimnadia*. *Proceedings of the Royal Society B-Biological Sciences* 273:725-734.

Weeks SC, Sanderson TF, Zofkova M, and Knott B. 2008. Breeding systems in the clam shrimp family Limnadiidae (Branchiopoda, Spinicaudata). *Invertebrate Biology* 127:336-349.

Weiblen GD, Oyama RK, and Donoghue MJ. 2000. Phylogenetic Analysis of Dioecy in Monocotyledons. *The American Naturalist* 155:46-58.

Wiens JJ. 2011. Re-evolution of lost mandibular teeth in frogs after more than 200 million years, and re-evaluating Dollo's Law. *Evolution* 65:1283-1296.

Wilkinson RD, Steiper ME, Soligo C, Martin RD, Yang Z, and Tavaré S. 2011. Dating primate divergences through an integrated analysis of palaeontological and molecular data. *Systematic Biology* 60:16-31.

Williams GC. 1975. *Sex and evolution*: Princeton University Press.

Willing E-M, Hoffmann M, Klein JD, Weigel D, and Dreyer C. 2011. Paired-end RAD-seq for de novo assembly and marker design without available reference. *Bioinformatics* 27:2187-2193.

Wojtasik B, and Brylka-Wolk M. 2010. Reproduction and genetic structure of a freshwater crustacean *Lepidurus arcticus* from Spitsbergen. *Polish Polar Research* 31:33-44.

Wolf DE, and Takebayashi N. 2004. Pollen limitation and the evolution of androdioecy from dioecy. *American Naturalist* 163:122-137.

Womack T, Slater BJ, Stevens LG, Anderson LI, and Hilton J. 2012. First cladoceran fossils from the Carboniferous: palaeoenvironmental and evolutionary implications. *Palaeogeography, Palaeoclimatology, Palaeoecology* 344–345:39-48.

Wray CG, Landman NH, Saunders WB, and Bonacum J. 1995. Genetic divergence and geographic diversification in Nautilus. *Paleobiology*:220-228.

Wray GA, Levinton JS, and Shapiro LH. 1996. Molecular evidence for deep Precambrian divergences among metazoan phyla. *Science* 274:568-573.

Yang Z, and Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution* 23:212-226.

Yang Z, and Rannala B. 2012. Molecular phylogenetics: principles and practice. *Nature Reviews Genetics* 13:303-314.

Yusa Y, Yoshikawa M, Kitaura J, Kawane M, Ozaki Y, Yamato S, and Høeg JT. 2012. Adaptive evolution of sexual systems in pedunculate barnacles. *Proceedings of the Royal Society B: Biological Sciences* 279:959-966.

Zerbino DR, and Birney E. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* 18:821-829.

Zierold T, Hanfling B, and Gomez A. 2007. Recent evolution of alternative reproductive modes in the 'living fossil' *Triops cancriformis*. *BMC Evolutionary Biology* 7.

Zierold T, Montero-Pau J, Hanfling B, and Gomez A. 2009. Sex ratio, reproductive mode and genetic diversity in *Triops cancriformis*. *Freshwater Biology* 54:1392-1405.

Zucker N, Cunningham M, and Adams HP. 1997. Anatomical evidence for androdioecy in the clam shrimp *Eulimnadia texana*. *Hydrobiologia* 359:171-175.

Zuckerkandl E, and Pauling L. 1965. Evolutionary divergence and convergence in proteins. In: Bryson V, and Vogel HJ, editors. Evolving genes and proteins. New York: Academic Press. p 97-166.

Glossary

awk: programming language designed for text processing. Implemented in Unix-based operating systems.

Crown group: in phylogenetic terminology a crown group contains all the living representatives of a clade back to their most recent common ancestor and extinct descendants of that common ancestor.

Dollo's Law: an evolutionary principle that states that complex characters cannot re-evolve once lost (Gould 1970).

Generalised mixed Yule coalescent (GMYC) species delimitation: maximum likelihood model based approach to identify the point in a phylogenetic tree where the branching pattern changes from that of species-level evolutionary processes (speciation and extinction) to population-level evolutionary processes (coalescent) in order to identify putative species (Pons et al. 2006).

Stem group: in phylogenetic terminology a stem group refers to extinct taxa that are paraphyletic to a given crown group (see above).

Yule model of speciation: a pure birth model of speciation where each branch in a phylogenetic tree has a birth rate associated to it that determines the rate at which the branch bifurcates.

Appendix 1: Additional information for Chapter 2

Table 1: GMYC species delimitation results (Chapter 1). Accession numbers within each ML cluster defined by the GMYC model. Accessions in bold were used in the multilocus phylogenetic analysis.

Cluster ID	Accession
L. apus	DQ834543.1
	DQ834544.1
	DQ148285.1
	EF189669.1
	JX110638
L. arcticus	HM425362.1 BOLD
	AF209067.1
	DQ834545.1
	HM425363.1 BOLD
	HM425361.1 BOLD
	DQ148286.1
	HM425364.1 BOLD
	HM425365.1 BOLD
	JX110641
L. cf. couesii (Italy)	DQ148290.1
	DQ148288.1
	DQ834549.1
	DQ834546.1
	DQ148289.1
	DQ834547.1
	DQ834548.1
	DQ834550.1
	DQ148287.1

L. couesii	**DQ310622.1**
	DQ889156.1
L. lemmoni	**GQ144447.1**
L. lubbocki	DQ148284.1
	DQ834542.1
	JX110643
	DQ148283.1
	DQ148282.1
	DQ834541.1
	DQ834540.1
L. packardi	**JX110642**
L. viridis	**JN175225**
	JN175226
	JN175227
	JN175228
	JN175229
L. cf. couesii (Sardinia)	**JX110640**
T. baeticus	FN691434.1
	JX110645
	FN691433.1
T. cancriformis	EF675884.1
	EF675896.1
	EF675892.1
	EF675885.1
	EF675895.1
	EF675887.1
	EF675888.1
	EF675893.1
	EF675886.1
	EF675891.1
	EF675894.1
	EF675897.1

EF675889.1

EF675890.1

FN691430.1

EF675899.1

EF675898.1

GQ328960.1

GQ144445.1

FN691431.1

DQ369312.1

EF675859.1

EF675860.1

EF675853.1

EF675832.1

EF675842.1

EF675843.1

EF675844.1

EF675850.1

EF675845.1

EF675831.1

EF675836.1

EF675835.1

EF675879.1

EF675851.1

EF675829.1

EF675849.1

EF675838.1

EF675837.1

EF675840.1

EF675861.1

EF675857.1

EF675830.1

EF675848.1

EF675854.1

EF675856.1

EF675852.1

EF675834.1

EF675846.1

EF675839.1

EF675833.1

EF675847.1

EF675841.1

EF675858.1

EF675855.1

DQ369315.1

JX110644

EF675869.1

EF675872.1

EF675866.1

EF675868.1

EF675878.1

DQ369317.1

EF675874.1

EF675875.1

EF675870.1

EF675864.1

EF675867.1

EF675873.1

EF675863.1

EF675871.1

EF675876.1

EF675865.1

EF675877.1

EF675862.1

DQ369314.1

	EF675827.1
	DQ369313.1
	FN691432.1
	EF675826.1
	NC 004465.1
	EF675828.1
	EF189678.1
	AB084514.1
	DQ148291.1
	DQ664196.1
	EF675880.1
	EF675881.1
	DQ369316.1
	EF675882.1
	EF675883.1
	JN175241.1
	JN175234.1
T. cf. australiensis sp. 1	**EF189677.1**
T. cf. australiensis sp. 2	**DQ310624.1**
	DQ889135.1
T. cf. australiensis sp. 3	**DQ310625.1**
T. cf. australiensis sp. A	**DQ343234.1**
T. cf. australiensis sp. B	**DQ343235.1**
T. cf. australiensis (Lake Carey)	**JN175235.1**
	JN175236.1
Triops cf australiensis (Baladonia Rock)	**JN175233.1**
	JN175237.1
	JN175244.1
T. cf australiensis (Ayers Rock)	**JN175245.1**
T. cf australiensis (Paynes Find)	**JN175242.1**
T. cf australiensis (Gibb Rock)	**JN190396.1**
	JN190398.1

T. cf australiensis (Walga Rock)	**JN175238.1**
	JN175232.1
	JN175230.1
	JN175231.1
	JN175240.1
	JN175239.1
	JN190397.1
	JN175238.1
T. cf. longicaudatus sp. 1	HQ908557.1
	HQ908544.1
	HQ908563.1
	HQ908559.1
	HQ908554.1
	HQ908552.1
	HQ908548.1
	HQ908547.1
	HQ908550.1
	HQ908567.1
	HQ908564.1
	HQ908551.1
	HQ908556.1
	HQ908546.1
	HQ908565.1
	HQ908558.1
	HQ908561.1
	HQ908549.1
	HQ908566.1
	HQ908555.1
	HQ908560.1
	HQ908553.1
	HQ908545.1
	HQ908562.1

T. emeritensis	EF675900.1
	FN691435.1
T. cf. granarius (Japan)	GQ144446.1
	JX110646
T. cf. granarius (Namibia)	**JX110639**
T. cf. granarius (South Africa)	**JN175223.1**
	JN175224.1
T. cf. longicaudatus sp. 2	HQ908538.1
	HQ908539.1
	HQ908526.1
	HQ908523.1
	HQ908542.1
	HQ908531.1
	HQ908524.1
	HQ908530.1
	JX110649
	HQ908519.1
	HQ908527.1
	HQ908525.1
	HQ908521.1
	HQ908534.1
	HQ908540.1
	HQ908541.1
	HQ908518.1
	HQ908522.1
	HQ908543.1
	HQ908537.1
	HQ908536.1
	HQ908533.1
	HQ908535.1
	HQ908529.1
	HQ908520.1

	HQ908517.1
	HQ908532.1
	HQ908528.1
	GU475465.1
	HM883938.1 BOLD
	HM883939.1 BOLD
	HM883941.1 BOLD
	HM883940.1 BOLD
	HM883942.1 BOLD
	GQ144444.1
	DQ310623.1
	JX110647
T. mauritanicus	EF675905.1
	EF675904.1
	EF675901.1
	EF675903.1
	EF675902.1
	FN691439.1
	FN691440.1
	FN691443.1
	FN691442.1
	FN691441.1
T. newberryi	HQ908510.1
	HQ908508.1
	HQ908509.1
	HQ908496.1
	HQ908507.1
	HQ908499.1
	HQ908501.1
	HQ908502.1
	HQ908512.1
	HQ908500.1

	HQ908498.1
	HQ908504.1
	HQ908505.1
	HQ908506.1
	HQ908503.1
	HQ908511.1
	HQ908513.1
	HQ908516.1
	HQ908514.1
	HQ908497.1
	JX110648
	HQ908515.1
	NC 006079.1
	AY639934.1
	JN175243.1
T. simplex	FN691438.1
	FN691436.1
	FN691437.1
T. cf. mauritanicus (E. Spain)	EF675906.1
	EF675907.1
	EF675908.1
T. cf. granarius (Russia)	**EF521890.1**
T. vicentinus	**FN691444.1**

Table 2: Accession numbers of sequences included in the supermatrix used for the divergence dating analysis (Chapter 1).

STU	12S	16S	COI	28S	EF1	RNA pol	Glyc synth
T. baeticus	KC466334	KC466343	JX110645		KC466360	KC466372	
T. cancriformis	KC466333	KC466342	JX110644	KC466348	KC466359	KC466368	KC466363
T. cf. australiensis sp. 1		EF189616	EF189677	EF189662	EF189595		
T. cf. australiensis sp. 2			DQ310624				
T. cf. australiensis sp. 3			DQ310625				
T. cf. australiensis sp. A	DQ343232		DQ343234				
T. cf. australiensis (Lake Carey)	JN175250		JN175235				
Triops cf australiensis (Baladonia Rock)	JN175259		JN175233				
T. cf australiensis (Ayers Rock)	JN175263		JN175245				
T. cf australiensis (Paynes Find)	JN175254		JN175242				
T. cf australiensis (Gibb Rock)	JN175265		JN190396				
T. cf australiensis (Walga Rock)	JN175253		JN175238				
T. cf. australiensis sp. B	DQ343233		DQ343235				
T. cf. granarius (Japan)	KC466335	KC466341	JX110646	KC466347	KC466358	KC466369	
T. cf. granarius (Namibia)	AM269423	AM269433	JX110639	AM269444			
T. cf. granarius (Tunisia)	AM269421	AM269431		AM269442			
T. cf. granarius (South Africa)	JN175248		JN175223				
T. cf. longicaudatus sp. 1			HQ908544				
T. cf. longicaudatus sp. 2	KC466336	KC466344	JX110649		KC466356	KC466371	
T. emeritensis	FN691428	AM183882	FN691435				
T. gadensis	FN691421	FN689863					
T. mauritanicus	AM184177	AM183873	FN691439				
T. newberryi	KC466337	KC466345	JX110648	KC466346	KC466357	KC466370	KC466364
T. simplex	AM184172	AM183867	FN691436				
T. cf. mauritanicus (E. Spain)			EF675907				

T. cf. granarius (Russia)			EF521890				
T. vicentinus	FN691426	FN689867	FN691444				
L. apus	AF494483	DQ148279	JX110638	KC466349	AF526293		
L. arcticus	AY159569 / AJ583699	DQ834538	JX110641	AF209047	KC466353	KC466365	KC466361
L. bilobatus	AJ000828						
L. couesii (Canada)	AJ000827		DQ310622				
L. cf. couesii (Apulia)	DQ148274	DQ148280	DQ834546				
L. cf. couesii (Sardinia)	KC466331	KC466339	JX110640	KC466351	KC466355	KC466366	KC466362
L. cryptus	AJ000824						
L. lemmoni	AY115604	AY115614	GQ144447				
L. lubbocki	KC466332	KC466340	JX110643	KC466350		KC466367	
L. packardi	KC466330	KC466338	JX110642	KC466352	KC466354		
L. viridis	JN175246		JN175225				
Artemia sp.	X69067	FJ007834	DQ401269	AY210805	GQ122208	U10331*	
Streptocephalus seali		JX439913	AY519832		AY305480	AY305628	GQ88703
Daphnia magna	JN903683	GQ343288	EU702133	AF532883	AB734039		GQ887666
Daphnia pulex	JN903685	JN874607	HM622593	AY630618	EFX85268	EFX75312	EFX74238
Eulimnadia sp.	AY779680	EF189604	FJ499139	FJ499231	FJ499103		
Limnadia lenticularis	AF494471	EF189609	FJ499183	FJ499284	AF063412	AF138989	GQ887703
Lynceus sp.	AF494479	EF189612	HQ966453	EF189653	AF526294	AY305581.1	GQ887686

Table 3: Optimum partitioning scheme and best fit models identified by PartitionFinder with model choice restricted to GTR and GTR + Γ.

Partition	Partition contents	Substitution model
1	12S	GTR +G
2	16S	GTR +G
3	28S	GTR +G
4	COI 1st position	GTR +G
5	COI 2nd position, EF1 2nd position, Glyc. Synth. 2nd position, RNA P. II 2nd position	GTR +G
6	COI 3rd position	GTR +G
7	EF1 1st position, Glyc. Synth. 1st position, RNA P. II 1st position	GTR +G
8	EF1 3rd position	GTR +G
9	Glyc. Synth. 3rd position, RNA P. II 3rd position	GTR +G

Table 4: Optimum partitioning scheme and best fit models identified by PartitionFinder with model choice restricted to those available in BEAST.

Partition	Partition contents	Substitution model
1	12S	GTR +G
2	16S	GTR +G
3	28S, Glyc. Synth. 1st position	GTR +G
4	COI 1st position	TrNef +G
5	COI 2nd position, EF1 2nd position, Glyc. Synth. 2nd position, RNA P. II 2nd position	GTR +G
6	COI 3rd position	TrN +G
7	EF1 1st position, RNA P. II 1st position	TrN +G
8	EF1 3rd position	GTR +G
9	Glyc. Synth. 3rd position, RNA P. II 3rd position	HKY +G

Table 5: Comparison of diversification models fitted to the BEAST time tree. Only Nostostraca and one outgroup were included in the analysis. Analysis conducted using LASER with the optimum diversification model selected based on AIC score (Rabosky 2006). The best scoring model is shaded.

Model	**Parameters**	**Model type**	**LH**	**r1**	**r2**	**r3**	**a**	**xp**	**k**	**st 1**	**st 2**	**AIC**	**dAIC**
pureBirth	1	RC	-81.70	0.0185	NA	NA	NA	NA	NA	NA	NA	165.40	23.95
bd	2	RC	-73.02	0.0000	NA	NA	0.9999	NA	NA	NA	NA	150.04	8.59
DDX	2	RV	-74.37	0.0044	NA	NA	NA	-0.5862	NA	NA	NA	152.74	11.29
DDL	2	RV	-81.70	0.0185	NA	NA	NA	NA	645575	NA	NA	167.40	25.95
yule2rate	3	RV	-70.34	0.0035	0.0299	NA	NA	NA	NA	73.28	NA	146.68	5.23
yule3rate	5	RV	-65.72	0.0035	0.0365	0.0043	NA	NA	NA	73.28	6.00	141.45	0.00

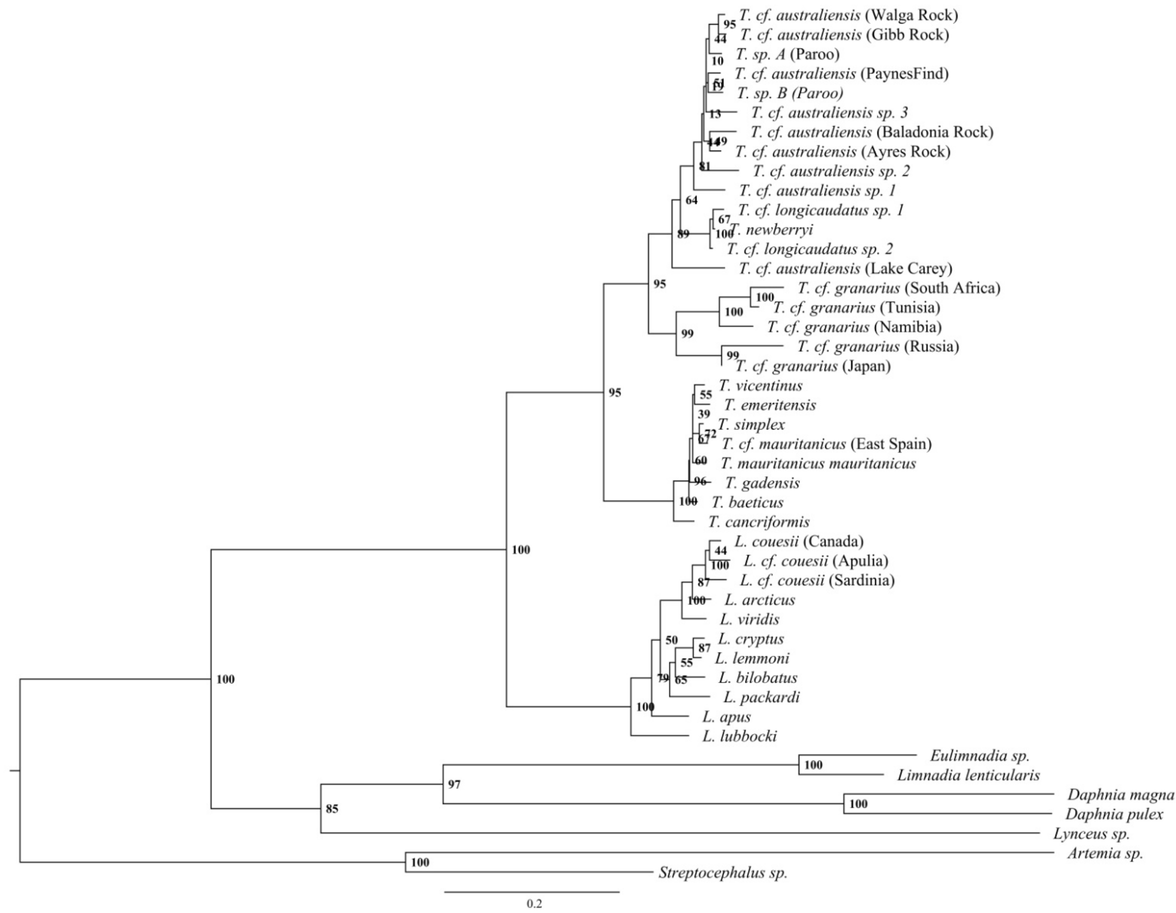


Figure 1: Maximum likelihood phylogeny of Branchiopoda inferred with RAxML using a GTR + Γ substitution model for each partition. Data was partitioned according to the best scoring scheme identified by PartitionFinder. Numbers at nodes give RAxML rapid bootstrap support values (100 replicates) with values less than 50 not shown.

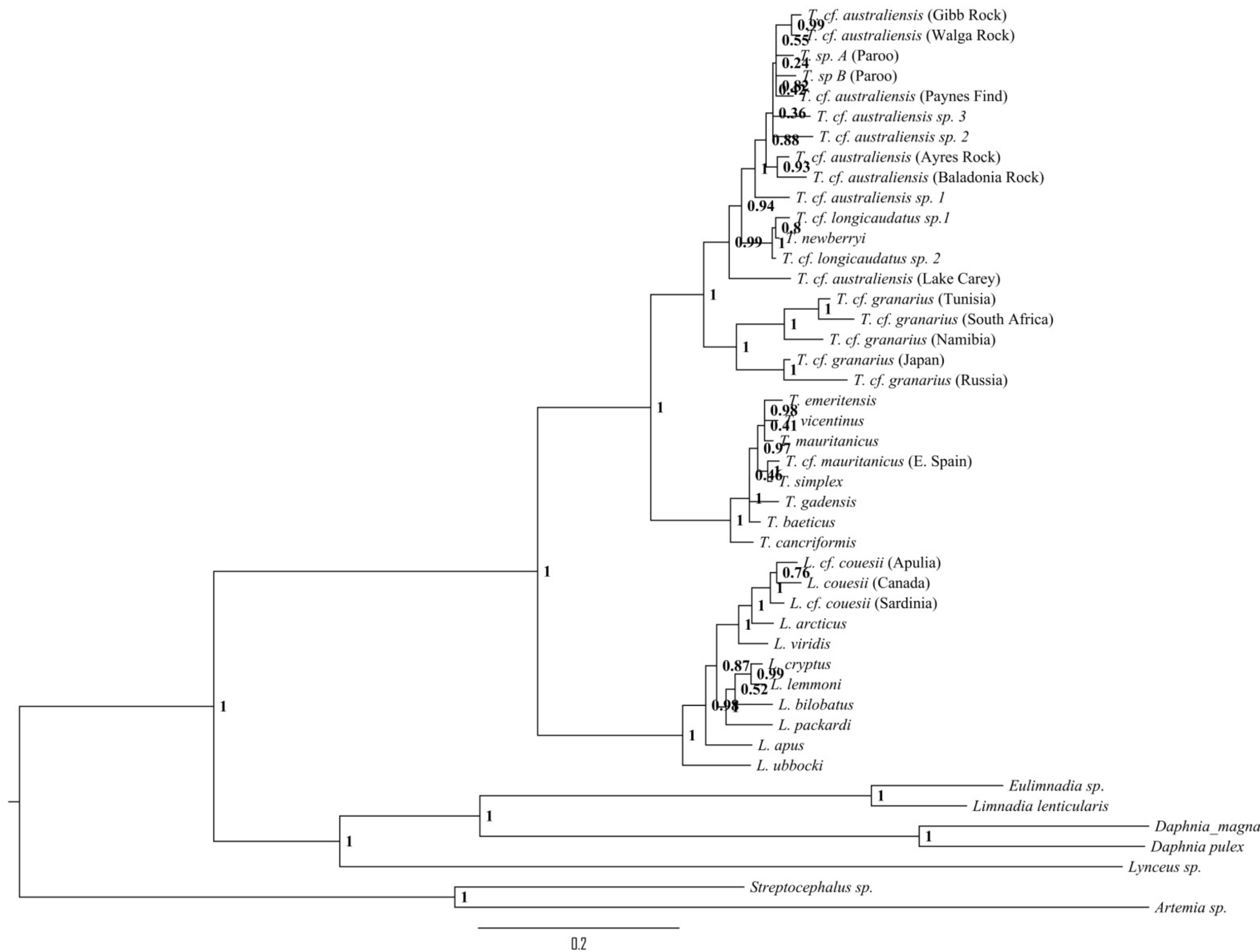


Figure 2: Bayesian phylogeny of Notostraca based on a 7 gene supermatrix inferred with MrBayes. Data was partitioned based on the optimum strategy identified by PartitionFinder under BIC. A GTR + Γ model of nucleotide evolution was specified for each partition. Values at nodes show posterior probabilities with values less than 50 not shown.

Appendix 2: Additional information for Chapter 3

Table 1: Details of samples collected for this study with sample type, location, coordinates, inferred STU following the GMYC barcoding analysis and details of genes sequenced with accession numbers for each sample. For sample type GB = GenBank data, RE = resting egg, LR = lab reared, WC = wild caught.

					Genes Sequenced with accession numbers						
Sample I.D.	**Sample type**	**Location**	**Coordinates**	**STU I.D. (From GMYC)**	**COI**	**12S**	**16S**	**28S**	**EF1-a**	**RNA PII**	**Glyc. Synth.**
T_m_D2	LR	Doñana national park, Spain		*Triops baeticus*	JX110645	KC466334	KC466343	-	KC466360	KC466372	-
T_c_ESP	LR	Espolla, Spain	42°09'02" N / 002°45'60" E	*Triops cancriformis*	JX110644	KC466333	KC466342	KC466348	KC466359	KC466368	KC466363
T_g_CC	RE	Cornish Crispa Co. commercial kit, "Triops granarius"		*Triops granarius* (Japan)	JX110646	KC466335	KC466341	KC466347	KC466358	KC466369	-
T_g_Namibia	WC	Arandis, Namibia	22°24'27'' S / 14°58'24'' E	*Triops granarius* (Namibia)	JX110639	-	-	-	-	-	-
T_long_CW	LR	Clearwater, Kansas, USA		*T. cf. longicaudatus sp. 2*	JX110647	KC466336	JX110639	-	KC466356	KC466371	-
T_bb	LR	Corni Crispa Co. Comercial kit, "Black Beauty Triops"		*T. cf. longicaudatus sp. 2*	JX110649	-	-	-	-	-	-
T_n	RE	Shallow Water,		*Triops*	JX110648	KC466337	KC466345	KC466346	KC466357	KC466370	KC466364

		Kansas, USA		*newberryi*						
L_apus	WC	Frankfurt-Oder, Germany	52°19'29" N / 14°33'33" E	*Lepidurus apus*	JX110638	-	-	KC466349	-	-	-
L_arc	WC	Strandvolldammane, Svalbard	78.96061 N / 11.50817 E	*Lepidurus arcticus*	JX110641	-	-	-	KC466353	KC466365	KC466361
Ll_MS	WC	Castel Porziano Estate, Italy	41.74748 N / 12.41860 E	*Lepidurus lubbocki*	JX110643	KC466332	KC466340	KC466350	-	KC466367	-
L_pack	LR	Jeppsen Prarie Reserve, California		*Lepidurus packardi*	JX110642	KC466330	KC466338	KC466352	KC466354	-	-
L_sp_Sardinia	WC	Giara di Gesturi, Sardinia	39°44'30" N / 8°59'51" E	*Lepidurus sp.* (Sardinia)	JX110640	KC466351	KC466355	KC466351	KC466355	KC466366	KC466362

Table 2: Primers and PCR conditions used for the amplification and sequencing of mitochondrial and nuclear markers. A general PCR program was used with annealing temperatures given below: 94 ºC 2 min, (94 ºC 30 s., annealing temp. 30 s. 72 ºC 1 min) x 35, 72 ºC 5 min.

Gene	Name	Primer sequence (5'->3')	Annealing temp. (ºC)	$MgCL_2$ conc. (mM)	Reference
12S	12S_col_F	ATGCACTTTCCAGTACATCTAC	50	1	(Colbourne & Hebert 1996)
	12S_col_R	AAATCGTGCCAGCCGTCGC			
16S	16Sar	CGCCTGTTTATCAAAAACAT	50	1	(Palumbi et al. 1991)
	16Sbr	CCGGTCTGAACTCAGATCACGT			
28S	28S_L1	AGCGGAGGAAAAGAAACTA	58	2	(Korn & Hundsdoerfer 2006)
	28S_H4lv1	ACGATCGATTTGCACGTCAG			
COI	LCO1490	GGTCAACAAATCATAAAGATATTGG	52	2	(Folmer et al. 1994)
	HC02198	TAAACTTCAGGGTGACCAAAAAATCA			
Elongation factor 1 α	EF1_F	TCGAAACCGCCAAGTTCTAC	Touch down 55 – 45*	3	This Study
	EF1_R	GTAACCACGACGCAATTCCT			
Glycogen synthase	Glyc_F	CGACGAAGAAGCTGGAAAAC	Touch down 55 – 45*	1.5	This Study
	Glyc_R	GACAAAAACTCGGGATGGAA			
RNA polymerase II	RNA_P_F	GGATTTCTCCWGGAGATACSAAGGTCATGG	Touch down 55 – 45*	2	This Study
	RNA_P_R	TCGACGTTCTGTTGACCKACGCAAGC			

* -0.4 per cycle for 25 cycles, then 10 cycles at 45ºC.

Table 3: Results of the STU cluster analysis on Notostraca COI sequences. Accession numbers of Notostraca COI sequences used in the GMYC analysis are provided, as well as the reference for the STU ID. New accessions for this study are in bold.

STU/Cluster ID	Accession	Reference
L. apus	DQ834543.1	(Mantovani et al. 2004; Mantovani et al. 2009)
	DQ834544.1	
	DQ148285.1	
	EF189669.1	
	JX110638	
L. arcticus	HM425362.1 BOLD	(Rogers 2001)
	AF209067.1	
	DQ834545.1	
	HM425363.1 BOLD	
	HM425361.1 BOLD	
	DQ148286.1	
	HM425364.1 BOLD	
	HM425365.1 BOLD	
	JX110641	
L. cf. couessi (Italy)	DQ148290.1	(Mantovani et al. 2009)
	DQ148288.1	
	DQ834549.1	
	DQ834546.1	
	DQ148289.1	
	DQ834547.1	
	DQ834548.1	
	DQ834550.1	
	DQ148287.1	
L. couessi	DQ310622.1	(Rogers 2001)
	DQ889156.1	
L. lemmoni	GQ144447.1	(Rogers 2001)
L. lubbocki	DQ148284.1	(Mantovani et al. 2009)
	DQ834542.1	

	JX110643	
	DQ148283.1	
	DQ148282.1	
	DQ834541.1	
	DQ834540.1	
L. packardi	**JX110642**	(Rogers 2001)
L. cf. couesii (Sardinia)	**JX110640**	
T. baeticus	FN691434.1	(Korn et al. 2010)
	JX110645	
	FN691433.1	
T. cancriformis	EF675884.1	(Korn et al. 2006)
	EF675896.1	
	EF675892.1	
	EF675885.1	
	EF675895.1	
	EF675887.1	
	EF675888.1	
	EF675893.1	
	EF675886.1	
	EF675891.1	
	EF675894.1	
	EF675897.1	
	EF675889.1	
	EF675890.1	
	FN691430.1	
	EF675899.1	
	EF675898.1	
	GQ328960.1	
	GQ144445.1	
	FN691431.1	
	DQ369312.1	

EF675859.1

EF675860.1

EF675853.1

EF675832.1

EF675842.1

EF675843.1

EF675844.1

EF675850.1

EF675845.1

EF675831.1

EF675836.1

EF675835.1

EF675879.1

EF675851.1

EF675829.1

EF675849.1

EF675838.1

EF675837.1

EF675840.1

EF675861.1

EF675857.1

EF675830.1

EF675848.1

EF675854.1

EF675856.1

EF675852.1

EF675834.1

EF675846.1

EF675839.1

EF675833.1

EF675847.1

EF675841.1

EF675858.1

EF675855.1

DQ369315.1

JX110644

EF675869.1

EF675872.1

EF675866.1

EF675868.1

EF675878.1

DQ369317.1

EF675874.1

EF675875.1

EF675870.1

EF675864.1

EF675867.1

EF675873.1

EF675863.1

EF675871.1

EF675876.1

EF675865.1

EF675877.1

EF675862.1

DQ369314.1

EF675827.1

DQ369313.1

FN691432.1

EF675826.1

NC 004465.1

EF675828.1

EF189678.1

AB084514.1

DQ148291.1

	DQ664196.1	
	EF675880.1	
	EF675881.1	
	DQ369316.1	
	EF675882.1	
	EF675883.1	
T. cf. australiensis sp. 1	EF189677.1	This study
T. cf. australiensis sp. 2	DQ310624.1	This study
	DQ889135.1	
T. cf. australiensis sp. 3	DQ310625.1	This study
T. cf. australiensis sp. A	DQ343234.1	(Murugan et al. 2009)
T. cf. australiensis sp. B	DQ343235.1	(Murugan et al. 2009)
T. cf. longicaudatus sp. 1	HQ908557.1	(Macdonald et al. 2011)
	HQ908544.1	
	HQ908563.1	
	HQ908559.1	
	HQ908554.1	
	HQ908552.1	
	HQ908548.1	
	HQ908547.1	
	HQ908550.1	
	HQ908567.1	
	HQ908564.1	
	HQ908551.1	
	HQ908556.1	
	HQ908546.1	
	HQ908565.1	
	HQ908558.1	
	HQ908561.1	
	HQ908549.1	
	HQ908566.1	
	HQ908555.1	

	HQ908560.1	
	HQ908553.1	
	HQ908545.1	
	HQ908562.1	
T. emeritensis	EF675900.1	(Korn et al. 2010)
	FN691435.1	
T. cf. granarius (Japan)	GQ144446.1	(Korn & Hundsdoerfer 2006)
	JX110646	
T. cf. granarius (Namibia)	**JX110639**	(Korn & Hundsdoerfer 2006)
T. cf. longicaudatus sp. 2	HQ908538.1	This study / (Macdonald et al. 2011)
	HQ908539.1	
	HQ908526.1	
	HQ908523.1	
	HQ908542.1	
	HQ908531.1	
	HQ908524.1	
	HQ908530.1	
	JX110649	
	HQ908519.1	
	HQ908527.1	
	HQ908525.1	
	HQ908521.1	
	HQ908534.1	
	HQ908540.1	
	HQ908541.1	
	HQ908518.1	
	HQ908522.1	
	HQ908543.1	
	HQ908537.1	
	HQ908536.1	
	HQ908533.1	
	HQ908535.1	

	HQ908529.1	
	HQ908520.1	
	HQ908517.1	
	HQ908532.1	
	HQ908528.1	
	GU475465.1	
	HM883938.1 BOLD	
	HM883939.1 BOLD	
	HM883941.1 BOLD	
	HM883940.1 BOLD	
	HM883942.1 BOLD	
	GQ144444.1	
	DQ310623.1	
	JX110647	
T. mauritanicus	EF675905.1	(Korn et al. 2010; Korn et al. 2006)
	EF675904.1	
	EF675901.1	
	EF675903.1	
	EF675902.1	
	FN691439.1	
	FN691440.1	
	FN691443.1	
	FN691442.1	
	FN691441.1	
T. newberryi	HQ908510.1	(Macdonald et al. 2011)
	HQ908508.1	
	HQ908509.1	
	HQ908496.1	
	HQ908507.1	
	HQ908499.1	
	HQ908501.1	
	HQ908502.1	

	HQ908512.1	
	HQ908500.1	
	HQ908498.1	
	HQ908504.1	
	HQ908505.1	
	HQ908506.1	
	HQ908503.1	
	HQ908511.1	
	HQ908513.1	
	HQ908516.1	
	HQ908514.1	
	HQ908497.1	
	JX110648	
	HQ908515.1	
	NC 006079.1	
	AY639934.1	
T. simplex	FN691438.1	(Korn et al. 2010)
	FN691436.1	
	FN691437.1	
T. cf. mauritanicus (Ares)	EF675906.1	This study
	EF675907.1	
	EF675908.1	
T. cf. granarius (Russia)	FN691444.1	This study
T. vicentinus	FN691444.1	(Korn et al. 2010)

Table 4: Samples and accession numbers of the sequences used for phylogenetic analysis. GB = GenBank data, RE = resting egg, LR = lab reared, WC = wild caught. Newly generated sequences are shown in bold.

	Genetic data used with accession number						
STU	**12S**	**16S**	**COI**	**28S**	**EF1α**	**RNA PII**	**Glyc. Synth**
T. baeticus	**KC466334**	**KC466343**	**JX110645**	-	**KC466360**	**KC466372**	-
T. cancriformis	**KC466333**	**KC466342**	**JX110644**	**KC466348**	**KC466359**	**KC466368**	**KC466363**
T. cf. australiensis sp. 1	-	EF189616	EF189677	EF189662	EF189595	-	-
T. cf. australiensis sp. 2	-	-	DQ310624	-	-	-	-
T. cf. australiensis sp. 3	-	-	DQ310625	-	-	-	-
T. cf. australiensis sp. A	DQ343232	-	DQ343234	-	-	-	-
T. cf. australiensis sp. B	DQ343233	-	DQ343235	-	-	-	-
T. cf. granarius (Japan)	**KC466335**	**KC466341**	**JX110646**	**KC466347**	**KC466358**	**KC466369**	-
T. cf. granarius (Namibia)	AM269423	AM269433	**JX110639**	AM269444	-	-	-
T. cf. granarius (Tunisia)	AM269421	AM269431	-	AM269442	-	-	-
T. cf. longicaudatus sp. 1	-	-	HQ908544	-	-	-	-
T. cf. longicaudatus sp.2	**KC466336**	**JX110639**	**JX110647**	-	**KC466356**	**KC466371**	-
T. emeritensis	FN691428	AM183882	FN691435	-	-	-	-
T. gadensis	FN691421	FN689863	-	-	-	-	-
T. mauritanicus	AM184177	AM183873	FN691439	-	-	-	-
T. newberryi	**KC466337**	**KC466345**	**JX110648**	**KC466346**	**KC466357**	**KC466370**	**KC466364**
T. simplex	AM184172	AM183867	FN691436	-	-	-	-
T. cf. mauritanicus (E Spain)	-	-	EF675907	-	-	-	-
T. cf. granarius (Russia)	-	-	EF521890	-	-	-	-
T. vicentinus	FN691426	FN689867	FN691444	-	-	-	-
L. apus	AF494483	DQ148279	**JX110638**	**KC466349**	AF526293	-	-
L. arcticus	AY159569 / AJ583699 *	DQ834538	**JX110641**	AF209047	**KC466353**	**KC466365**	**KC466361**
L. bilobatus	AJ000828	-	-	-	-	-	-
L. couesii (Canada)	AJ000827	-	DQ310622	-	-	-	-

L. cf. couesii (Italy)	DQ148274	DQ148280	DQ834546	-	-	-	-
L. cf. couesii (Sardinia)	**KC466351**	**KC466355**	**JX110640**	**KC466351**	**KC466355**	**KC466366**	**KC466362**
L. cryptus	AJ000824	-	-	-	-	-	-
L. lemmoni	AY115604	AY115614	GQ144447	-	-	-	-
L. lubbocki	**KC466332**	**KC466340**	**JX110643**	**KC466350**	-	**KC466367**	-
L. packardi	**KC466330**	**KC466338**	**JX110642**	**KC466352**	**KC466354**	-	-
Leptestheria sp.(Outgroup)	DQ872782	EF189606	DQ872786	-	FJ499045	-	-

* two concatenated sequences were used for this gene fragment.

Table 5: Detailed, population by population, Notostraca sexual system data. STU I.D., population locations, the presence of genetic barcoding data (COI, 12S or 16S) to support species assignment, sex ratio (percent male), and sample sizes (where available), histology data (presence of ovotestis in individuals) and the inferred sexual system used in the character evolution analysis is given. Where genetic data for a population is published in a separate study to the sexual system data both references are given.

Significant Taxonomic Units (STU)	Location	Genetic barcode for population	Sex ratio	N	Reproduction in isolation	Ovotestis	Sexual system	References
T. baeticus	Southern Iberia – Multiple populations	Y	51	255			Gonochoric	(Korn et al. 2010)
T. cancriformis	Across Europe	Y	0 -53		Yes[1]	Yes [1]	Gonochoric / AD	(Bernard 1891; Longhurst 1954; Longhurst 1955c; Zierold et al. 2007; Zierold et al. 2009)
T. cf. australiensis sp. 1		-					No data	
T. cf. australiensis sp. 2		-					No data	
T. cf. australiensis sp. 3		-					No data	
T. cf. australiensis sp. A	Bourke, Paroo, Australia	Y	47	15		No	Gonochoric	(Murugan et al. 2009)
T. cf. australiensis sp. B	Bourke, Paroo, Australia	Y	0	22		Yes	AD	(Murugan et al. 2009)
T. emeritensis	Badajoz, Spain – Multiple populations	Y	55	29			Gonochoric	(Korn et al. 2010)
T. gadensis	Cadiz, Spain – Multiple populations	Y	36	52			Equivocal	(Korn et al. 2010)
T. cf. granarius (Japan)	3 locations: Shizuoka, Kagawa, Fukuoka	Y	49	388	No		Gonochoric	(Longhurst 1954; Longhurst 1955c; Mitsumoto & Yahata 2006; SunoUchi et al. 1997)
T. cf. granarius (Namibia)		Y	Even[2]		No		Gonochoric	(Korn & Hundsdoerfer 2006; Longhurst 1954; Longhurst 1955c)

T. cf. granarius (Tunisia)		Y	Even[2]		No		Gonochoric	(Korn & Hundsdoerfer 2006; Longhurst 1954; Longhurst 1955c)
T. cf. longicaudatus sp. 1 (Long bodied - Macdonald)	Northern Chihuahuan desert, USA	Y	48	21			Gonochoric	(Macdonald et al. 2011)
T. cf. longicaudatus sp. 2	Zacatecas, Mexico	Y	53	108		No	Gonochoric	(Garcia-Velazco et al. 2009; Murugan et al. 2002)
T. cf. longicaudatus sp. 2	Clearwater, Kansas	Y	50	6			Gonochoric	Own data
T. cf. longicaudatus sp. 2	Northern Chihuahuan desert, USA	Y	0	212			AD	(Macdonald et al. 2011)
T. newberryi	Multiple populations from Northern Chihuahuan desert, USA	Y	0 - 27	427			AD	(Macdonald et al. 2011)
T. newberryi	Multiple populations from California, Nevada and Utah	N	0 – 16	487	Yes		AD	(Sassaman 1991)
T. mauritanicus		-					No data	
T. simplex	Kairouan, Tunisia	Y	53	17			Gonochoric	(Korn et al. 2006)
T. vicentinus	SW Portugal – Multiple populations	Y	50	70			Gonochoric	(Korn et al. 2010)
T. vicentinus	Vila do Bispo, SW Portugal, 2 populations	Y	58	118			Gonochoric	(Machado et al. 1999)
L. apus	Frankfurt, Germany	Y	0	62			AD	Own data
L. apus	Germany	N	5	411			AD	(Stephan 2008)
L. apus	Gosseberg,	N	1	2712			AD	(Simon 1886)

	Austria						
L. apus	Ivry, France	N	18	174		AD	(Simon 1886)
L. arcticus	Rasshua,Kuril Islands, Russia		7	60		AD	
L. arcticus	Cambridge Bay, Canada	N	7	14		AD	(Linder 1952)
L. arcticus	Ooglamie, Alaska	N	0	113		AD	(Linder 1952)
L. arcticus	Broughton Island, Canada	N	1	109		AD	(Bushnell & Byron 1979)
L. arcticus	Spitsbergen, Svalbard	N	0	200		AD	(Wojtasik & Brylka-Wolk 2010)
L. bilobatus	Irish Canyon, Moffat, Colorado, USA	Y	35	116		Equivocal	(King & Hanner 1998; Saunders 1980)
L. couesii	Frenchman River, Montana, USA	N	51	63		Gonochoric	(Packard 1875)
L. cf. couesii (Italy)		-				No data	-
L. cryptus		-				No data	-
L. lemmoni	Grand Couleee, Washington State, USA	N	56	96		Gonochoric	(Linder 1952)
L. lemmoni	Lassen County, California, USA	N	Even[2]	247		Gonochoric	(Lynch 1966)
L. lubbocki	Israel, 9 populations	N	Median = 54			Gonochoric	(Kuller & Gasith 1996)
L. lubbocki	Castel Porziano Estate, Italy	Y	Even [2]		No	Gonochoric	(Mantovani et al. 2004; Mantovani et al. 2009; Scanabissi & Mondini 2002)
L. packardi	Tehama County, California	N	~ 50 (over several months)			Gonochoric	(Ahl 1991)
L. cf. couesii (Sardinia)	Giara di Gesturi, Sardinia	Y	54	208		Gonochoric	(Margraf & Maass 1982)

Table 6: Latitude data for taxa with known sexual system. Sexual system (AD) was coded either as androdioecious or not.

STU	AD	Latitude
T. baeticus	No	36
T. cancriformis	Yes	42
T. cf. australiensis sp. A	No	29
T. cf. australiensis sp. B	Yes	29
T. cf. granarius (Japan)	No	36
T. cf. granarius (Namibia)	No	22
T. cf. granarius (Tunisia)	No	35
T. cf. longicaudatus sp.2	Yes	37
T. cf. longicaudatus sp. 1	No	32
T. emeritensis	No	39
T. newberryi	Yes	38
T. simplex	No	36
T. vicentinus	No	37
L. apus	Yes	52
L. arcticus	Yes	78
L. cf. couesii (Sardinia)	No	39
L. couesii (Canada)	No	53
L. lemmoni	No	34
L. lubbocki	No	41
L. packardi	No	38

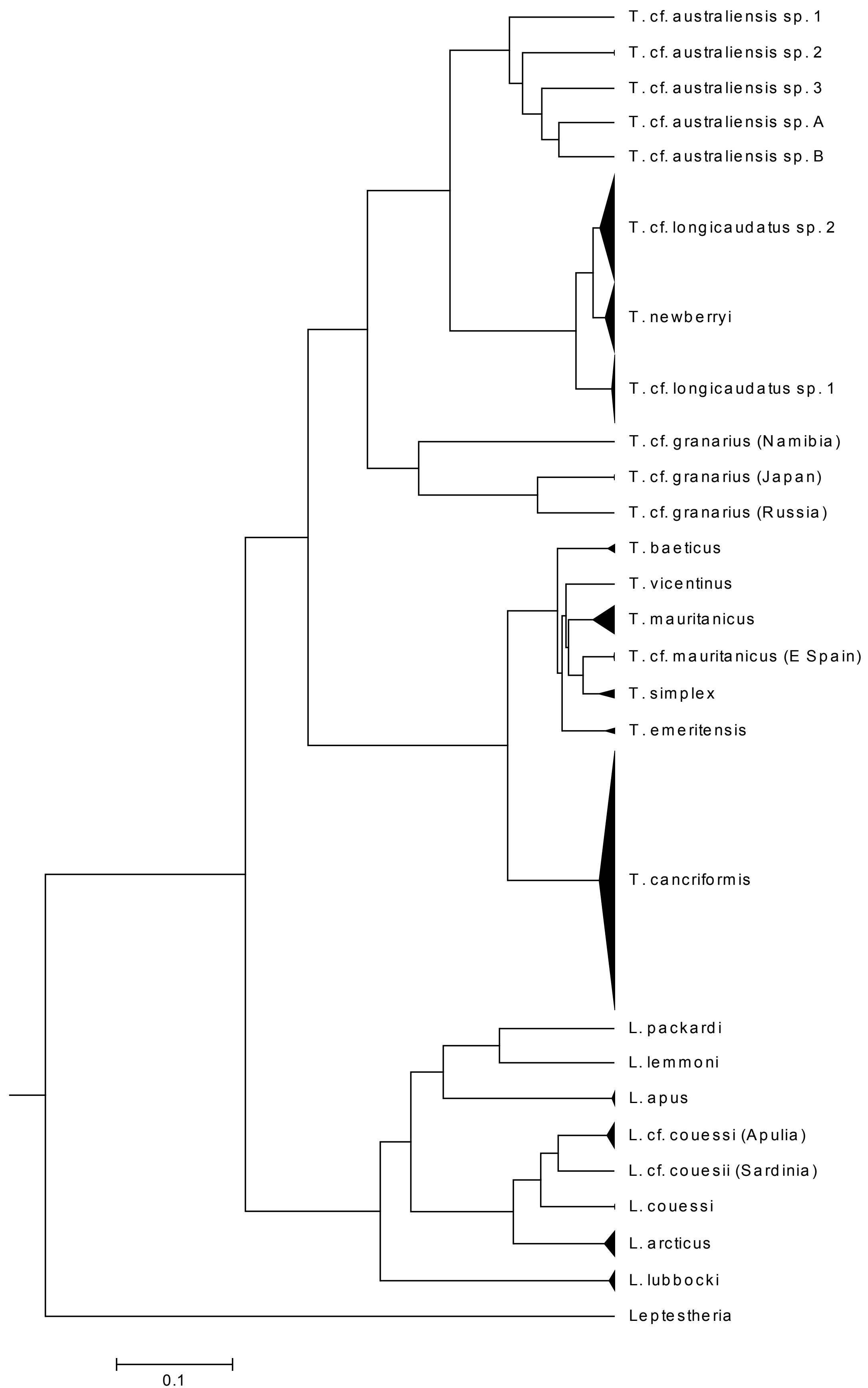


Figure 1: Ultrametric consensus phylogeny created in BEAST based on aligned COI sequences from GenBank and our newly generated sequence data showing collapsed clusters identified by GMYC analysis. Accessions within each cluster are listed in Appendix 2 Table 3.

Appendix 3: Additional analysis for Chapter 3

Relaxed clock phylogenetic analysis

BEAST v1.7.4 (Drummond et al. 2012) was used to create two ultrametric phylogenies; one based on the full dataset and a second using a reduced dataset containing only the mitochondrial (mt) genes 12S, 16S and COI. We used an uncorrelated log normal relaxed clock and Yule speciation prior for both analyses. As with the unconstrained ML and Bayesian phylogenetic analyses each gene was treated as a separate partition with substitution model parameters estimated independently (unlinked). All partitions were initially assigned a GTR + gamma model of sequence evolution. However, test runs showed poor mixing of some GTR model parameters for the 12S, 16S, 28S and RNA polymerase II partitions. We therefore simplified the substitution model to HKY + gamma for these partitions (12S and 16S only in the mt analysis). Two independent MCMC chains were run for 100 million iterations, sampling every 10,000 iterations, for each analysis. Runs were checked for convergence and adequate mixing using Tracer v1.5. ESS values were greater than 1000 for all parameters in all runs. A posterior sample of 8,000 trees from one of the runs for each analysis was used to create a maximum clade credibility tree with TreeAnotator (http://beast.bio.ed.ac.uk/TreeAnnotator) (see Appendix 3, Figures 1 and 2). These two trees were then used to run the ML character mapping analysis in BayesTraits (Appendix 3, Table 1).

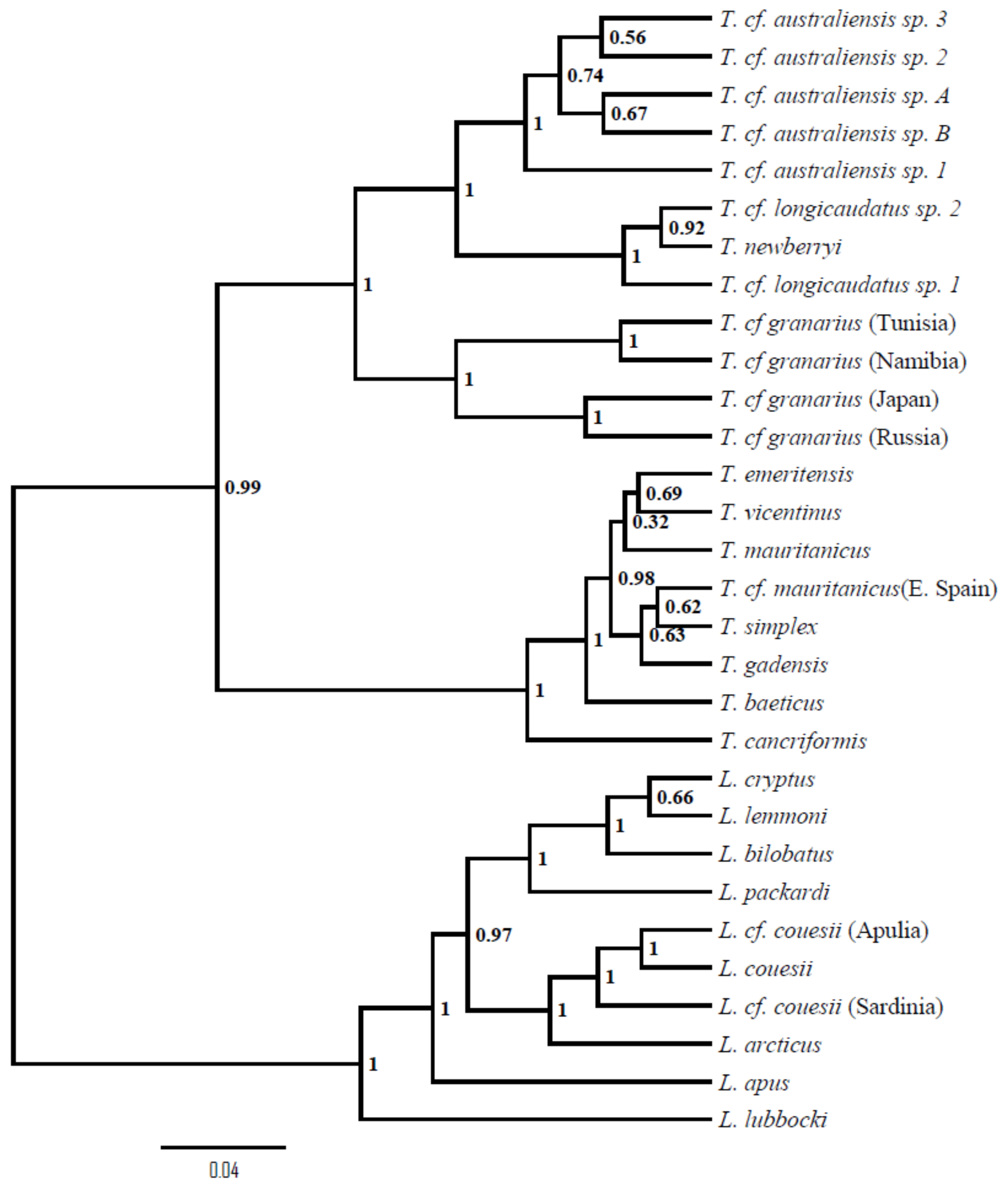


Figure 1: Ultrametric BEAST maximum clade credibility tree based on the full dataset. Values at nodes show posterior probabilities.

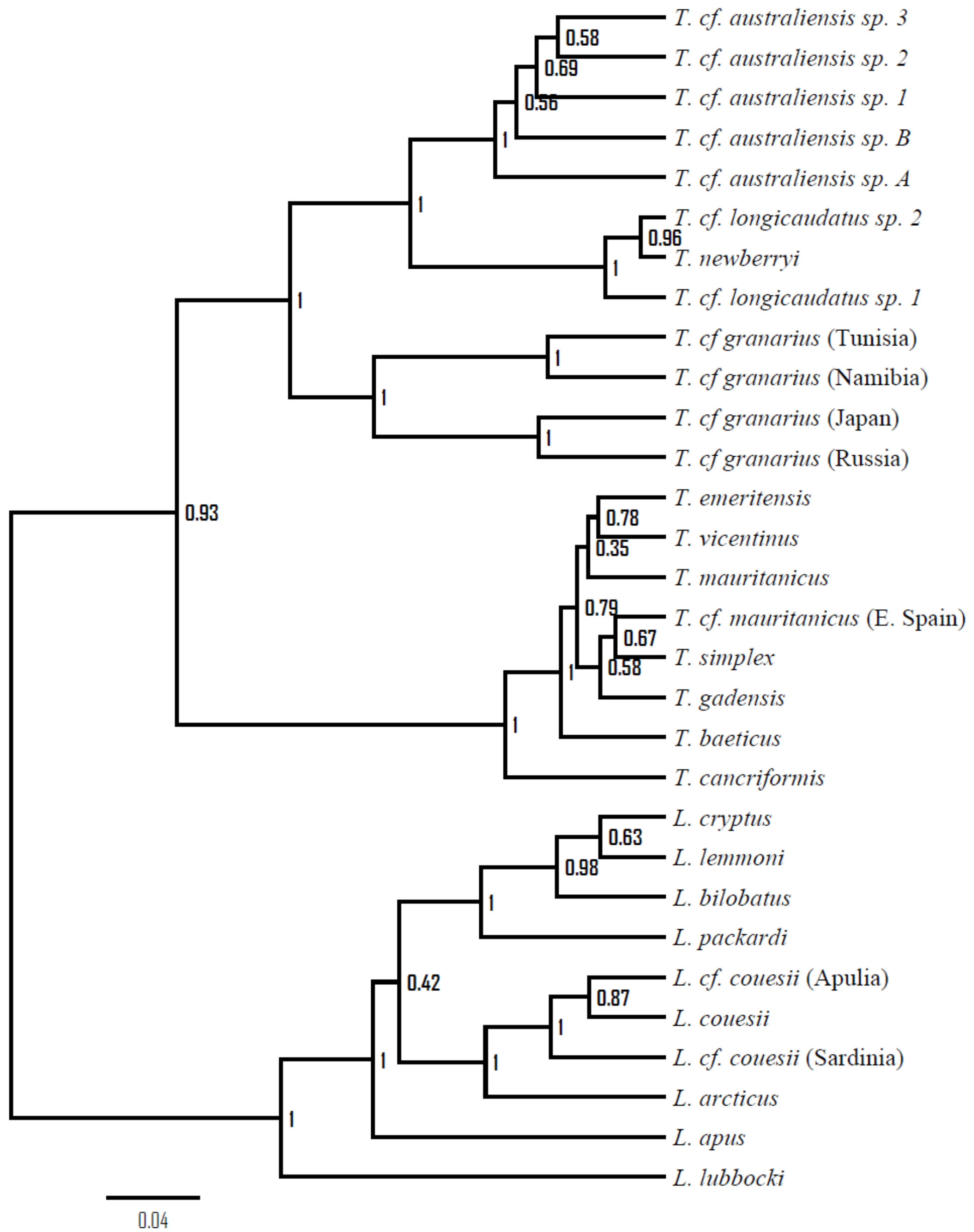


Figure 2: Ultrametric BEAST maximum clade credibility tree based on the reduced dataset (only mitochondrial genes). Values at nodes show posterior probabilities.

Table 1: Comparison of BayesTraits character mapping results using three alternative phylogenetic trees – an ML unconstrained phylogeny using the full dataset, an ultrametric phylogeny created with BEAST under a lognormal relaxed molecular clock based on the full dataset and an ultrametric phylogeny created with BEAST under a lognormal relaxed molecular clock based on a reduced dataset containing only the mitochondrial genes 12S, 16S and COI. Under each phylogeny, models of transitions in sexual system were compared using the ML implementation of BayesMultistate. lnL = log-likelihood of model, q_{GA} = transition rate from gonochorism to androdioecy, q_{AG} = transition rate from androdioecy to gonochorism, p = p-value for D with 1 degree of freedom comparing the restricted models to the unrestricted model. The best fitting model under each phylogeny is highlighted in green.

Phylogenetic tree	Model	lnL	q_{GA}	q_{AG}	p
Unconstrained ML (RAxML), full dataset	Unrestricted	-7.229	23.729	110.640	-
	Equal rates	-10.262	5.730	5.730	0.0138
	AD to gonochorism only	-16.855	0.000	6.673	< 0.0001
	Gonochorism to AD only	-12.0864	1.755	0.000	0.0018
Ultrametric (BEAST), lognormal relaxed molecular clock, full dataset	Unrestricted	-8.841	25.144	87,939	-
	Equal rates	-11.013	6.776	6.776	0.0371
	AD to gonochorism only	-13.715	0.000	8.686	0.0018
	Gonochorism to AD only	-12.334	2.681	0.000	0.0082
Ultrametric (BEAST), lognormal relaxed molecular clock, Mt only	Unrestricted	-8.841	133.513	467.296	-
	Equal rates	-10.975	2.036	2.036	0.0389
	AD to gonochorism only	-13.340	0	6.395	0.0027
	Gonochorism to AD only	-12.209	1.975	0	0.0095

Appendix 4: Additional information for Chapter 4

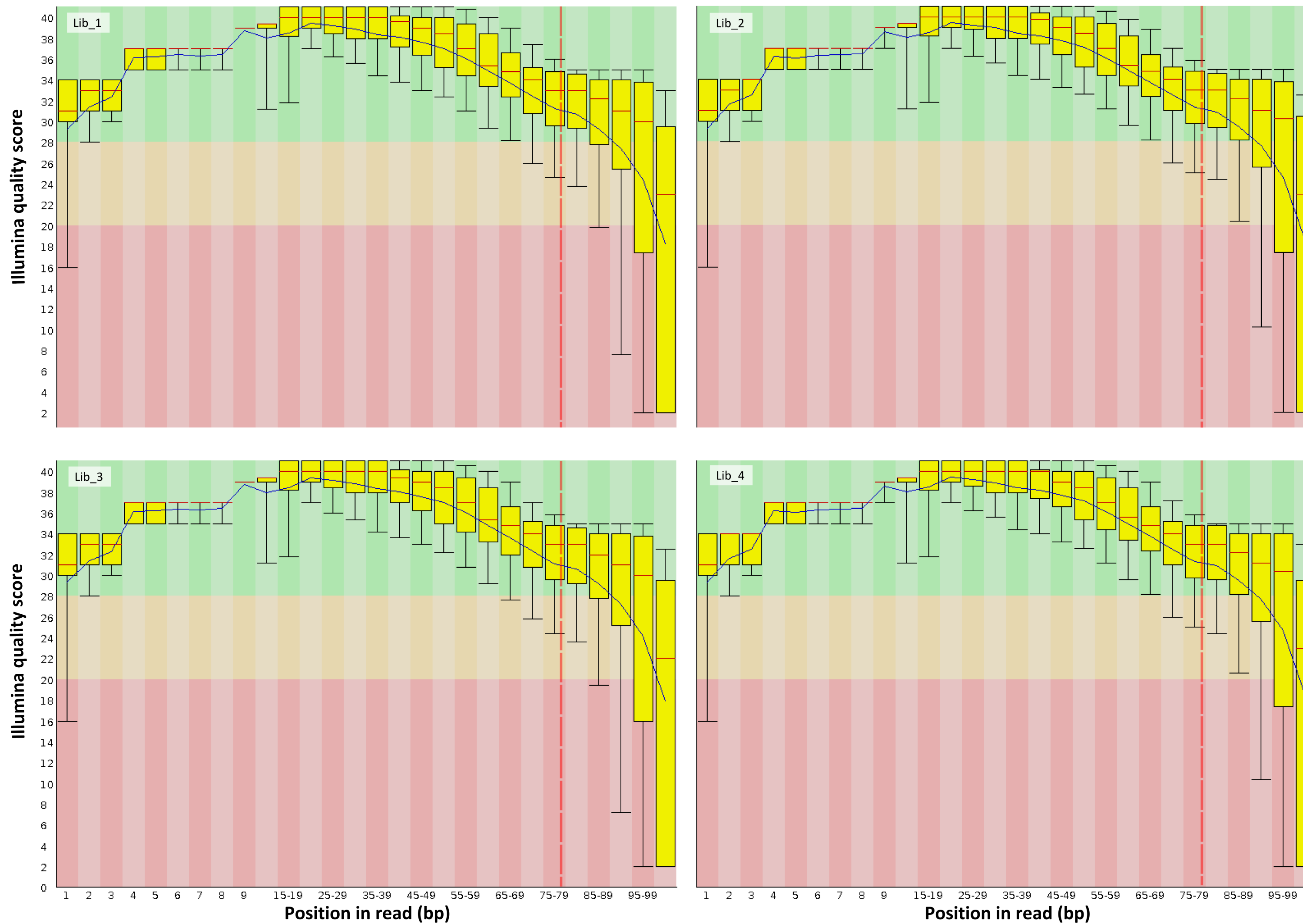


Figure 1: Box plots generated by FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc) showing per base quality scores for raw first end reads from each of the four RAD libraries. The dashed red lines indicate the length at which reads were trimmed to during processing. Quality scores are in Sanger / Illumina v1.9 encoding.

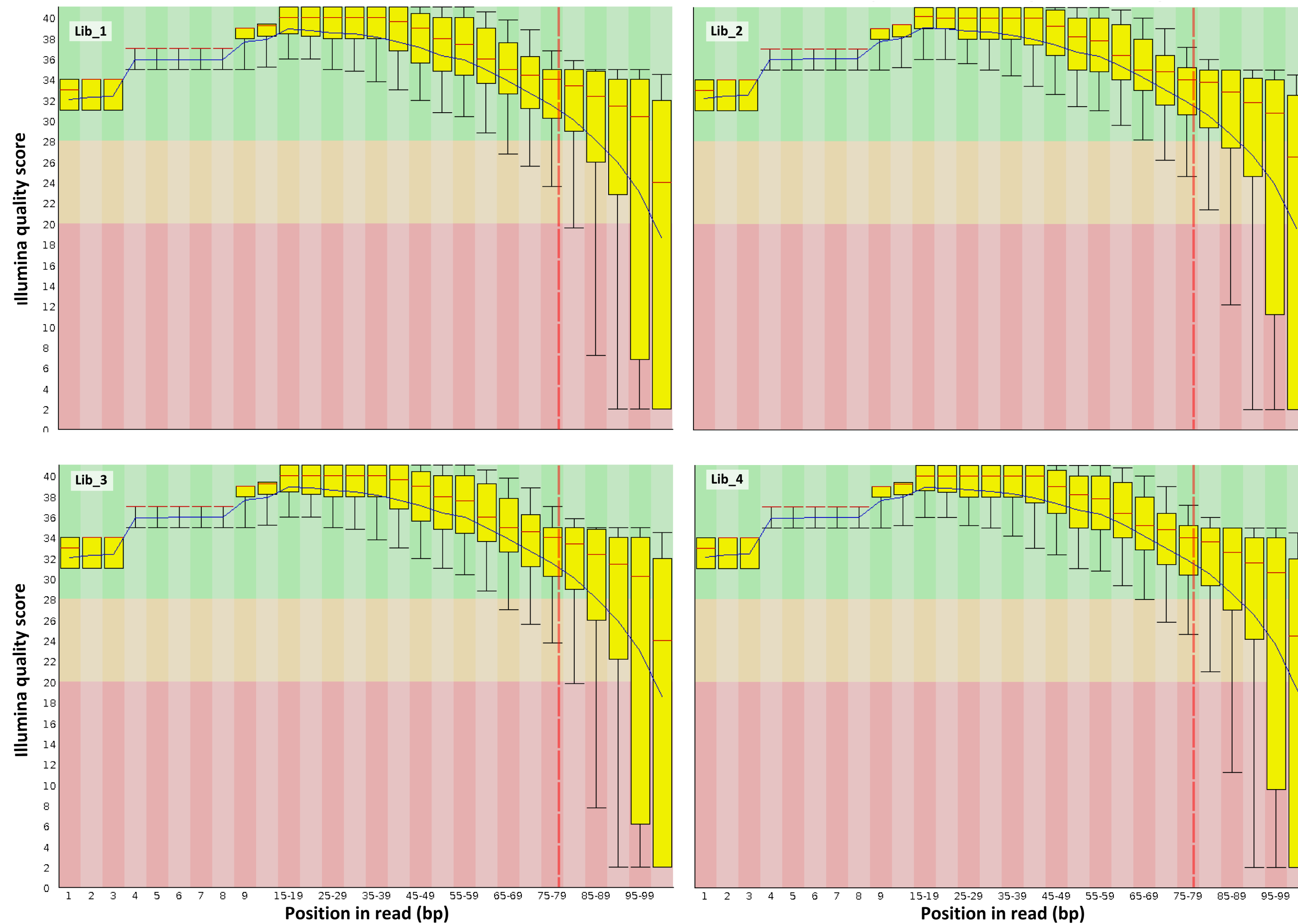


Figure 2: Box plots generated by FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc) showing per base quality scores for raw second end reads from each of the four RAD libraries. The dashed red lines indicate the length at which reads were trimmed to during processing. Quality scores are in Sanger / Illumina v1.9 encoding.

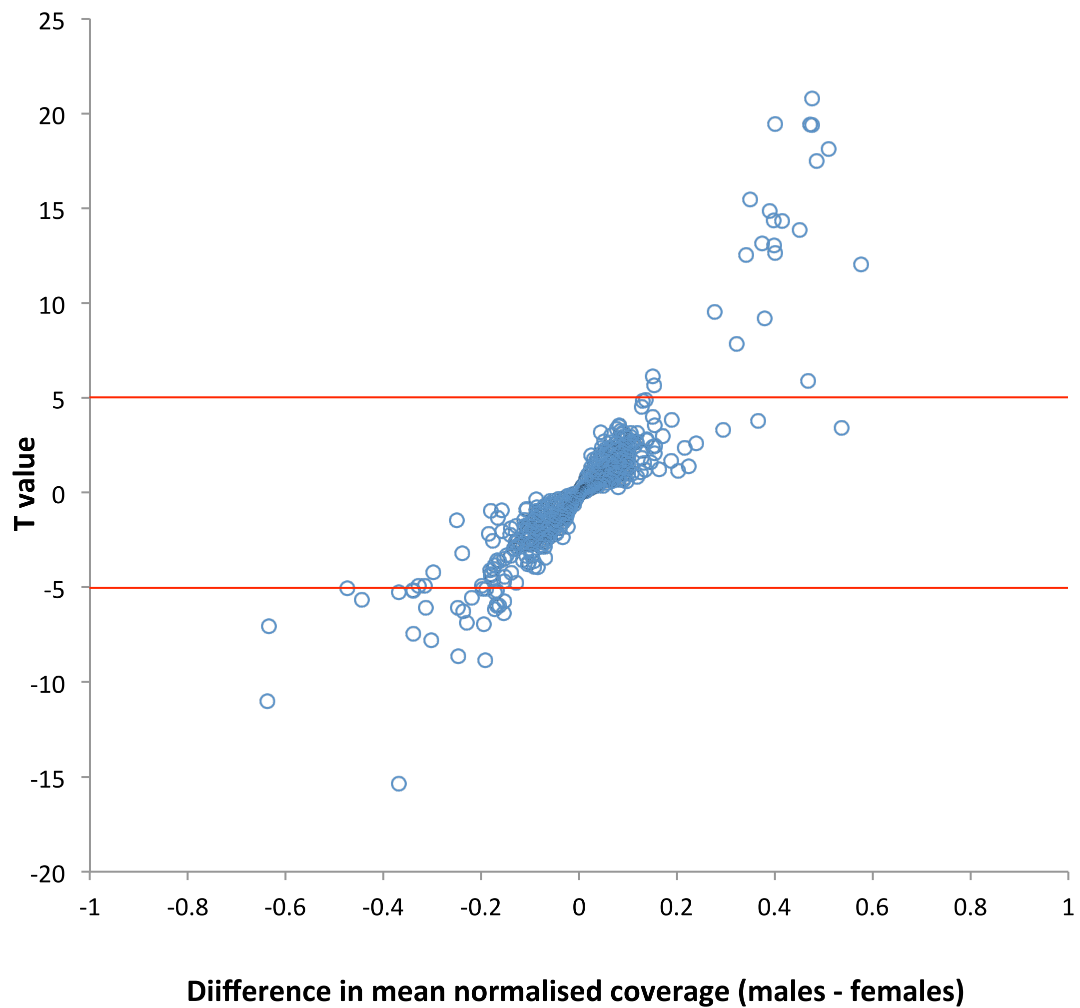


Figure 3: Male vs.female T values (calculated from normalised coverage measures) against mean difference in normalised coverage (males - females). Horizontal lines show critical values for t where alpha=0.05/1180 (Bonferoni corrected for 1180 tests). Negative t-values: normalised coverage male < female; positive t-values: male > female.

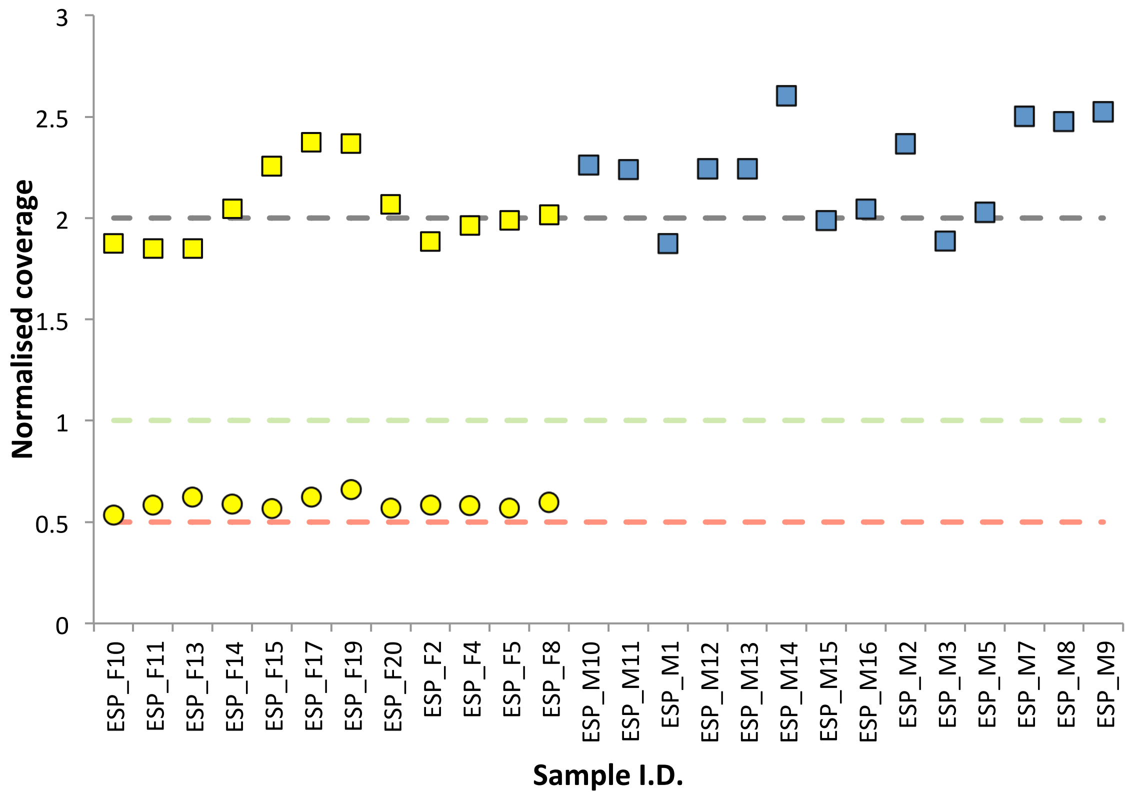


Figure 4: Normalised coverage in ESP females (yellow) and males (blue) for RAD loci #834 (circles) and #835 (squares). Dashed lines show expected coverage for hemizygous loci (red), autosomal single copy loci (green) and autosomal duplicated (once) loci (grey). #834 is orthologous to #835 but was split by Stacks due to a sex specific indel in the first end. #834 is W specific, being found only in ESP females (WZ) and never in ESP males (ZZ),

Appendix 5: Consensus sequences of sex-linked RAD loci identified in Chapter 4

#11

TGCAGGTCCTCCGCAAGACGCAGAGACTGGAGCAACTTATCAGTTTCTTCCAATCCGATCCCGAGATCCA

#75

TGCAGGCGTTTCTGGCTCTGGTTCTTTGACGACATCTTTGAGTGAGGCTTGTAAGCGATTAATAAGGTTA

#143

TGCAGGGAGTCGGCTTGCAGGACAGCCCGCTGACTTGCGCTGTTGAGCGCTGTCTGGCGCTTCCTGAGCT

#191

TGCAGGATTCCGGGAATTGGCTGCGGCGTTGCTCATAAGGCTGGGAACGTTTATGAACGTTTATGACAGG

#227

TGCAGGAGGTTCGTTATATAGAGATTTGCTAATACTTTTATGCGTTCTCTGACCTGAGGGTTAGTAGGCA

#317

TGCAGGACGGGTCATCAGGCTGGTCAGGCTCAGGCGCCAGACTAGGTAGCGGACTGTCGCACTGTTCTAT

#321

TGCAGGAACAAGTCCACCAACTGAAGCTGAAGATTGCGGAAGAAGAGTTGTTCCGTAAGAAAGCTCAAGA

#365

TGCAGGATCTCCTGAACAGCGTTCCCGTACAAACTGAAGCCATGGCCCAGATTCCGATTCCAGCCGTCGC

#368

TGCAGGTACGCGTTTTTCACAAGGTTTGTTACCAGTGCCTTAGCAAGAATGTGGTTTGGGAGTATCTCTC

#448

TGCAGGTTTATGCCAGAGCCGTTGTATTAGGCGAATAAGATACAATAGAAGAGAACAGTGCAAATAATAC

#494

TGCAGGGAAACAGGCTAGTCCACGAAGCAATAATACAACCCTCTTCGTACTATCGTTTCATCGTTGCTTA

#513

TGCAGGCATAATTGGTATCTAACTAAAAGACTGCTGGCCCCATCCTTTTTGCTTTTCATGCTTGTGTTTT

#550

TGCAGGACTAACACGCCTGTTTACTCTACGGATAATGAAGCTTTCCTTCAAGAGCACGCACTGGACAGGT

#592

TGCAGGTAAGCCTGACAGTTTCTTTCCCCTTTTTCCTTTCCTAGACATGGGTACACACACTTGGCACACG

#623

TGCAGGTTTTCTTCGGGGACTATTTAGCTCCGTTTCTACTCCAGAAGCCGGTTAAAATATGTCTACTGCT

#732

TGCAGGGCGCAAAGACACACTAAAACGCGGAACATGGTTTCTGCCAAGTATCGGCAACGGGTTTCTGTGA

#765

TGCAGGGTTCGCTAACCAGAAAACTGTTCGATACACCCAAAGTTCGTTGTATCGGACGCTTTTAGGCGCA

#779

TGCAGGGAAAAGCTGTCGTACAGCTGGGCTCTTGGTTACACCGGCTAAGGAGGCCACTGTGACGGCTGG
G

#781

TGCAGGCTAGTGGCCTTTCGTGTGAAATTTCTTAAATTGCCTAAAGGAAGGTTGTTAATTGTCCATAGAC

#835

TGCAGGAAGCATACGGAATACCAGAAGTCAAACCAGGAAACTTATTCGATAATTTAACAGCACCATGGGA

#943

TGCAGGTCCATTATCGCGGATTCTGACATGGCCTCCTCAAGTCGAAATTCAAGCATGGCTATCTCGAAGC

#991

TGCAGGAAGGTTTTCTAATGAGATTTATTGAGAAACTTTTATATTTCTATTTATTTATTTAATGCTGATC

#1021

TGCAGGTGTAGACAGCAGCGCCCCCAAACCGGAAGCAACCCCTGCTCACCAGGCGGCGAGTGCCCCGCCA

#1073

TGCAGGTAATGAAGGGATTTTTTTTCCTATTACAAACAAAGAGAAAATCGAACTAGTTTATTTTCAATCA

#1346

TGCAGGTCAAAGCTCAAGTTTTGCAACAAGTACGGGAACTCATGTTGCAATGCGATCAAACCATGAAAGC

#1369

TGCAGGTATAATTTCTGCGACGCTGCATTGGGCGTCCACCGCTTGTGCACCCAGCAAGGTGCTAGCCCAT

#1398

TGCAGGAAAAAAACCGGGGTCAATTTGAATTTTTACTTCGCTCGCATAATCAACTGTGTGTTTGAAGGAC

#1519

TGCAGGGCTGGTCTCGGTTACCGATCTATTTGGGAATATGCCTTTTCGTATCCTCCAGGTTAGACACCCG

#1524

TGCAGGAGAAACTCACGCCTCCGTAGAAAATGCTACTAGCGATGTTCTAACACGAGCATTATATCAACGA

#1546

TGCAGGTTTTTTTTTCAAAGATACGCGTATTACCGATTTTCTTTCACTTTTCGTGTTCCTGTGATTATAC

#1722

TGCAGGCTGCGAAATTTTTTCTTCTTTATTAGTTCGCTTTTATCCTTCACATCTTTTGCTTTCCTAACTA

#1780

TGCAGGGACTTAATACTCAGGTTAAAAATCATGTCATTTGGGAATATTTTCTTTCGCCAAACCCAGTTAA

#1791

TGCAGGGAGTTAAAACTTCATGCTCGGGGGATGTTGGAAAACGGGAGAGAGAGCAAAAAAAAAAAAGAC
A

#1913

TGCAGGTTTTTCTTTTTCAAAGATACGCGTATTACCGATTTTCTTTCACTTTTCGTGTTCCTGTGATTAT

#1953

TGCAGGTTTAAGCGTGTAGCAAACTTGCAAGCCAACTTTACAAATTCAAATAATGCACAACTTACGTTGC

#1981

TGCAGGCGAAAAATAAAAGTAAAAAAATCCCTCCCACAGAAAAAAAGGAATGTTAGTTTTATGTATCTTT

#2099

TGCAGGTGGCATTGGCCATCAATTTAGGTAATTACTAGCTTTCAAAGTAATTCAGGTATATATGACCAAT

#2110

TGCAGGGATCTTACTCCTTTTCAAAATGAGATTTTTTGTTTTATTTGTATTTTATTGTATTGTTGGAGCG

#2137

TGCAGGAACGTCGCCGAACGTTAGCGGGTATTTTTTCGTCATCTGTTGGGCAATGTAAGCCGCGCTTGGG

#2139

TGCAGGGGTTTTGTAAGGTAATAAAAGTAATTGACGATTATGACGGGTAGCAATAAGCAATTCAGTTTGT

#2141

TGCAGGGGAGAAATAAAAAACGCAGAGGTAACATGGTACCAGTTTTAAAGCGCTTGCGATGGGTGTCACC

#2215

TGCAGGCTGCGCCAATTTGTTTTGGTTATTTAGCTTACAACAAACAGTTAGAGATATAATATGTATTTAT

#2223

TGCAGGCCGCGCTAGTCTACAGTCCCAAAACCTCTCCCTTCGATTTATGGTACTCGAAGACACACGTGTA

#4909

TGCAGGATGTCACGGAGGACTTCGATGGTTACGAAATGAAGGCCAGGAGCAAATGCCCAGACTCAAGTTA

#11478

TGCAGGCCCATAGTGACATTCCCGCAAAACACCTTGGTGGAACGGGCGAAACAGGACGAGAAGGACTTTC

#11483

TGCAGGGTGGTTTATGGTTATGTCCCGAGAACAGGCCTCATTTATCGTATATAGGTAAAGCACCGTACTC

#11488

TGCAGGGCCCAAAAGTAAGCCATTGTGGCGACTACCCATATGAGGCGCATGCACTGCCGCGTTCTCGTAG

#11495

TGCAGGAATAAAAGCTAACCAAAGCCCTCGACACGCCAGTTGGATAGCCCTACTGTATAATGGTATCACT

#11499

TGCAGGAAAGATTAAATTAACTGTTTGCGCACGTTCTTACTATAATATTCTTTTAGAAAAACCAACAAAC

#11536

TGCAGGGACTTAATACTCAGGTTAAAAAGTCATGTCATTTGGGAATATTTTCTTTCTCCAAACCCAGTTA

#11561

TGCAGGTCTCCGAATATATGCCCGCCTCGAAGAAATCTTCCCCACCTTTATATCCCCGTGATGTTGGCAT

#11562

TGCAGGAGGGCACCTCGGTCGATCGAAGACGCTCGAGAAAGTACGTGCACGATACTGGTGGCCGACGTT
A

#2117

TGCAGGATTGGCGTGTCAGGCGTTTTAACCGACCCAACGAGCAAAGTGCACTGAACAATTTAACCTTTAT

#1920

TGCAGGGGGTGACGACGAGCACGGGCCTGTAGTCTTCCATCTGGAGGGCGAACTCATTACGAATGTGTAG

#2086

TGCAGGAACTGGAGCCGTTTTGGGCGACGGTGCACCTAAGTGGGATATAGGGAAATCACATGCTAAACG
G

#976

TGCAGGTGATCAAAACAGAGTGCGTTTTGCTCCTCCTCAGGAGAACTCGACACGAACGGCATAGCTAGGC

#246

TGCAGGAACGCAAATCAGTTAATGCAACTTTCCAATGCCACTTTCCCTCCTCAGTTTAAAAATGTTTTAA

#984

TGCAGGAGGCTCAGCTGGTTTACGAGAGAGAGGATCAGCATCATTGTGGGCCATTCCTATGGGTAATAGT

#775

TGCAGGTTCACAATTCCTTGGAACACGTTCCATGATAACGAAGATGGCGCAACGCGAAATTTGAAGTTGT

#1562

TGCAGGAAGGTTTTCTAATGAGATTTATTGAGAAACTTTCTATTTTTATTTATTTATTTAATGCTGATCA

#519

TGCAGGCTGCGCCAAAATGTAAGAATGTGTTCTTACATTTAGAAATGGTTCTTTTTAGGCGTAGGCGTTA

#157

TGCAGGTTTGTTGTGCTCACTTTGACATTTAGGCATTTCTTGCACCTGTGAAAATTCATCCATTAGCGAA

#1073

TGCAGGTAATGAAGGGATTTTTTTTCCTATTACAAACAAAGAGAAAATCGAACTAGTTTATTTTCAATCA

#468

TGCAGGCTGCGAAATTTTTTTCTGCTTTCTTTGTTTGCTTTATCCTTTACATCTTTTGCTTTCCTAACCA

#2104

TGCAGGAGGTTCGTTATATAGAGATTTGCTGTAATACTTTTATGCGTTCTCTGACCTGAGGGTTAGTAGG

#1447

TGCAGGGCTGGTCTCGGTTACCGATCTATTTGGGAATATGCCTTTTCGTATTCCTCCAGGTTAGACACCC

#842

TGCAGGGCTTGGAGAGCGTGGCCTGAATTCGGGTGCTTCTGCACGAAAGCTACTACCAGTTCTTGAAAAG

#1279

TGCAGGCGAATATAAATTTTCAGAACTATTCCAAGAATAACTCCGCATTCAAGCATTTTGACATTTTTAG

#11485

TGCAGGTACTTCTTCAGAGATTCTAGAAGGTTCATTGCAGTAAGTATGTTCACTTCACTTGACTGCAGAA

#689

TGCAGGTGACAGACGGCTTTCATCTTTAGGAAAAACTCACTTTAGGAGTGAAAAAGGCTTTTAGGGACTT