INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Evolution of Polyploidy on Two Continents: Phylogenetic Resolution and Implications

Approved by
Dissertation Committee:

David M. Hillis, Co-Supervisor

James J. Bull, Co-Supervisor

Michael J. Ryan

Paul A. Krieg

Donald Morizot

J. Dale Roberts

Evolution of Polyploidy on Two Continents: Phylogenetic Resolution and Implications

by

Barbara Kim Mable, BSc, MSc

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 1996

Acknowledgements

There are many people who have assisted me in various ways throughout the course of my dissertation studies. In Texas, my general thanks go to the people in the Hillis/Bull and the Ryan labs, who have provided constant support, advice, and stimulation. Although each member of these labs provided moral support or assistance at some stage, there are several people who deserve particular mention. Cathy Marler provided valuable advice in the beginning stages of my research and showed me around the local frog ponds; John Huelsenbeck showed me the wonders of simulation, let me parasitize his programs, and taught me why likelihood is superior to any other method known to man; Tod Reeder introduced me to PCR and sequencing; Keith Crandall was a constant source of advice and generously sent me to collect crayfish in Australia so that I could follow up on loose ends of my frog project; Marty Badgett provided continual assistance in the lab beyond expectations of a normal laboratory technician. I would also like to thank Aaron Zorn (not from the Hillis/Bull or Ryan labs), who nurtured me through cloning and technical failures, taught me how research was done in a real molecular lab, and provided continual entertainment during the most hopeless periods of my laboratory research.

During the six months I spent in Australia, although everyone in Craig Moritz' lab was extremely helpful and friendly and made my stay there productive and enjoyable, there are several people who deserve particular thanks. First of all I

cannot thank Craig Moritz enough for his generosity, advice, and support. He gave me free reign to do whatever lab work I wanted, arranged for me to lecture a course that he was involved with, always treated me more as a colleague than a student, and continued to provide advice after I returned to Texas. Christina James made my initiation into the lab easy and provided technical assistance and lab failure sympathy; Leo Joseph provided advice and entertainment; Andrew Hugall and Anita Heideman were generous with their time and expertise. There are also some people from outside of the Moritz lab who deserve special acknowledgement. Mike Mahony from the University of Newcastle and Steve Donnellan from the South Australian Museum generously provided me access to their tissue collection and their data; Jan Birrell extracted DNA for the original samples that I used for the Australian frog project.

I would very much like to thank the members of my committee, who have been very understanding about the changing nature of my thesis plan. Mike Ryan initially supported me as a graduate student and was understanding when my research plan deviated from the focus of research in his laboratory; Paul Krieg provided continual technical advice and accepted me into the social world of the second floor; Dale Roberts, at the University of Western Australia, provided extremely useful advice on Australian frog biology, provided hours of interesting discussions in his laboratory, and collected specimens for me; Don Morizot, at Science Park, provided realistic criticism and suggestions. I would especially like to thank my advisors, David Hillis and Jim Bull. David Hillis accepted me into his lab even though I came in through the back door, provided the opportunity for me to spend six months in Australia on an informal exchange with Craig Moritz' lab at

the University of Queensland, gave me the opportunity to work on a major publication, and was always generous with resources; Jim Bull was a constant source of ideas, advice and encouragement. I would also like to thank Jim Bogart, my master's advisor from the University of Guelph, who continued to provide advice and suggestions throughout my graduate career, and included me in a collecting trip throughout the northcentral states and Canada.

Finally, the people who have provided moral support over the years (often at the Crown and Anchor) were essential to the completion of my dissertation project but are too numerous to mention without risk of serious omissions. However, Dan Haydon deserves particular praise for not only providing me with computer programs to make my life easier, providing a haven for me to write up in, and doing the boring task of copy editing, but especially for bearing the brunt of the fallout from the stresses of the last few years of my graduate career.

Evolution of Polyploidy on Two Continents: Phylogenetic Resolution and Implications

Publication No._____

Barbara Kim Mable, PhD

The University of Texas at Austin, 1996

Supervisors: David M. Hillis and James J. Bull

The North American grey treefrogs are a cryptic diploid-tetraploid complex that has been extensively studied but the nature of origins of the tetraploids is still not certain and there is a question of whether the currently applied names, *Hyla chrysoscelis* (2n=24) and *Hyla versicolor* (4n=48), adequately describe the number of evolutionary lineages in the group. The Australian burrowing frog genus *Neobatrachus*, which has not been studied as extensively, consists of six diploid species (2n=24) and four nominal tetraploid species (4n=48). Sequences from mitochondrial genes were used to reconstruct historical relationships of populations (grey treefrog complex) or nominal species (Australian burrowing frog complex) within each of these groups to identify the probable number of origins of tetraploidy within each group and to determine if a diploid progenitor could be identified among exant congenerics. In the North American group, mitochondrial data

suggested that tetraploids from northern populations are not closely related to the diploids with which they are currently sympatric and a potential diploid ancestor was not identified. Conclusions based on allozyme evidence did not generally concur with mitochondrial data. This discrepancy could provide evidence that tetraploids arose through hybridization but could also reflect different evolutionary forces occurring within the same populations, or complications associated with differential expression of gene products in tetraploids. However, conclusions based on all data currently available suggested that the number of lineages present in the complex has been underestimated. In the Australian group, mitochondrial data suggested that there have been at least two (and possibly three) origins of tetraploids. A potential diploid relative was identified for only one of the tetraploid lineages; tetraploids in the other lineages may have replaced their diploid progenitors. Sequences from intronic regions of presumed single-copy nuclear genes were also obtained but their utility for phylogenetic comparisons was limited by problems with pseudogenes, possible heterogeneity of multiple gene copies, lack of introns in the species studied, and most importantly, lack of nucleotide variation within the taxa of interest. Implications for species concepts and the utility of the particular molecular markers chosen for comparison in the study are discussed.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Table Of Contents

INTRODUCTION

Background

Species Concepts and Speciation Processes

In the study of evolution, it is important to identify genetic changes within species as well as those that could produce new species. However, while genetic changes within species can be studied in a reasonable timeframe, the evolution of reproductive isolation generally takes much longer (Coyne and Orr 1989). Most models of speciation have been based on biogeographical evidence for allopatric speciation (e.g., Mayr 1942) for which it is difficult to establish how many genes might be involved in generating isolating mechanisms (reviewed by Coyne 1992). Interpretation of factors that result in speciation may also vary with the definition of species that is used. For example, Mayr's (1942) biological species concept is that "species are groups of interbreeding natural populations that are reproductively isolated from other such groups". This definition predicts that once we understand the origin of reproductive isolating factors we understand the origin of species (Coyne 1992). However, there are problems with this kind of interpretation when it is applied to populations that have diverged in allopatry, because intrinsic barriers to reproduction do not need to be invoked to maintain isolation and homogamy in these cases (Littlejohn 1981). According to the evolutionary species concept a species is defined as "a single lineage of ancestral-dependent populations which maintains its identity from other such lineages and which has its own evolutionary tendencies and historical fate" (Wiley 1978). This definition ascribes less importance to isolating factors in the identification of species units, especially for allopatrically derived populations.

Sympatric speciation models, in contrast, require independent evolution of gene pools under conditions of constant potential contact. Therefore, by all definitions, functional reproductive isolation between diverging species would be essential to the process (unless the new species replaces its progenitor). However, the likelihood of divergence in sympatry has been debated and usually past geographic separation has been invoked as an untestable alternative (Coyne 1992). If speciation has occurred in sympatry, it seems likely that rapid or instantaneous development of isolating mechanisms would greatly increase the possibility that the diverging species would not interact.

Polyploidy and Speciation

In plants, the major mechanism for sympatric speciation is through polyploidization. Historically, this type of speciation has not been recognized as a major factor in animal evolution because of the suggestion that imbalances in sex determination (Muller 1925) or in dosage compensation mechanisms (Orr 1990) would make the possibility of polyploidy improbable. However, Ohno *et al.* (1968) have proposed that gene duplication has been a major factor in the evolution of vertebrates from fish through mammals. They suggest that both tandem duplication of single genes and duplication of entire genomes has occurred throughout vertebrate history, and that the larger amount of DNA in mammals compared to the "lower" groups may be explained on the basis of duplication of the genome of a crossopterygian ancestor. This is suggested both by the variable amount of DNA in different species and by the many examples of nonlinked duplicate genes for various proteins (Kobel and Pasquier 1986). The existence of polyploid species in such

diverse groups as insects (Bogart 1967), crustaceans (Salemma 1984), fish (Ohno *et al.* 1968; May *et al.* 1979), reptiles (Bogart 1980), and amphibians (M.L. Beçak *et al.* 1970b; Bogart and Tandy 1976; Bogart 1980; Kawamura 1984; Kuramoto 1990) emphasizes the possible importance of speciation through gene duplication of entire genomes (i.e., polyploidy) in animals. Since some of the polyploid species putatively can be traced to a relatively recent diploid ancestor (or sibling species), they provide a model to study the origin and maintainance of duplicated genomes and their relationship to speciation patterns. In particular, identification of the prezygotic (behavioural or ecological) and postzygotic (genetic incompatibility and hybrid sterility) factors that have allowed the derived species to coexist with the putative ancestor would help to determine whether genome duplication can lead to automatic speciation in sympatry or whether further refinement and physical isolation of the daughter and parent genomes are required before independent evolution can be achieved.

While many of the polyploid vertebrate species have adopted alternative modes of reproduction (parthenogenesis, hybridogenesis and gynogenesis), there are many groups which display normal bisexual reproduction (e.g., Bogart 1980; Schultz 1980). These include fish in the families Salmonidae and Catostomidae (Ohno *et al.* 1968), and anurans in the families Pipidae, Leptodactylidae, Hylidae, Bufonidae, Phyllomedusidae, Ranidae (reviewed by Bogart 1980; Kawamura 1984), Mybatrachidae (Mahony and Robinson 1980; Mahony 1986; Mahony and Roberts 1986), and Microhylidae (Kuramoto and Allison 1989). Although these groups diverged at different times through geologic history, the general consensus of researchers working on these groups is that polyploids should be generalistic

pioneers that are able to exploit new habitats under times of environmental upheaval (such as glaciation periods) due to the increased flexibility provided by extra gene copies (e.g., Salemma 1984). Although precise origins have not been established for any group, many have been tied to a glacial or postglacial period.

Polyploidy and Gene Regulation

Many of the more ancient polyploids in these groups (salmonids, catostomids, pipids) have undergone a functional rediploidization of their genomes followed by subsequent silencing of duplicate gene expression. This loss is thought to be due to the costs of maintaining an extra genetic load, such as the necessity to become tolerant to a larger number of self-antigens in the immune system (Kobel and Pasquier 1986) and random loss of duplicate gene expression through genetic drift effects (Bailey *et al.* 1978). However, even in the older polyploids (ca. 30 million years in pipids; 50 million years in salmonids and catostomids) there is evidence of retention of duplicate genes. *Xenopus laevis* (family Pipidae) displays disomic inheritance at most loci but displays several genes that have retained duplicate function (Kobel and Pasquier 1986). Polyploid catostomid and salmonid species show limited residual tetrasomic inheritance but maintain duplicate gene expression at approximately 50% of their loci and retain some multivalents at meiosis (Bailey *et al.* 1978). Allendorf (1978) has suggested that the same evolutionary pressure that results in the accumulation and fixation of nonfunctional ("null") mutations may also lead to structural divergence and the fixation of distinct alleles at duplicate loci. Duplicate gene copies may ultimately diverge enough that they experience differential expression due to regulatory differences that often appear as tissue- or developmental

stage-specific expression (Ferris and Whitt 1977a, 1977b; Danzmann and Bogart 1983; Kobel and Pasquier 1986). In *X. laevis*, for instance, malate dehydrogenase shows differential expression in various tissues (Kobel and Pasquier 1986); two forms of the gene c-myc are controlled by different promoters and are expressed at different stages in development (Principaud and Spohr 1991; Vriz *et al.* 1989); and two insulin genes are present in duplicate form (Schuldiner *et al.* 1991). The evolutionary forces that result in loss of duplicate gene expression and/or divergence of duplicated loci are not known. The regulation of duplicated genes is currently an area of strong interest in developmental and cellular biology that could be extended by knowledge of the processes that control regulation of genes in polyploids.

Gene regulation has been suggested as an important factor (or the most important factor) in organismal evolution (Britten and Davidson 1969; Parker *et al.* 1984; Cavener 1989). An important question arising from this idea is whether the regulatory changes that make two groups distinct functionally and morphologically represent a magnification of the polymorphisms found within species or if additional elements are necessary for the process of speciation and subsequent genetic isolation (Paigen 1989). Tetraploidization provides a plethora of evolutionary possibilities and many duplicated genes undergo a progressive restriction in patterns of expression as speciation proceeds that tend to be tissue rather than enzyme specific (Paigen 1989). Although polyploid animals usually have fewer, larger cells than their diploid counterparts, they also tend to regulate cell size to less than twice the volume to preserve normal body size (Bachmann and Bogart 1975). Whether this is an automatic result of polyploidization or is a result of postspeciation regulatory changes is not known. Comparing the degree of regulatory and structural divergence between

polyploid species and their diploid ancestors could contribute to knowledge not only about control of gene expression in polyploids but of the general processes governing the regulation of duplicated gene products in diploid organisms. However, difficulties in absolutely identifying "null" alleles (i.e., nonfunctional duplicate gene copies) and the obscure relationships of the more ancient polyploids to diploid ancestors has made extrapolation of these comparisons to speciation processes in groups that have already undergone rediploidization unfeasible.

Polyploidy and Reproductive Isolation

Study of a more recent polyploid that has not yet undergone extensive functional rediploidization could provide more information on factors contributing to genetic isolation and on the nature of regulatory changes that may occur immediately following the speciation process. Among anurans, putatively recently evolved bisexual diploid-polyploid pairs have been described in the families Leptodactylidae, Hylidae, Ranidae, Bufonidae (Bogart 1980), Myobatrachidae (Mahony and Robinson 1980; Mahony 1986), and Microhylidae (Kuramoto and Allison 1989). Among these groups, the North American hylid representatives (*Hyla versicolor* complex) and one of the South American leptodactylid complexes (*Odontophrynus americanus* complex) have been studied the most extensively. In both of these systems, reproductive isolating mechanisms of the polyploid species from its diploid counterpart have been characterized. The tetraploids produce mating calls that are distinctive from the diploids and that can be used by females to discriminate species (Littlejohn 1960; Gerhardt 1982; Doherty and Gerhardt 1984; Klump and Gerhardt 1987; Gerhardt and Doherty 1988). The ratio of the pulse rates of the tetraploids and

diploids corresponds closely with the difference in cell size between them (Bachmann and Bogart 1975) and tetraploids in both groups produce calls with reduced pulse rates compared to closely related diploids at the same temperature (Gerhardt 1978; Bogart and Wasserman 1972; Bogart 1980). This evidence has been used to suggest that acoustic isolation could have been automatic with the increase in genome size if there was a corresponding change in the female auditory system (Bogart and Wasserman 1972; Bogart 1980). Naturally occurring triploids in the grey treefrog complex have been found very rarely in the field (Gerhardt *et al.* 1994) although artificially produced hybrids are quite vigorous (e.g., Mecham 1965; Mable 1989). The rarity of hybrids in the field could be because mating calls provide an effective barrier to reproduction.

Triploid hybrids resulting from artificial hybridization of diploids and tetraploids within the grey treefrog complex are sterile, resulting in postzygotic isolation if accidental mismatings do occur (Johnson 1966; Bogart 1980; Mable 1989). Hybrids involving tetraploid females and diploid males are easily produced artificially and these crosses often produce more robust offspring than control crosses but sexual development does not appear to be normal and reciprocal crosses (i.e., involving diploid females and tetraploid males) do not develop past the gastrula stage (*Odontophrynus*: W. Beçak *et al.* 1968; *Hyla*: Mable 1989; Mable and Bogart 1995). An interesting paradox is that successful backcrosses may be obtained in crosses involving tetraploid *H. versicolor* females and diploid males that are outside of the chromosomally characterized (Wiley 1982) *versicolor* species group (that includes *H. chrysoscelis*), while crosses within this group are less successful (Mable 1989; Mable and Bogart 1995). Similarly, crosses involving female *O. americanus* tetraploids and

O. cultripes males (the most distantly related of the diploids within the genus according to Bogart and Wasserman 1972) resulted in hybrids with better gonadal development than crosses involving diploid and tetraploid populations of *O. americanus* (M.L. Beçak *et al.* 1970a) and produced haploid, diploid and triploid gametes (M.L. Beçak and Beçak 1970). This suggests that meiotic failure due solely to triploidy may not be sufficient to explain the complete hybrid sterility observed in these two cryptic diploid-tetraploid species pairs. These observations suggest that postzygotic isolation may have been the result of postspeciation divergence rather than a direct consequence of genome duplication; premating isolation may have been sufficient to initially reduce gene flow between species. The strong parallels in isolating mechanisms between the North American polyploid complex and the South American complex (Bogart 1967) suggest a pattern that may explain how polyploidization can result in "successful" instant speciation of ecologically similar species under sympatric conditions that may be reinforced by subsequent changes occurring with divergence. This pattern may not be a universal feature of anuran polyploids (e.g., diploid and tetraploid *Phyllomedusa* produce calls with similar pulse rates: Haddad *et al.* 1994), but provides an interesting hypothesis to examine.

Polyploidy and Ecological Isolation

In order for a new polyploid species to arise sympatrically with its diploid progenitor it would not only have to have effective isolation mechanisms but would probably have to experience some sort of competitive advantage to exist in the midst of a very similar species using the same ecological resources. *H. versicolor* and *H. chrysoscelis* are morphologically indistinguishable and can be found breeding in the

same ponds on the same nights in a broad range of sympatry ranging from eastern Manitoba to central Texas (Bogart, personal communication; Ralin and Selander 1979; Gerhardt *et al.* 1994; Ptacek *et al.* 1994). They may have some differences in tolerance to dessication (Ralin 1981), and may have slight differences in microhabitat usage (Ralin 1968) where they occur together but are in most respects (e.g., metabolic rate: Kamel *et al.* 1985; body size: Ralin and Rogers 1979; Bogart 1980) ecologically equivalent. The ecology of *Odontophrynus americanus* has not been studied as extensively, but morphological identity with a diploid relative has been described, although diploids and tetraploids appear to have been collected from different areas (M.L. Beçak *et al.* 1970a). Acoustical but not morphological distinction from the diploid with which they are in closest association is also found in other tetraploid anuran groups (e.g., *Phyllomedusa burmeisteri* complex: Batistic *et al.* 1975). These trends suggest that tetraploids may speciate without a significant change in ecological or physiological attributes and it therefore seems likely that competition between the newly arising species and its diploid progenitor could be a significant detriment unless polyploidy conferred some initial advantage. Salemma (1984) suggested that polyploids arising in sympatry would be most vulnerable in the initial stages of their history and therefore must either fill an ecological niche that is different than their progenitors or have an improved capacity for colonization and reproduction in order to disperse and occupy new habitats and must have sufficient genetic variation in their populations for further evolutionary change (Salemma 1984).

Consequences of Polyploidy

Salemma (1984) suggested that polyploid marine amphipods (genus *Pontoporeia*) represent glacial relicts and arose in a time of environmental upheaval when they experienced an advantage compared to their diploid counterparts. Blair (1958) considered the *H. versicolor-H. chrysoscelis* species pair to be an example of speciation resulting from climatic events of the Pleistocene (Bogart 1980) and the period proceeding the Wisconsin glaciation (ca. 150,000 years ago) has been suggested as the most probable for explaining current distribution patterns (Ralin *et al.* 1983; Little *et al.* 1989). These predictions are consistent with the view that polyploids represent "generalistic pioneers" who may take advantage of changes in the distribution of their ancestral species caused by drastic environmental crisis that allow them to colonize empty niches (Salemma 1984). The reasons invoked for this advantage under unstable conditions have generally been linked to the increased capacity for heterozygosity, and resulting potential for increased genetic plasticity associated with total genome duplication. Ohno *et al.* (1968) suggested that a single gene duplication has the capacity to confer a heterozygous advantage to every member of a species; duplication of the entire genome would vastly increase this capacity. Bogart and Wasserman (1972) suggested that diploids may be relatively more successful in more stable regions or when conditions are stable for a long period of time, while tetraploids may be able to exploit a larger range of environmental conditions than the potentially less variable diploid species under more unstable conditions. Genetic load effects in tetraploids could be responsible for reducing competitive advantage under stable conditions, while increased capacity for the

evolution of novel gene functions could increase this advantage under more variable conditions.

This potential trade-off could explain why most older tetraploids eventually reduce their overall gene expression to a diploid level but retain some degree of duplicate expression beyond what is usual for their diploid counterparts. Genetic load effects could also result in strong selection for reduction of redundant gene expression in newly arising polyploid species prior to a return to disomic inheritance. Extensive gene divergence and loss of expression are not thought to be possible until functional diploidization of the genome has occurred (Bailey *et al.* 1978), but changes in regulatory control may be apparent even within younger tetraploids which still exhibit duplicate gene expression and multivalent chromosome associations. Tetraploids in the *O. americanus* complex of South America still show multivalents at meiosis (M.L. Beçak *et al.* 1966; 1967) and have twice the amount of DNA as their diploid counterparts (W. Beçak *et al.* 1967). However, while the tetraploid exhibits duplicate expression at certain loci (e.g., albumin), at others (e.g., LDH) it shows reduced expression as judged both by allozyme patterns and enzyme activity measurements (W. Beçak 1969). RNA and hemoglobin content were also found to be independent of ploidy in *Odontophrynus*, as diploids and tetraploids showed similar values even though the DNA content is doubled (W. Beçak and Goissis 1971). These differences in regulation were attributed to super-repression of structural genes by the increased number of regulatory loci in the tetraploid, asynchrony in expression of duplicated genomes, or a reduced transcription rate of the duplicated genome to compensate for the excess genetic material (W. Beçak and Goissis 1971). *H. versicolor* has been shown to exhibit disomic, tetrasomic and

intermediate patterns of inheritance that varies among tissues, enzyme systems and individuals (Danzmann and Bogart 1982a,b, 1983; Marsden *et al.* 1987) and is therefore thought to be in an initial stage of rediploidization. It has less than twice the amount of DNA as *H. chrysoscelis*, reflecting an approximately 10% loss, and the ratio in content of other cytological parameters is further reduced (DNA 1.9; nuclear histones 1.7, nuclear RNA 1.6, total nuclear protein 1.5; nuclear sizes 1.4; Bachmann and Bogart 1978). Deviations from expected inheritance patterns in artificial hybrids between *H. versicolor* females and diploid males also indirectly suggest that the tetraploid genome may have more extensive control over gene expression and may be able to preferentially express specific gene combinations (Mable 1989; Mable and Bogart 1995).

Origins of Tetraploids

Mechanism and Number of Origins

The formation of multivalents in meiosis has been used as evidence that tetraploids in these relatively young groups have arisen through autopolyploidy (e.g., Bogart 1980), in which rediploidization would not likely occur until sufficient divergence of chromosome copies enables recognition as distinctive entities. However, it is difficult to imagine why structural divergence of chromosomes would occur in pairs if associations are completely multivalent. Consequently, most older polyploids that have rediploidized are thought to be of hybrid (allopolyploid) origin (e.g., Bogart 1980), with reduction or inactivation of expression either in one of each pair of parental chromosomal types or in a complete set from one of the parents. If tetraploids arose through hybridization of previously isolated diploid populations or

distinct species, deviations from full duplicate gene expression in new tetraploids could result from regulatory differences between the parental genomes. Interpretation of gene expression patterns would be very different for these two modes of speciation but evidence has not conclusively distinguished the alternative modes of speciation in most of these younger tetraploids. There is recent evidence that polyploids in the South African frog genus *Tomopterna* (formerly *Pyxicephalus*, as described in Bogart 1980) have arisen through allopolyploidy but this would be the first evidence that a "young" anuran tetraploid (i.e., still exhibiting duplicate expression at most loci) has arisen through hybridization (Channing and Bogart, unpublished). Therefore, before conclusions can be drawn on how loss of gene expression might evolve in polyploid lineages, more information is required on how the species actually formed.

Interpretation of reproductive isolating mechanisms and subsequent conclusions on whether they arose as an automatic consequence of polyploidy or developed through later selection for character displacement would differ depending on how many tetraploid lineages actually exist in these cryptic complexes. Few of these groups have been studied extensively enough to establish whether polyploidy was a rare event that occurred once or a few times in the evolutionary history of a particular group or whether polyploidy has been a common mode of speciation that has occurred multiple times or may still be occurring in extant populations.

North American Grey Treefrogs

The most extensive body of literature addressing questions of origins of polyploidy in anurans is on relationships among populations of the North American

grey treefrogs, but the evolutionary origins of the tetraploids have still not been elucidated satisfactorily. Evidence from allozymes (Ralin 1977; 1978; Ralin *et al.* 1983), immunological distance data (Maxson *et al.* 1977), and chromosomal polymorphisms (Wiley 1983; Wiley *et al.* 1989) combined with characterization of mating calls (Gerhardt 1974) has been used to suggest that the diploid "species" in this group, *H. chrysoscelis*, may represent two distinct lineages that were geographically isolated during the Illinoian glaciation (ca. 570,000 years ago; Ralin and Selander 1979) and are currently divided by the range of the tetraploids (Bogart and Wasserman 1972; Maxson *et al.* 1977; Ralin and Selander 1979; Ralin *et al.* 1983; Ptacek *et al.* 1994). Authors have not agreed on whether the diploid lineages deserve elevation to species status, and predictions on origins of tetraploids (sometimes based on different interpretations of the same data set) have ranged from 1) single autopolyploid origin from a single genetically intermediate diploid population (Ralin *et al.* 1983) with subsequent parallel electromorph variation of diploids and tetraploids living in similar environments resulting in similar allozyme frequencies (Romano *et al.* 1987); 2) multiple autopolyploid origins from diploid hybrids in a narrow area of contact of eastern and western diploid "species" (Romano and Vaughn 1986); 3) unspecified number of allopolyploid origins through hybridization between eastern and western "species" of diploids (Maxson *et al.* 1977); and 4) allopolyploid or autopolyploid origin through an intermediate triploid stage (Bogart and Wasserman 1972). The currently applied names therefore probably define classes (diploids vs tetraploids) rather than actual species.

The strongest evidence to date comes from mitochondrial data which supports the separation of diploids into eastern and western lineages and predicts that the

tetraploids have arisen multiple times, possibly from different female ancestors (Ptacek *et al.* 1994). If this is true, characterization of molecular markers in the two diploid lineages ("eastern" and "western") could be used to predict whether the tetraploids, grouped under the name *H. versicolor*, have arisen independently from one or the other (autopolyploid) or through a hybridization event between the two diploid taxa (allopolyploidy). However, a definite diploid ancestor has not been identified and most previous studies have ignored populations of tetraploids and diploids from the northern extremes of the ranges. Additional mitochondrial DNA data from these populations and characterization of nuclear gene markers is required before an overall understanding of evolution in this group can be reached. Characterization of the historical relationships within this group is essential to the interpretation of isolating mechanisms and the evolution of genetic regulatory mechanisms controlling rediploidization. Definition of population interactions would also help to superimpose a biogeographical component on the establishment of feasible hypotheses on how and why tetraploids have arisen within this group.

One reason that the evolutionary history of this group has not been resolved after nearly 30 years of study is because of the almost complete morphological crypticity of the potentially independently evolving lineages and the resulting difficulty of field identification. Diploids and tetraploids can be distinguished by mating call variation (e.g., Johnson 1966), nucleolar number (Cash and Bogart 1978), erythrocyte size (Mable 1989; Matson 1990), flow cytometry (Krishnan 1975; Gerhardt *et al.* 1994), and to some extent intensity of staining patterns in allozyme bands (Danzmann and Bogart 1982a). However, few of these characters are useful

for museum specimens and absolute confirmation of chromosome number requires sacrificing live animals.

Australian Burrowing Frogs

The Australian genus *Neobatrachus* (Family Myobatrachidae) has not been studied extensively as the *Hyla* complex but represents a diploid-polyploid system with an additional level of complexity compared to the groups discussed above. The genus currently consists of six diploid and four tetraploid species (Mahony and Robinson 1980; Mahony 1986; Mahony and Roberts 1986; Roberts *et al.* 1991; Roberts 1978). This is one of the few anuran species groups that is known to have more than one morphologically and/or behaviourally recognizable species of tetraploid within the same complex. The potential exists to examine not only the factors contributing to and the consequences of sympatric speciation of polyploid lineages, but possibly to examine the divergence of tetraploids from other tetraploids under allopatric conditions as well. Diploids and tetraploids are found across the arid zones of Australia, with the highest number of species in the southwestern part of Western Australia. Many of the species are found sympatrically, hybridization occurs in some but not all regions of overlap, and viable hybrid offspring appear to be produced in at least some cases (Main *et al.* 1959; Mahony 1980). However, it is not known exactly which species are involved and the degree of backcrossing in these hybrid areas. Roberts (unpublished) has characterized the mating calls of all ten species but did not have a phylogenetic framework in which to examine patterns of relationships among the diploids and tetraploids in the group. However, most species have recognizably different calls. The diploids are well characterized by

morphological and behavioral differences and can be distinguished by fixed differences at allozyme loci (Mahony, Donnellan and Roberts unpublished) but the relationship of the tetraploids to the diploid lineages has not yet been established. The time of divergence of tetraploids within the group is not known but the formation of tetravalents at meiosis I suggests that they have not undergone a rediploidization of their genome (Mahony 1986; Mahony and Robinson 1986). However, it is not known how many lineages of tetraploids actually exist in this complex nor is the nature of their origins known.

GOALS AND DISSERTATION OUTLINE

The underlying force behind this dissertation was to address the very broad question: "Why are there tetraploid species?" More specific goals were to explore the questions 1) How has polyploidy arisen (i.e., sympatry or allopatry; automatic or derived isolating mechanisms; hybridization or autopolyploidy)? and 2) What are the consequences of polyploidy after it has been established in a population (i.e., changes in gene regulation; genetic load effects; advantages compared to diploids; disomic vs tetrasomic inheritance)? The apparently recent evolution of polyploidy in certain anuran groups and the well-characterized species isolating mechanisms make them a good system in which to address these questions. However, because historical relationships among diploid and tetraploid lineages have not been clearly established from across the ranges of any of these species complexes, reconstruction of phylogenetic history was a necessary first step in addressing any questions related to the nature of speciation in these groups.

The first goal of the thesis was therefore to use mitochondrial and nuclear gene sequences to reconstruct the evolutionary relationships of diploids and tetraploids within two groups of anuran diploid-polyploid species complexes: the well studied but cryptic North American grey treefrog complex (Chapter 2) and Australian burrowing frogs in the genus *Neobatrachus* (Chapter 3). Although the original plan was to compare nuclear genes at the protein level (i.e., allozymes) and at the molecular level (i.e., sequences of nuclear genes) with gene trees obtained from mitochondrial DNA, technical and conceptual problems limited conclusive evidence based on nuclear gene sequences. Results obtained from attempts to obtain molecular sequence data from nuclear genes are discussed in a separate chapter (Chapter 4) but do not contribute significantly to the understanding of evolution within either species group. The final chapter (chapter 5) consists of an overall discussion of how conclusions on relationships among diploids and tetraploids in the two genera based on mitochondrial and/or allozyme evidence contribute to the understanding of speciation processes within these complexes. Although resolution of phylogenetic relationships was not sufficient in either group to address questions of consequences of polyploidy (i.e., the intended second goal), the final chapter also includes a general discussion of implications for the process of genome duplication in general. Limitations of the molecular characters chosen for comparison and for species concepts are also discussed.

Chapter 2

Mitochondrial DNA and Allozyme Evolution in North American Grey Treefrogs (Genus *Hyla*)

Table of Contents

ABSTRACT

The grey treefrogs of North America, *Hyla chrysoscelis* (2n=24) and *Hyla versicolor* (4n=48), are a cryptic diploid-tetraploid species pair with a broad distribution across North America. The purpose of this study was to compare mitochondrial gene sequences, allozyme allele frequencies and nuclear gene sequences from populations surrounding the Great Lakes region that have not been previously characterized. These data were used to reconstruct historical relationships of diploids and tetraploids to identify the probable number and type of origins of tetraploidy in the complex. Sequences from the mitochondrial gene cytochrome oxidase I suggested that northwestern tetraploids were probably not derived from the diploids with which they are currently sympatric; tetraploids from this region displayed haplotypes that were more similar to each other and to tetraploids from other regions than they were to any of the diploid populations examined. Conclusions based on allozyme evidence did not completely concur with mitochondrial data; taxonomic groupings tended to be more by geographic area than by ploidy level and in several cases tetraploids appeared to be most closely related to the diploids with which they are currently sympatric. This discrepancy could be evidence that tetraploids have arisen through hybridization but could also reflect differences in selective forces acting on proteins and mitochondrial genes within the same populations, differences in rates of evolution of synonymous nucleotide substitutions in mitochondrial genes versus amino acid substitutions sufficient to result in protein mobility changes in allozymes, or could be evidence of introgression between diploids and tetraploids. Both mitochondrial and allozyme data suggested that the actual number of lineages in the complex has been underestimated for the tetraploids and possibly for the diploids.

INTRODUCTION

Among anuran frogs, bisexually reproducing polyploid species occur in at least five families on four continents. The grey treefrogs (Family Hylidae) of eastern North America have been studied the most extensively. Although evidence suggests that the evolutionary history of this group may warrant division into a complex of diploid and tetraploid species, at present two names have been applied to the group: a tetraploid, *Hyla versicolor* Cope (4n=48), and a diploid, *Hyla chrysoscelis* LeConte (2n=24) (Wasserman 1970). The tetraploids are thought to have arisen during the Pleistocene glaciations (ca. 135,00 years ago) (Blair 1964; Ralin *et al.* 1983; Little *et al.* 1989) and they do not appear to have advanced to rediploidization or loss of duplicate gene expression at most loci (Danzmann and Bogart 1982a; 1982b; 1983; Marsden *et al.* 1987). Diploids and tetraploids within this group show no distinctive differences in morphology (Ralin and Rogers 1979), ecology (Ralin 1968) or physiological features (Kamel *et al.* 1985). They are found sympatrically in many areas and may be found breeding in the same ponds on the same night, but appear to be reproductively isolated by differences in mating call parameters (Johnson 1966) which females use to choose conspecifics (Klump and Gerhartd 1987; Gerhardt and Doherty 1988; Gerhardt 1991). Sterile triploid hybrids result when accidental matings do occur (Johnson 1966; Bogart and Wasserman 1972; Mable 1989; Mable and Bogart 1995). The factors that currently isolate individuals of the two ploidy levels therefore are clear.

Since polyploidization is one of the few known mechanisms for speciation in sympatry, reproductive isolation between the progenitor and daughter species would be essential, especially considering the high degree of similarity of the two species in

ecology and breeding biology. Whether this was an automatic product of genome duplication or evolved later to reduce the potential for introgression with the parental population is unclear. Prediction of type and number of origins of the tetraploid lineages is crucial to the interpretation of reproductive isolating factors leading to speciation and evaluation of subsequent genetic consequences of polyploidy.

Previous studies have suggested that the present distribution of the tetraploids divides distinctive eastern and western populations of diploids that have been characterized by unique allozymes (Ralin 1978; Ralin and Selander 1979; Ralin *et al.* 1983; Romano *et al.* 1987), immunological distance data (Maxson *et al.* 1977), mating calls (Gerhardt 1974; Ralin 1977), chromosomal polymorphisms (Wiley 1983; Wiley *et al.* 1989), restriction site polymorphisms of nuclear ribosomal DNA genes (Romano and Vaughn 1986) and mitochondrial DNA sequence variation (Ptacek *et al.* 1994). These studies have not agreed on whether the tetraploid arose through a genetically intermediate form of the diploid or from one of the two diploid "forms" (autopolyploid) (Ralin 1977; Ralin and Selander 1979; Ralin *et al.* 1983; Romano *et al.* 1987); through hybridization of the two diploid forms (allopolyploid) (Maxson *et al.* 1977; Romano and Vaughn 1986); or whether polyploids arose through a direct doubling of the chromosomes or an intermediate triploid stage (Bogart and Wasserman 1972). Discrepancies in interpretation of results by different researchers have led to conflicts on whether the division of the diploids into two species is warranted and how allopolyploidy versus autopolyploidy should be recognized and defined. For example, the presence of multivalents at meiosis (Bogart and Wasserman 1972) suggests similarity between all four homologous chromosomes in tetraploids that would be expected with an autopolyploid origin.

However, the chromosomes of eastern and western diploids differ only by a shift in NOR (nucleolar organizer region) position that is polymorphic within eastern diploid populations (Wiley 1983) and these differences may not be sufficient to prevent multivalent chromosome segregation in tetraploids arising through hybridization. Evidence from restriction site polymorphism in ribosomal genes suggests that tetraploids contain ribosomal repeats from both eastern and western diploids and therefore may have arisen through hybridization (Romano and Vaughn 1986). However, diploid populations are not themselves absolutely distinct from each other and so this also does not in itself prove a hybrid origin.

It also has not been completely resolved whether the tetraploid had a single origin or whether polyploidy arose multiple times. Morphometric analyses (Ralin and Rogers 1979), combined with allozyme data (Ralin and Selander 1979) and mating call variation (Ralin 1977) have been used to suggest that southern populations of tetraploids are genetically intermediate between eastern and western diploids in Texas, but tetraploids from the northeast (New York) may be more closely related to populations of diploids from eastern Texas than to western Texas (Ralin *et al.* 1983). In addition, tetraploids from Virginia were polymorphic for two NOR positions, one of which is not found in other tetraploid populations (Wiley *et al.* 1989). This could suggest a separate origin of tetraploids in this region (Romano *et al.* 1987). While this evidence seems to suggest that multiple origins are most likely, these authors (Ralin *et al.* 1983; Romano *et al.* 1987) consistently claim that their allozyme data support a single origin and that similarites are due to convergence resulting from parallel selection. The least ambiguous evidence to date comes from comparison of mitochondrial DNA sequence data, which suggests multiple origins of the polyploid

(Ptacek *et al.* 1994). This study concluded that southwestern tetraploids (Oklahoma, Texas, Louisiana) were most closely related to eastern diploids, that eastern (Virginia, West Virginia, Maine) tetraploids were most closely related to western diploids, and that northwestern tetraploids (several counties in Missouri, Minnesota, and Ontario) were not closely related to any of the diploid populations sequenced. The lack of a diploid ancestor for the northwestern tetraploids may be a result of the limited number of samples included from the northern parts of the range. Because of this, the actual number of tetraploid lineages may have been underestimated.

Distributions of diploids and tetraploids in the Southern states from Florida to central Texas, as well as many of the central states up to Illinois, Indiana, Ohio and Pennsylvania have been described in the literature and many of these populations have been characterized using morphometric analyses (Ralin and Rogers 1979), allozymes (Ralin 1978; Ralin and Selander 1979; Ralin et al 1983; Romano *et al.* 1987), mitochondrial DNA variation (Ptacek *et al.* 1994), and mating call variation (Ralin 1977; Jaslow and Vogt 1977; Bogart and Jaslow 1979; Gerhardt 1978). Both diploids and tetraploids have been found in Kansas (Hillis *et al.* 1987), Virginia, Missouri, Oklahoma, Texas and Louisiana, often under sympatric conditions (Bogart, personal communication), but only diploids have been found in the southeastern states. Diploids and tetraploids in the Appalachian highlands (Ohio, West Virginia and Pennsylvania) appear to be mainly allopatric, with diploids occurring throughout the unglaciated Allegheny Plateau and tetraploids occurring in the glaciated central lowland of Ohio and throughout the Greenbriar and Potomac River drainages of the Ridge and Valley Province in West Virginia. A narrow band

of sympatry occurs in West Virgina but more extensive sympatry occurs throughout Pennsylvania (Little *et al.* 1989).

Distribution of diploids and tetraploids in the more northern and western parts of the ranges and definition of the extent of sympatry in these areas have been more anecdotal, and few studies have examined genetic attributes or mating call variation in these regions. Ralin and Selander (1979) identified the need for examining northern populations because of allozyme and morphological evidence that they may have had an independent origin. They examined northeastern populations of tetraploids from New York and Vermont (Ralin *et al.* 1983) but did not sample further west. The Great Lakes appear to define the northern limit of the ranges of diploids in the central parts of their range (Bogart, personal communication) and may represent a geographical barrier that resulted in isolation as a result of glacial retreat. Tetraploids have been found in populations throughout central and southern Ontario (i.e., across the northern border of the Great Lakes) and extend as far west as eastern Manitoba and as far east as New Brunswick at roughly the same latitude. Sympatric populations of diploids and tetraploids have been found in most of the states forming the southern border of the Great Lakes and in Manitoba, which represents the northwestern limit of the ranges of both diploids and tetraploids. In their analysis of cytochrome b sequences, Ptacek *et al.* (1994) included a single tetraploid from Ontario and a diploid and tetraploid from allopatric populations in Minnesota but did not include any other populations north of Missouri. Variation across the northern distribution of tetraploids and in sympatric northern populations has therefore not been assessed at the molecular level using mitochondrial markers or allozymes and the number and type of origins in this region remain unclear.

Thus, despite almost 30 years of study, the nature of origins of tetraploidy in this complex has still not been resolved satisfactorily. The purpose of this study was to examine variation in both mitochondrial DNA sequences and allozyme allele frequencies in populations of diploid and tetraploid grey treefrogs from the northern and western portions of their ranges that have not been previously characterized. Specific goals were to determine the likelihood that northern tetraploid populations had a unique origin and to determine if a possible diploid ancestor could be identified. Attempts to use sequences from single-copy nuclear genes to study relationships among these populations are discussed in Chapter 4.

MATERIALS AND METHODS

Samples

Samples were obtained from populations from across the northwestern and north central ranges of grey treefrogs by opportunistic collecting during the breeding season (Table 2.1; Figure 2.1). Most samples were collected during field surveys in the springs of 1988 and 1992. Some samples were obtained from a tissue bank collection in the laboratory of James P. Bogart at the University of Guelph that represented samples collected since 1967 (Table 2.1). Electrophoresis was performed on samples collected prior to the 1992 field survey; mtDNA sequence analysis was performed on DNA obtained from toes of individuals collected in 1992. The same populations used for the allozyme analysis were visited in 1992 but it was not possible to collect from all of these because males were not calling at the time of the survey. In addition, a "new" population in Ontario (Rainy River) and several

populations from Texas that had not been included in the allozyme study were added to the mtDNA study.

The populations sampled represent a mixture of allopatric and sympatric populations mainly from the northern part of the range of diploids and tetraploids in North America. In Table 2.1, "sympatric" refers to populations where diploids and tetraploids were found calling in the same ponds; "parapatric" refers to populations where diploids and tetraploids were collected in the same region but not from the same ponds; and "allopatric" refers to populations outside of the known range of the other ploidy level. Samples from north of the Great Lakes in Ontario (Ottawa, Guelph, Aberfoyle, Rosseau, Rainy River) represent areas where tetraploids are completely isolated from diploid lineages. Samples from Manitoba and Minnesota represent the westernmost distribution where diploids and tetraploids are sympatric and probably represent the "western" diploid lineage. Samples from the southern border of the Great Lakes (Wisconsin, Michigan) represent mainly sympatric populations where diploids and tetraploids are found together, often in the same ponds; these populations could contain either "eastern" or "western" diploids, as they have not been previously characterized. Samples from Bastrop Co., Texas represent populations where only tetraploids have been found within at least the last four years but which previously have been described as containing both diploids and tetraploids. Samples from Travis Co., Texas represent the "western" lineage of diploids, in a location where tetraploids have not been reported. Diploids from Comal Co. and Lamar Co., Texas have not been previously characterized but the former probably represents the "western" lineage and the latter the "eastern". Populations from

Virginia, West Virginia and Maryland represent mixed sympatric and allopatric populations where the ranges of "eastern" diploids and tetraploids overlap.

Diploids and tetraploids were initially identified by differences in mating calls based on temperature regressions (Gerhardt 1978) and were verified by chromosome counts (Bogart 1967; see also Mable 1989 for protocols), blood cell measurements (Mable 1989; Matson 1990), or dosage in allozyme patterns (Danzmann and Bogart 1983). Mating call is a reliable criterion for males and so only a representative sample of males from each population was karyotyped but identification was confirmed by karyotype for all females. Samples obtained from the tissue bank collection relied on allozyme dosage to confirm ploidy level but careful records had also been kept at the time of collection on species identification.

It is interesting that the Virginia and Manitoba populations were the only ones for which numbers of diploids and tetraploids calling in the ponds (and subsequently sampled) were relatively equal. In all other sympatric populations or parapatric regions, tetraploids tended to vastly outnumber diploids. Gerhardt *et al.* (1994) found relatively equal numbers of tetraploids and diploids in a pond in Missouri but this proportion tended to vary across years. Most of the populations I sampled were dominated by tetraploid choruses, but I cannot determine if this was due to decline of diploids, recent invasion of small populations of diploids, or if diploids and tetraploids call at different times. Sample sizes for electrophoresis generally reflect relative population sizes of calling males at the time of collection.

Mitochondrial DNA

Sequencing

Toes were clipped from several individuals from each population. Toes were ground in TE buffer and DNA isolated using PCI extractions as described in Sambrook *et al.*(1989) and Hillis *et al.* (1996a). Yields averaged between 35 and 100 ng of DNA per toe, which is sufficient for PCR amplification using most primers. The advantage of using toes rather than tissue samples is that populations can be screened without killing or removing animals from the population.

The polymerase chain reaction (PCR) (Saiki *et al.* 1985, 1986) was used to amplify a 600-bp region of cytochrome oxidase I (COI) using primers modified to frog (*Xenopus*) sequences from the COIe-5' and COIa-3' primers described in Palumbi (1996). Primer sequences were: 5' CCT GCA GGA GGA GGA GAY CC 3' (forward primer) and 5' TGT ATA AGC GTC TGG GTA GTC 3' (reverse primer). PCR products of the size fragment corresponding to this region were gel purified using low-melting-point agarose. Target bands were cut from the gel and DNA recovered using a commercially available purification method (Prep-a-gene®; BioRad). Purified PCR templates were sequenced using a thermal cycle sequencing method (Craxton 1991; Murray 1989), based on the BRL dsDNA Cycle Sequencing system. This technique is based upon the dideoxy chain termination method of Sanger *et al.* (1977) but utilizes end-labeled primers in a linear amplification procedure that results in efficient sequencing of even very small amounts of template (see also Hillis *et al.* 1996a). Sequences were generated for both strands of DNA from each individual to detect possible PCR artifacts or ambiguities in sequences generated.

Analyses

Sequences were aligned using MacVector (International Biotechnologies, Inc) and multiple alignments were optimized with the program Clustal V (Higgins and Sharp 1988, 1989; Higgins *et al.* 1992). Alignments were verified by eye based on codon positions. Using a test version of PAUP*, version 4.0d42 (Swofford 1996), sequences were analyzed in three sets of analyses: using maximum parsimony, maximum likelihood and minimum evolution as optimality criteria. In each case, heuristic searches were conducted with initial trees obtained by simple stepwise addition, followed by branch swapping using the TBR (tree bisection-reconnection) implemented by PAUP*. Only minimal length trees were saved and zero length branches were collapsed. Multistate characters within taxa were interpreted as uncertainties and an accelerated transformation procedure (ACCTRAN in PAUP*) was used for character optimization (PAUP*). Bootstrap proportions based on 100 replicates were used to assess the relative degree of support for branching relationships predicted from each reconstruction method. Transition: transversion biases and base pair frequencies were estimated using maximum likelihood (PAUP* 4.0d42; Swofford 1996) based on the trees found in an initial search using parsimony with characters unordered and unweighted. The average number of synonymous to nonsynonymous substitutions was estimated using the method of Nei and Gobojori (1986) using a program written by D. Haydon at Oxford University.

Allozymes

Electrophoresis

Tissue samples from heart, muscle, spleen and liver were taken from adults and frozen with an equal volume of double deionized water in 1.5 ml microcentrifuge tubes at -70°C. Heart, muscle and spleen samples were combined in the same tubes.

Electrophoresis was performed according to Bogart *et al.* (1985) and Utter *et al.* (1975), using buffers and stains described by Selander *et al.* (1971), Shaw and Prasad (1970), and Clayton and Tretiak (1972). The Clayton and Tretiak (1972) amine-citrate buffer system was altered from pH 6.1 to pH 6.5. After an initial survey of 32 presumptive locus products, 16 clearly resolved, unambiguously scored, polymorphic loci were selected for phylogenetic analysis (Table 2.2). Enzyme Commission (EC) code numbers (International Union of Biochemistry, Nomenclature Committee 1984) for these loci follow. 6PGD (phosphogluconate dehydrogenase EC 1.1.1.44); ACON (aconitase hydratase or Acoh; EC 4.2.1.3; = aconitase); GOT (aspartate aminotransferase EC 2.6.1.1 or AAT; = glutamate-oxaloacetic transaminase); IDH (isocitrate dehydrogenase EC 1.1.1.42); MDH (malate dehydrogenase EC 1.1.1.37); MPI (mannose-6-phosphate isomerase EC 5.3.1.8); PGI (glucose-6-phosphate isomerase EC 5.3.1.9 or GPI; =phosphoglucose isomerase); PGM (phosphoglucomutase EC 5.4.2.2); SOD (superoxide dismutase EC 1.15.1.1); αGPD (a-glycerophosphate dehydrogenase or glycerol-3-phosphate dehydrogenase EC 1.1.1.8 or G3PDH).

After electrophoresis and staining, each gel was photographed, and the negatives projected through a microfiche reader. Relative mobilities of the

electromorphs were determined by measuring the distance from the origin to the band using a sonic digitizer. Mobilities of the electromorphs were defined according to the mobility of the furthest migrating, which was assigned an arbitrary value of 100. Phenotypes of heterozygous tetraploids were determined by relative intensity of staining of allozyme bands within loci (see Danzmann and Bogart 1982a, b). The electrophoretic data were stored and analyzed for gene frequencies using GENEPAC, a compiled computer software package written for the IBM PC by James P. Bogart, at the University of Guelph which was designed specifically to compute comparable allele frequencies for diploid and tetraploid genotypes.

Analysis

Frequency data were analyzed initially using BIOSYS-1 (Swofford and Selander 1981) to construct UPGMA trees from distance matrices based on Nei's genetic distance (Nei 1972), and Rogers' genetic distance (1972). The data were also analyzed using a parsimony-based method that allows incorporation of frequency data into a step matrix used to establish character weightings. This method was developed for use with allozyme data by Paul Chippindale and is similar to one proposed by Berlocher and Swofford (1996). Manhattan distances were constructed for each allozyme locus between all pairs of taxa using the formula

$$D_M = \frac{1}{2}\sum|x_i - y_i|$$

(Swofford *et al.* 1996). Distances were rounded to three figures and then multiplied by 100 to standardize values across loci. For each locus, these values were used to construct a step matrix to weight each unique combination of allele and frequency data according to the Manhattan distances among taxa for each unique array. The data matrix was constructed by assigning each allele-frequency array for a given locus a letter corresponding to that used in the step matrix for that locus. A sample partial data file is provided in Appendix A. The advantage of this type of approach is that it provides a weighting for presence/absence of alleles that preserves the information in the frequencies of allelic combinations within and between populations. This was especially important for the grey treefrogs because few populations showed fixed differences at the loci examined and diploids and tetraploids shared most alleles. These data were then analyzed using parsimony as described for the mtDNA data (PAUP* 4.0d42; Swofford 1996).

The broadest surveys of allozyme variation in grey treefrogs across North America are those by Ralin *et al.* (1983) and Romano *et al.* (1987). To more realistically compare their results with the results presented here, the allele frequencies presented in each of these papers were reanalyzed using the parsimony-based Manhattan distance stepmatrix approach. Although direct comparisons could not be made with their results and mine, this reanalysis was included for general comparison of diploid and tetraploid relationships in other regions.

RESULTS

Mitochondrial DNA

A 589-bp region of cytochrome oxidase I displayed 28 variable sites among the taxa sequenced. Sequence alignments in a sample PAUP file are shown in Appendix B. Attempts to sequence COI from allozyme tissues were unsuccessful so all sequences shown are from DNA isolated from toes. Sequence divergence among pairs of taxa sequenced ranged from 0 to 4.2% (Table 2.3b). Most substitutions were synonymous transitions (Table 2.3a). Transversions occurred at three sites, all sites except one having the same type of change in all taxa. Based on the trees with the highest likelihood from the most parsimonious trees found in an initial search, the transition:transversion ratio was estimated to be 7.25. However, this value varied across trees from 5.78 to 7.25. Base pair frequencies did not appear to be homogeneous (A=0.27459; C=0.21798; G=0.17509; T=0.3324); there appeared to be a bias against Gs and in favour of Ts. The average ratio of synonymous to nonsynonymous changes across the data set was 82:1.

Pairwise sequence divergence among populations was used as an initial indication of relationships among taxa (Table 2.3b). Haplotypes of tetraploids from Minnesota and Northwestern Ontario (Rainy River) were identical, as were those from the diploid and tetraploid from Wisconsin. Except for the Wisconsin population, sequences from Northern tetraploids were more similar to each other than to any of the diploids. Pairwise sequence divergence among the northern tetraploids was 1.0 %, while divergence from the closest diploid was 2.4% (i.e., Michigan diploid to Minnesota/northwestern Ontario tetraploid). The largest divergence was between northern tetraploids and diploids from the Lamar County, Texas population

(4.1-4.2%). The tetraploid from Bastrop County, Texas was not as similar to the northern tetraploids as they were to each other (1.2-1.5%) but was still more closely related to them than to any of the diploids. In relation to the diploids, this tetraploid was most closely related to the Wisconsin diploids (2.2%), and most distantly related to the Lamar County, Texas diploid population (3.2%).

Diploid populations tended to be less divergent from each other in general than were tetraploid populations. Diploid populations from Wisconsin and Travis County, Texas were most similar to one another (0.2%) and diploids from Michigan and Lamar County, Texas populations were most distant (2%).

An exhaustive search using parsimony with characters unordered and unweighted yielded a single most-parsimonious tree (PAUP* 4.0d42; Swofford 1996). This tree topology did not appear to be sensitive to the type of character weighting scheme or to the optimality criterion used, as all analyses resulted in the same branching relationships. The midpoint rooting tree for the unweighted parsimony search is shown in Figure 2.2a (with bootstrap support indicated). There was strong support that tetraploids were more closely related to each other than to the diploids (100%), some support that tetraploids from Ontario and Michigan were more closely related to each other than to tetraploids from Minnesota (65%) and that tetraploids from Texas were distinctive from those from the more northern populations (96%). There was also strong support that diploids from Wisconsin, Michigan, and the populations from Travis County, Texas were each other's closest relatives (91%), some support that the Lamar County diploids were not included within this clade (89%) and strong support that the Minnesota diploids fell outside of

this clade (89%). Visualization of degree of divergence between diploids and tetraploids is clearer using an unrooted phylogram (Figure 2.2b).

Allozymes

Allele frequencies for allozyme loci used in the analysis and population sample sizes are shown in Table 2.4. Table 2.4a shows the populations that were collected in the field and Table 2.4b shows the additional populations used from a tissue bank collection. Samples from the tissue bank collection were not stained for Acon-2, Pgi-1, Pgi-2 or αGPD and scoring of alleles from Acon-1 did not appear to be the same. Genetic distances based on Roger's genetic distance, average heterozygosity, and Nei's genetic distance, are shown for the field-collected populations that were complete for all loci (Table 2.5). Average heterozygosity within populations ranged from 0.079 to 0.173 (Table 2.5).

Diploid and tetraploid grey treefrogs from northwestern, north/central, and northeastern populations appeared to share the same alleles and there were no fixed allelic differences at any of the loci examined. Rare alleles were found in some of the larger sample sizes but this may be a sampling effect and not reflect meaningful phylogenetic information. Percentage of polymorphic loci appeared to be more related to sample size than to ploidy level. For example, diploids from Minnesota had half the polymorphic loci compared to tetraploids from those populations (40 and 80, respectively), but the respective sample sizes were 3 and 46. However, diploids and tetraploids from Virginia showed approximately the same percentage of polymorphic loci (60 and 65, respectively) and the sample sizes were 24 and 25.

A phenetic analysis using UPGMA with Roger's genetic distance for the field-collected samples suggested a general division into eastern and western groups, each containing both ploidy levels (Figure 2.3a). The eastern group was composed of Wisconsin, Virginia, Maryland and Ontario populations and the western group of Minnesota and Manitoba populations. Within the eastern group, the Maryland population (where tetraploids are not found), appeared to be the most divergent. The next deepest split was a cluster involving tetraploids from Ontario and tetraploids from Wisconsin (labelled "Northern 4n"). The Wisconsin sample was very small (N=4) but did appear to be more similar to Ontario tetraploids than to the two populations of diploids from Wisconsin, which clustered together. The Virginia tetraploids however, appeared to be more similar to diploids from the same population. This was the only sympatric population (i.e., 2n and 4n calling in the same pond) sampled from the eastern populations. Within the western cluster, sympatric diploids and tetraploids from Manitoba appeared to be more similar to each other than to either ploidy group from Minnesota. Comparisons between diploids and tetraploids from Minnesota could be biased by large sample size discrepancies (3 diploids versus 46 tetraploids), but there appeared to be a deeper divergence between them than among the other diploid-tetraploid pairs.

Using the same populations, an heuristic search using parsimony (as described in the methods section) and the locus-specific stepmatrices described above yielded a single most parsimonious tree (Figure 2.3b) that did not completely concur with the UPGMA tree. The tree did support the same general division of diploids into eastern (Virginia and Maryland) and western (Minnesota and Manitoba) groupings, but diploids from Wisconsin appeared to be more closely related to

western diploids than to eastern (Figure 2.3b). Relationships of tetraploids and diploids within regions also varied but tetraploids appeared to be divided into the same general eastern and western groups. Within the clade from Minnesota and Manitoba, diploids appeared to be more closely related to each other than to tetraploids from the same population (these were both sympatric populations). This was also true of the populations from Virginia and Maryland. However, tetraploids from Wisconsin and Ontario (Aberfoyle) did not appear to be closely related to any of the diploids. Bootstrap analysis (100 replications) of this data, supported only two clades: a clade involving Manitoba and Minnesota populations (50%) within which there was a trichotomy for diploids from both populations and tetraploids from the Manitoba population (75%); and a clade involving diploids from Virginia and Maryland (61%). No other relationships were strongly supported.

Populations from the tissue bank collection were included in an additional parsimony analysis to include a broader survey of sympatric versus allopatric populations. The single most parsimonious tree from a heuristic search suggested a division of populations into three major clades: one involving diploids from Wisconsin and diploids and tetraploids from Manitoba and Minnesota ("western"); a second composed of tetraploids from Ontario (Ottawa, Guelph, Aberfoyle, Rosseau), Michigan, and Wisconsin, and both diploids and tetraploids from West Virginia ("northeastern"); a third involving diploids from Virginia and Maryland, and possibly including tetraploids from Virginia ("eastern") (Figure 2.4). In addition to the clades supported in the analysis of field collected populations, bootstrap analysis (100 replicates) indicated support for a clade involving diploids and tetraploids from West Virginia (77%).

Reanalysis of the frequency data presented in Ralin *et al.* (1983) using the parsimony-based method yielded a single most parsimonious tree that suggested that tetraploids from Texas and Illinois were most closely related to "western" diploids, Illinois diploids fell within the "eastern" diploid clade, and that tetraploids from Vermont and New York ("northeastern" tetraploids) were more closely related to each other than to any of the other taxa (Figure 2.5). The use of stepmatrices without a defined outgroup can lead to visual representations of relationships that may be misleading. An unrooted phylogram indicates that tetraploids could be interpreted to be intermediate between eastern and western diploids, as was suggested in the original paper (Figure 2.5b). Unfortunately, the frequency data provided in the original paper sums over "eastern" diploids (pooled across populations from Mississippi, Ohio, South Carolina, and Georgia), "western" diploids (pooled across Texas diploid populations), and Texas tetraploid populations, even though the trees presented are based on frequencies in individual populations so direct comparisons cannot be made.

Reanalysis of the data presented in Romano *et al.* (1987) using the parsimony-based method described here resulted in five equally parsimonious trees, one of which is shown in Figure 2.6a. The strict consensus tree was not well resolved and bootstrap analysis supported only a single node, involving diploids and tetraploids from Tyler, Texas (Figure 2.6b). However, populations can generally be divided into eastern and western, each including both diploids and tetraploids (Figure 2.6a).

A combined analysis using populations from both Ralin *et al.* (1983) and Romano *et al.* (1987) was also conducted. Using all populations from both papers

resulted in 26 equally parsimonious trees with little resolution. Taxon sampling appeared to have a significant effect on the trees resolved as relationships and number of equally parsimonious trees found varied greatly when certain populations were removed. Figure 2.7b shows the single-most parsimonious tree found when pooled samples from the Ralin *et al.* (1983) data (Texas tetraploids, "eastern" diploids, "western" diploids) were removed from the analysis. This tree suggests division of the populations into a "western" clade involving diploids and tetraploids from Missouri, diploids from Bastrop, Texas, tetraploids from Travis Co., Texas, Oklahoma diploids and tetraploids, and some of the illinois diploids and tetraploids (Mason Co. diploids and tetraploids; Fayette County tetraploids); and a "northeastern" clade involving populations of diploids and tetraploids Moultrie and Macon Counties in Illinois (from the same regions or the same samples used by Ralin et al. (1983) "2n Illinois" and "4nIllinois"), Indiana, Vermont, New York and Ohio. Diploids and tetraploids from Tyler, Texas and diploids from Fayette, Illinois may be separate from the "western" clade and are located further east than the diploids in the "western" clade but their placement is uncertain. Note that the placement of these populations is the major difference between the combined and uncombined analyses.

DISCUSSION

Allozymes versus mtDNA

Allozymes suggest that eastern and western groupings of tetraploids can be distinguished in which tetraploids are more similar to their sympatric or parapatric diploid counterparts than they are to tetraploids from the opposite part of the range,

whereas mtDNA sequences suggest that eastern and western groupings of diploids can be identified but tetraploids are more similar to each other than they are to their sympatric diploids. Some authors claim that a "total evidence" approach to phylogenetic analysis will usually be more informative than separate analyses that support different hypotheses of relationships among taxa (e.g., Kluge 1989). However, this kind of approach does not seem reasonable when biochemical or evolutionary reasons for discrepancies between types of data can be identified that may not reflect phylogenetic history. Combining data in these cases risks losing information gained from each because averaging over different evolutionary processes could obscure the actual history of taxonomic interactions.

The discrepancy between mitochondrial DNA and allozymes could have several explanations. The first explanation is that tetraploids arose through a hybridization event and differences in conclusions reflect differences in the inheritance of mitochondrial versus nuclear markers. Mitochondrial genes represent maternal gene phylogenies (Avise and Lansman 1983) and indirectly could document the origin of the cytoplasm that gave rise to the tetraploids (Hedges *et al.* 1992). The predominant mode of polyploid speciation in animals appears to have been through hybridization (Bogart 1980) but allopolyploid origins in tetraploid frogs has only recently been demonstrated conclusively, in South African *Tomopterna* (Channing and Bogart, personal communication). In that study, differences in allozymes versus mtDNA were used to predict the direction of hybridization based on fixed differences between parental species. For grey treefrogs, fixed differences between populations are rare and detecting evidence of hybridization using allozymes would thus be more difficult, so the likelihood of hybridization cannot be determined in this case. A

previous mtDNA study using cytochrome b (Ptacek *et al.* 1994) suggested that southwestern tetraploids arose from eastern diploids while eastern tetraploids arose from western diploids and that the few northwestern populations included (Minnesota and Southern Ontario) may have had a separate origin from an unidentified diploid haplotype. My COI data suggest that northwestern tetraploids are not similar to any of the diploid populations examined. Therefore, it could be that the maternal ancestor no longer exists in this area. If tetraploids arose as a result of a hybridization event, the allozyme data could suggest that tetraploids in the northwestern region each had different paternal progenitors with which they have remained in close contact. An intriguing idea is that polyploidy arose in this group whenever males hybridized with females from a particular diploid population that has yet to be identified in studies of extant populations. This would correspond with Ralin *et al.*'s (1983) suggestion that tetraploids arose from a genetically intermediate population of diploids that may itself have arisen through hybridization or may reflect a cline in diploid genotypes from east to west. Patterns of movement of treefrogs have not been well documented, but high site fidelity has been suggested anecdotally. If tetraploids tend to remain in the ponds in which they arose and are reproductively isolated from their diploid counterparts automatically with the change in chromosome number and thus do not experience introgression with their paternal progenitors, then it would not be unreasonable to assume that they would share many of the same alleles but have different mitochondrial haplotypes. This would require that the maternal and paternal originators shared similar protein alleles but were divergent in mitochondrial DNA, as might be expected if hybridization occurred early in their divergence from one another.

Allozymes represent gene products and may be altered by differences in regulatory control mechanisms that could lead to dominant expression of one parental allele over the other in hybrids. An alternative explanation therefore is that similarity to the diploid with which tetraploids are sympatric may be due to differential expression of proteins. The use of allozymes for study of polyploid individuals is complicated by the increased complexity that might be expected when four copies of genes coding for a particular protein exist within an individual. Some authors have rejected the use of frequency data in polyploid lineages because of the difficulty of determining dosage (Mahony *et al.*, unpublished). Dosage of allozyme bands can actually be determined fairly accurately in heterozygous polyploids (Danzmann and Bogart 1982a,b; 1983), but only if it is assumed that all genic forms present are actually expressed and are resolvable under the electrophoretic conditions used. For example, in a dimeric enzyme such as 6PGD, three bands are present in diploids heterozygous for two electromorph alleles: two homomeric bands and a heteromeric band. Intensity of staining of the two homomeric bands would be expected to be equal. In a tetraploid, a heterozygote for two alleles would still show three bands but the intensity of staining would vary depending on whether the alleles were present in equal copy number or if one allele was expressed from three chromosomal copies and the other from only one. If a tetraploid did not express four copies of the gene, an asymmetric staining pattern could reflect a "null" allele at one of the loci and the number of copies of the other allele would be overestimated.

A third explanation is that similarity of allozymes in sympatric populations is a reflection of similar environmental pressures resulting in selection for particular protein combinations and does not reflect phylogenetic divergence. Romano *et al.*

(1987) used the finding that every population of tetraploid that was sympatric with a diploid contained alleles present in other tetraploid populations but not in the diploid sympatric with it to support a hypothesis that tetraploids had a single origin. The rationale behind these conclusions was that fixed allelic differences among populations of grey treefrogs were rare and it was mainly frequencies that varied. Since frequencies of alleles may be subject to fluctuation and may respond more quickly to selection, they suggested that it was more parsimonious to conclude a single origin and invoke parallel selection to explain why tetraploids appeared in more than one diploid grouping. Although the reanalysis of their data and the study of northern populations presented in this thesis suggest multiple origins, the hypothesis of parallel selection cannot be completely refuted because of the finding that tetraploids found in sympatry with diploids tended to be more similar to those diploids, while tetraploids found in allopatry appeared to be more closely related to each other. The mitochondrial data are also consistent with a single origin but this study did not include as broad a survey across the range of diploids and tetraploids as the allozyme study and previous mitochondrial evidence (Ptacek *et al.* 1994) provides fairly conclusive evidence of multiple origins of polyploidy in regions not sequenced for COI. Selection at the biochemical level is very difficult to demonstrate in natural populations, although a recent study of allozyme polymorphisms in *Colias* butterflies suggests that it is not entirely impossible (Watt *et al.* 1996). While few studies have shown conclusive evidence of this kind of selection at the protein level, it cannot be excluded as a possible factor.

A fourth explanation has to do with relative rates of evolution. It has been well documented that wide differences may exist in the rates of evolution apparent for

mitochondrial and nuclear DNA. Tetraploids in the grey treefrog complex are thought to have arisen during the Pleistocene glaciations and are therefore relatively young "species". Mitochondrial genes in general tend to evolve at a faster rate than nuclear genes and nucleotide divergence might be expected to appear much sooner than in the nuclear genome (e.g., Brown *et al.* 1979). In addition, allozymes represent protein products and therefore only indirectly represent changes in the nuclear genome. Changes in protein mobility would not be expected to be evident until nonsynonymous nucleotide substitutions (i.e., those resulting in amino acid substitutions) occur. Although the mitochondrial gene examined here (cytochrome oxidase I) has been shown to have sufficient variation for population level comparisons within species in other frog groups (e.g., C. James, personal communication), in comparisons of the diploid and tetraploid populations used here the average ratio of synonymous substitutions per synonymous site to nonsynonymous substitutions per nonsynonymous site was 82:1. While this may be evidence that COI is under strong selective constraint and cannot tolerate radical changes in amino acid composition, it is also possible that this reflects the recent divergence of populations in this complex. The young age of tetraploids may mean that most changes in protein coding nuclear genes may not be detectable using allozymes because there is insufficient divergence in the diploid and tetraploid taxa examined.

Finally, it is possible that there has been continued introgression and gene flow between diploids and tetraploids within sympatric populations that have resulted in more similar allele frequencies within populations than between ploidy levels in some cases. One theory of how tetraploids have arisen is through an intermediate

triploid stage. Triploids have been found only very rarely in natural populations (Gerhardt *et al.* 1994). Artificial hybridization experiments have shown consistently that hybrids between tetraploid females and diploid males are sterile and hybrids between diploid females and tetraploid males do not advance beyond the gastrula stage (Mable 1989; Mable and Bogart 1995). This suggests that if gene flow were to occur from diploids to tetraploids (i.e., if hybrid sterility was not complete), it would be through paternal contributions only. If this were the case, one might expect to find similarity of allele frequencies of nuclear genes that were not apparent in mitochondrial genes. Paternal leakage of mitochondrial genes has been documented but appears to be a relatively rare occurrence. Given the sterility of artificially produced hybrids and the rarity of naturally occurring triploids (Gerhardt *et al.* 1994), it does not seem likely that gene flow is a significant factor in current populations. The cause of the hybrid sterility is not known and it may have developed as the result of gene divergence or structural changes in chromosomes (possibly to reduce the amount of extraneous genetic material carried) rather than being an automatic consequence of polyploidization. In addition, Gerhardt *et al.* (1994) suggested that mismatings were more frequent in the population that they surveyed than expected by chance but actual numbers of mismated pairs were not high. It may be that introgression and gene flow occurred in the past but have ceased in present populations and that differences in proteins may not yet be apparent.

Thus, because it is difficult to evaluate which of these alternatives is most likely, the following discussion will be restricted to conclusions that would be drawn by examining each data set alone and to conclusions that are common to both data sets.

Geographic Variation: Allozymes

Using both UPGMA and parsimony methods, conclusions based on allozyme variation in grey treefrogs from across the northern regions of their range are consistent with the hypothesis that there is a geographic division of diploids and tetraploids into eastern and western groupings. They do not, however, agree absolutely on where these divisions lie. In both sets of analyses the "western" group appears to be composed of populations from Manitoba and Minnesota and the eastern group of populations from Virginia and Maryland but they do not agree in the placement of populations from Wisconsin or in specific relationships between diploids and tetraploids. Both analyses do support the conclusion that diploids and tetraploids within regions are more closely related to each other than to their own ploidy level from different regions. In addition, the parsimony analysis including the tissue bank samples suggests that tetraploids from the more central part of the range (Ontario, Michigan, West Virginia) may be more closely related to each other than to either the "eastern" or "western" groups (Figure 2.7a). Inclusion in this clade of diploids from West Virginia could indicate that this population represents the intermediate diploid form that Ralin *et al.* (1983) and others (e.g., Romano *et al.* 1987) have predicted was the most likely ancestor for the tetraploid lineages.

Comparison of these data with the reanalysis of the data of Romano *et al.* (1987) and Ralin *et al.* (1983) allows at least a tentative prediction of geographic population divisions based on allozyme frequencies (Figure 2.7). The parsimony data presented here shows evidence of "eastern" and "western" forms that include both diploids and tetraploids and suggests that there may be an additional allozyme

"type" in the central parts of the range, at least in the more northern states and in populations from Ontario (i.e., north of the Great Lakes). The relationship of this third "type" to the other "lineages" does not appear to be well resolved. It is unfortunate that differences in interpretation of allozyme migration and differences in buffer and running conditions do not allow combination of results from different researchers into the same phylogenetic analysis. Nevertheless, conceptually combining conclusions based on these studies leads to the conclusion that populations in the central part of the range (eastern Texas, eastern Illinois, West Virginia, Michigan, Ontario and Wisconsin) represent intermediates between "western" (western Texas, Missouri, Minnesota, and Manitoba) and "eastern" (Virginia, Maryland, eastern Illinois, New York, Vermont) populations (Figure 2.7). This would be consistent with the conclusions of Ralin and coworkers that tetraploids may have arisen from a genetically intermediate diploid in the central part of the range as has been discussed above. The inclusion of tetraploids in each of the three clades identified suggests that they have experienced multiple origins, possibly from different diploid progenitors. Considering that tetraploids found sympatrically with diploids tend to be more similar to their diploid neighbours than tetraploids found allopatrically, the possibility still remains that parallel selection on electromorph frequencies explains this allozyme pattern. The analysis of allozyme data using parsimony presented here therefore cannot conclusively resolve relationships much further than the work of previous researchers, except to add further populations to the piecemeal data pool.

Geographic Variation: Mitochondrial DNA

Because of the problems with the assumption that expressed genes reflect actual differences in allelic sequences, mitochondrial data may be more reliable. However, because mtDNA is nonrecombining and maternally inherited, all parts of the molecule share the same historical pattern of common descent (Wilson *et al.* 1985) and phylogenies reconstructed using genes from this molecule are gene trees and may not reflect interactions between species. Mitochondrial haplotypes from northwestern populations of grey treefrogs did not identify a female ancestor for the tetraploids and suggest that populations of tetraploids from the northwest are more similar to each other than to any of the diploids, exept for the Wisconsin tetraploids. The sequence identity of the diploid and tetraploid sequences from Wisconsin would therefore be quite interesting if it were a reflection of actual differences because it would provide evidence that this tetraploid had probably experienced a very recent origin that was separate from the other tetraploids. However, because of the possibility of cross contamination or misidentification of the ploidy of the samples (I sequenced toes collected from calling males in the field) this result requires further investigation. Unfortunately, this population only had a few males calling during the time of collection and only a single diploid and tetraploid were collected. Therefore, the possibility of recent origins will not be proposed at this time until confirmation of this result has been obtained.

In all cases except Wisconsin populations, mitochondrial haplotypes of tetraploids appear to be more closely related to each other than to any of the diploids. In fact, in most of the sympatric populations, the tetraploids appear to be more distantly related to the diploid with which they are sympatric than to diploids from

other parts of the range. This supports the findings of Ptacek *et al.* (1994) that tetraploids tended to be most distantly related to diploids with which they were sympatric (Ptacek *et al.* 1994). Specifically, eastern tetraploids appeared to be most closely related to western diploids while southwestern tetraploids tended to be most closely related to eastern diploids. This study also suggested that there were at least three independent lineages of tetraploids, and that these lineages were more closely related to either the eastern or western diploid lineage than they were to each other (Ptacek *et al.* 1994). Although only a single tetraploid from another part of the range was included in my study (i.e., the tetraploid from Bastrop Co., Texas), my data did not suggest multiple origins of the tetraploids. Sampling of populations included in this study may not have been extensive enough to include a diploid that was more closely related to the Bastrop tetraploid than were the northern tetraploids. Although there appeared to be a distinction between diploids from Minnesota and the rest of the diploids that could be interpreted as "eastern" and "western", the Travis Co., Texas population was included the "western diploid" clade in the study of Ptacek *et al.* (1994). Therefore, it is likely that all diploids used here represented western diploid populations, that were not closely related to tetraploids from Bastrop in their analyses (Ptacek *et al.* 1994). It is unfortunate that sequences could not be obtained from the samples available from populations that had been described as "eastern", as comparisons with allozyme results might be more directly comparable. Therefore, the mtDNA portion of my study does not contribute much new information except that a potential diploid progenitor for the northern populations has still not been identified and that diploids from Wisconsin and Michigan may belong to the "western" diploid lineage.

Evolutionary Implications

Previous studies have assumed that the northern movement of tetraploids was restricted or driven by glacial retreat. It seems equally likely that tetraploids in this region outcompeted their diploid progenitors in a time when environmental upheaval was common. This could explain the lack of identification of a diploid progenitor in the northern extremes of the range. Salemma (1984) suggested that new tetraploid species should do best in early succession environments in which competition is reduced but may be more adaptable than diploids because of their increased capacity for heterozygosity and genetic flexibility. There is some mainly anecdotal evidence that tetraploids from Ontario have higher freeze tolerance than diploids from any region and from tetraploids from Texas (personal observation). If this is true, it may explain how tetraploids isolated north of the Great Lakes were able to exploit an environment uninhabitable by their diploid progenitors. The similarity of mtDNA in tetraploids from north and south of the Great Lakes suggests that they had a common origin.

Mating calls of diploids have been shown to display differences in call duration that appear to vary in accordance with the predicted boundaries of separation of "eastern" and "western" lineages (Gerhardt 1978). Data on variation in mating calls from northwestern tetraploids have not been collected. Although pulse rate is the major parameter which females use to discriminate calls of conspecifics (Doherty and Gerhardt 1984), a survey of mating call variation based on predicted lineage divisions from mitochondrial DNA studies would help to test the hypothesis that tetraploid calls change automatically with the change in cell size (Bogart and

Wasserman 1972; Bogart 1980). If northwestern tetraploids were found to have calls that were similar to those from independently derived lineages, support would be gained for the idea that polyploidy can result in automatic isolation (if a corresponding change also occurred in the female auditory system). It would also be interesting to perform hybridization experiments between tetraploids from different lineages and between tetraploids and diploids from different parts of the ranges to determine if hybrid sterility occurs in all combinations. No differences in hybrid success have been found using tetraploid females from Southern Ontario and diploid males from Maryland, Virginia and Texas (Mable 1989; Bogart, personal communication). Hybridization experiments have not been described using diploid males from the more western parts of the ranges or using tetraploid females from across the proposed independent lineage boundaries.

CONCLUSIONS

This study did not add radically new information to the understanding of geographic distribution of lineages of diploid and tetraploid grey treefrogs across North America. However, synthesis of the information presented by previous researchers and inclusion of northwestern populations that have not been previously described does provide a bit clearer picture of the complexity of this complex (Figure 2.8). One conclusion that seems to be clearly reinforced is that if a lineage based naming system were used in the taxonomy of this group, the two names currently applied would not be sufficient. Rather than consisting of a single diploid (*H. chrysoscelis*) and a single tetraploid (*H. versicolor*) species, the findings of Ptacek *et al.* (1994) suggested that the group consists of at least two diploid lineages and that

tetraploids arose independently 3 times. My results suggest that this may have been an underestimation. In the allozyme analyses, while populations in the more central parts of the range (e.g., Illinois, eastern Texas, Wisconsin) did tend to fluctuate across analyses, it could be concluded that there are more than two groupings of diploids and tetraploids: an eastern group from the eastern coastal region (e.g., Virginia, Maryland); a western group from Minnesota and Manitoba that may or may not be the same as western populations further south (down through Texas); and a north/central and/or northeastern group (Ontario, Michigan, Ohio, New York). My mtDNA study was very limited but also suggested that a northern grouping of western diploids might be distinctive from the more southern diploids in the west (Figure 2.8). While it might be tempting to assign names (possibly based on historical subspecies names) to mitochondrial lineages, this does not seem justified before more extensive comparisons are made between phylogenies based on both nuclear and mitochondrial genomes from across the entire range of the grey treefrog complex.

Table 2.1. Localities of populations used in allozyme and/or mtDNA analysis of grey treefrogs.

Sample	Ploidy	Locality	Condition	Source	N	Allozymes	mtDNA
1	2n	Walworth, Wisconsin	Allopatric	Field	11	2nWisconsin1	
2	4n	Columbia, Wisconsin	Parapatric	Field	4	4nWisconsin2	
3	2n	Columbia, Wisconsin	Parapatric	Field	7	2nwisconsin2	
4	2n	Mahnomen, Minnesota	Sympatric	Field	3	2nMinnesota	2nMinnesota
5	4n	Mahnomen, Minnesota	Sympatric	Field	46	4nMinnesota	4nMinnesota
6	2n	Whiteshell, Manitoba	Sympatric	Field	12	2nManitoba	
7	4n	Whiteshell, Manitoba	Sympatric	Field	30	4nManitoba	
8	2n	Prince Georges, Virginia	Sympatric	Field	24	2nVirginia	
9	4n	Prince Georges, Virginia	Sympatric	Field	25	4nVirginia	
10	2n	Shady Side, Maryland	Allopatric	Field	22	2nMaryland	
11	4n	Aberfoyle, Ontario	Allopatric	Field	31	4nSOntario	4nSOntario
12	2n	West Virginia	Sympatric	Freezer	4	2nWVirginia	
13	4n	West Virginia	Sympatric	Freezer	6	4nWVirginia	
14	4n	Ottawa, Ontario	Allopatric	Freezer	35	4nEOntario	
15	4n	Guelph, Ontario	Allopatric	Both	25	4nCOntario	
16	4n	Muskoka, Ontario	Allopatric	Both	9	4nNOntario	
17	4n	Michigan	Allopatric	Freezer	14	4nMichigan1	
17	2n	Jackson Co., Michigan	Sympatric	Field	1		2nMichigan
17	4n	Jackson Co., Michigan	Sympatric	Field	14	4nMichigan2	4nMichigan
18	2n	Travis Co., Texas	Allopatric	Field	1		2nTravis,TX
19	4n	Bastrop Co., Texas	Sympatric?	Field	1		4nBastrop,TX
21	2n	Lamar Co., Texas	Allopatric	Field	1		2nLamar,TX
22	4n	Rainy River, Ontario	Allopatric	Field		4nNWOntario	
23	2n	Juneau Co., Wisconsin	Sympatric	Field	1		2nWisconsin
24	4n	Juneau Co., Wisconsin	Sympatric	Field	1		4nWisconsin

Table 2.2. Enzyme systems and buffers used in a study of allozymes of northern populations of grey treefrogs.

Electromorph Designation	Allele Label[d]	Buffer System	pH	Electromorph Designation	Allele Label[d]	Buffer System	pH
6Pgd^{100}[a]	A	Poulik[b]	8.2	Mdh-2^{75}	B	Tris Citrate[b]	6.5
6Pgd^{85}	B			Mpi-2^{100}	A	Tris Citrate[b]	6.5
Acon-1^{100}	A	Tris Citrate[b]	6.5	Mpi-2^{90}	B		
Acon-1^{85}	B			Mpi-2^{80}	C		
Acon-1^{75}	C			Mpi-2^{70}	D		
Acon-2^{100}	A			Pgi-1^{100}	A	Amine Citrate[c]	6.5
Acon-2^{50}	B			Pgi-1^{90}	B		
Acon-2^{20}	C			Pgi-2^{100}	A		
Got-1^{100}	A	Tris Citrate[b]	6.5	Pgi-2^{80}	B		
Got-1^{48}	B			Pgi-2^{40}	C		
Got-1^{30}	C			Pgm-1^{100}	A	Amine Citrate[c]	6.5
Got-2^{100}	A	Tris Citrate[b]	6.5	Pgm-1^{90}	B		
Got-2^{75}	B			Pgm-2^{100}	A		
Idh-1^{100}	A	Tris Citrate[b]	6.5	Pgm-2^{80}	B		
Idh-1^{90}	B			Sod-1^{100}	A	Amine Citrate[c]	6.5
Idh-1^{78}	C			Sod-1^{90}	B		
Idh-2^{100}	A	Tris Citrate[b]	6.5	Sod-1^{60}	C		
Idh-2^{90}	B			Sod-1^{40}	D		
Idh-2^{80}	C			Sod-1^{20}	E		
Mdh-1^{100}	A	Amine-Citrate[c]	6.5	aGpd^{100}	A	Poulik[b]	8.2
Mdh-1^{52}	B			aGpd^{90}	B		
Mdh-2^{100}	A	Amine-Citrate[c]	6.5	aGpd^{75}	C		

[a] Mobilities of electromorphs relative to the furthest migrating *Hyla versicolor* allele for each enzyme

[b] Selander et al. 1971

[c] Clayton and Tretiak 1972

[d] Allele label refers to letter assigned to each electromorph

Table 2.5. Table of distances for alloyzme loci used to screen populations of grey treefrogs. Roger's Genetic Distance is above the diagonal, Nei's Genetic Distance is below the diagonal, and Average Heterozygosity is on the diagonal. Distance measures were computed using diagonal BIOSYS-1 (Swofford and Selander 1981) and GENEPAC (J.P. Bogart, unpublished).

Population	Wisconsin 2na	Wisconsin 4nb	Wisconsin 2nb	Minnesota 2n	Minnesota 4n	Manitoba 2n	Manitoba 4n	Virginia 2n	Virginia 4n	Maryland 2n	Sontario 4n
N	(11)	(4)	(7)	(3)	(46)	(12)	(30)	(13)	(25)	(22)	(31)
Wisconsin.2na	(0.095)	.078	.055	.133	.085	.112	.109	.075	.061	.099	.096
Wisconsin.4nb	.012	(0.079)	.114	.159	.110	.121	.112	.094	.076	.129	.085
Wisconsin.2nb	.019	.002	(0.131)	.115	.094	.108	.110	.086	.104	.109	.103
Minnesota.2n	.035	.049	.020	(0.190)	.130	.091	.101	.142	.126	.165	.168
Minnesota.4n	.007	.005	.008	.021	(0.138)	.091	.083	.109	.105	.107	.119
Manitoba.2n	.026	.021	.027	.007	.009	(0.170)	.052	.134	.116	.152	.144
Manitoba.4n	.026	.019	.029	.019	.008	.000	(0.173)	.126	.105	.137	.144
Virginia.2n	.007	.009	.013	.037	.007	.028	.022	(0.102)	.064	.078	.089
Virginia.4n	.006	.012	.019	.031	.013	.025	.026	.011	(0.139)	.100	.102
Maryland.2n	.018	.035	.024	.060	.025	.047	.038	.043		(0.173)	.120
SOntario.2n	.029	.016	.029	.061	.025	.050	.050	.028	.031	.043	(0.140)

Table 2.4. Summary of allele frequencies at allozymes used in a survey of northern populations of grey treefrogs.

(a) Field collected samples.

Locus	Alleles	wisconsin.2n	wisconsin.4n	wisconsin.2n	minnesota.4n	minnesota.2n	manitoba.4n	manitoba.2n	virginia.4n	virginia.2n	maryland.2n	aberfoyle.4n
	(N)	(11)	(4)	(7)	(46)	(3)	(30)	(12)	(25)	(24)	(22)	(31)
6Pgd	A	0.5	0.938	0.429	0.717	0.333	0.858	0.75	0.51	0.542	0.568	0.581
	B	0.5	0.062	0.571	0.283	0.667	0.142	0.25	0.49	0.458	0.432	0.419
Acon-1	A	0	0	0	0.098	0.167	0.108	0.0125	0.02	0.021	0	0.024
	B	1	1	1	0.902	0.833	0.892	0.875	0.9	0.979	1	0.96
	C	0	0	0	0	0	0	0	0.08	0	0	0.016
Acon-2	A	0.727	1	0.571	0.755	0.667	0.638	0.75	0.94	0.938	0.545	1
	B	0.273	0	0.143	0.092	0	0.129	0.125	0.006	0	0.455	0
	C	0	0	0.286	0.153	0.333	0.233	0.125	0	0.062	0	0
Got-1	A	0	0.25	0	0	0	0	0	0	0	0.045	0.5
	B	1	0.75	1	0.062	1	1	1	1	1	0.955	0.476
	C	0	0	0	0.938	0	0	0	0	0	0	0.024
Got-2	A	0.909	0.938	1	0.984	1	0.967	1	0.94	0.979	1	1
	B	0.091	0.062	0	0.016	0	0.033	0	0.06	0.021	0	0
Idh-1	A	0	0	0	0.016	0	0.108	0	0.03	0	0	0.016
	B	1	1	1	0.929	1	892	1	0.93	0.896	0.864	0.976
	C	0	0	0	0.022	0	0	0	0.04	0.104	0.136	0.008
	D	0	0	0	0.033	0	0	0	0	0	0	0
Idh-2	A	0	0	0	0.049	0	0	0	0.08	0	0	0
	B	0	0	0	0.109	0.333	0.362	0.25	0.92	0.104	0.182	0.048
	C	1	1	1	0.842	0.667	0.638	0.75	0	0.896	0.818	0.952
Mdh-1	A	0.955	0.938	0.714	0.766	0.5	0.792	0.667	1	1	1	0.919
	B	0.045	0.062	0.286	0.234	0.5	0.208	0.333	0	0	0	0.081
Mdh-2	A	1	1	1	0.995	1	1	1	1	1	1	1
	B	0	0	0	0.005	0	0	0	0	0	0	0
Mpi-2	A	0	0	0.071	0.031	0.666	0.125	0.208	0	0.02	0	0
	B	0.778	0.687	0.858	0.74	0.167	2514	0.417	0.56	0.938	0.909	0.79
	C	0.222	0.313	0.071	0.229	0.167	0.361	0.375	0.41	0.042	0.091	0.113
	D	0	0	0	0	0	0	0	0.03	0	0	0.097
Pgi-1	A	0	0	0	0.005	0	0.033	0	0.05	0	0.455	0
	B	1	1	1	0.995	1	0.967	1	0.95	1	0.545	1
Pgi-2	A	0	0	0	0.092	0.167	0.317	0.417	0.05	0	0.042	0.038
	B	1	1	1	0.886	0.833	0.683	0.583	0.95	1	0.917	0.962
	C	0	0	0	0.022	0	0	0	0	0	0.041	0
Pgm-1	A	1	0.063	0	0.033	0.167	0.005	0.042	0.22	0.25	0.205	0.097
	B	0	0.937	1	0.967	0.833	0.95	0.958	0.78	0.75	0.795	0.903
Pgm-2	A	1	1	0.857	0.951	1	0.967	1	0.98	0.896	0.841	0.584
	B	0	0	0.0143	0.049	0	0.033	0	0.02	0.104	0.159	0.416
Sod-1	A	0.045	0.063	0	0.016	0	0.017	0	0.148	0	0.091	0.04
	B	0	0	0	0	0	0.008	0	0.125	0.063	0.023	0.04
	C	0.818	0.937	0.786	0.984	1	0.975	1	0.727	0.917	0.886	0.92
	D	0.136	0	0.214	0	0	0	0	0	0.02	0	0
	E	0.001	0	0	0	0	0	0	0	0	0	0
aGpd	A	0	0	0.071	0	0	0.025	0.167	0	0	0	0.024
	B	1	1	0.929	1	1	0.975	0.833	1	0.979	0.955	0.927
	C	0	0	0	0	0	0	0	0	0.021	0.045	0.049

(b) Populations obtained from a tissue bank collection.

Locus	Alleles	WVirginia.2n	WVirginia.4n	EOntario.4n	COntario.4n	NCOntario.4n	Michigan1.4n	Michigan2.4n
	(N)	(4)	(6)	(35)	(25)	(9)	(14)	(13)
6Pgd	A	0.750	0.417	0.643	0.580	0.611	0.589	0.269
	B	0.250	0.583	0.357	0.420	0.389	0.411	0.731
Acon-1	A	0.500	0.125	0	0.063	0.972	0	0
	B	0.500	0.875	1.000	0.938	0.028	0.982	0.961
	C	0	0	0	0	0	0.018	0.039
Acon-2	A	0	0	0	0	0	0	0
	B	0	0	0	0	0	0	0
	C	0	0	0	0	0	0	0
Got-1	A	0	0	0	0	0	0	0
	B	0.625	0.583	0.229	0.340	0.389	0.232	0.411
	C	0.375	0.417	0.771	0.660	0.611	0.768	0.589
Got-2	A	1	1	1	1	1	1	1
	B	0	0	0	0	0	0	0
Idh-1	A	1	1	0.893	1	1	0.982	0.929
	B	0	0	0.107	0	0	0.018	0.071
	C	0	0	0	0	0	0	0.071
	D	0	0	0	0	0	0	0
Idh-2	A	0	0	0	0	0	0	0
	B	0	0	0	0	0	0	1
	C	1	1	1	1	1	1	0
Mdh-1	A	1	1	0.950	0.910	0.778	0.732	0.942
	B	0	0	0.050	0.090	0.222	0.268	0.058
Mdh-2	A	1	1	1	1	1	1	1
	B	0	0	0	0	0	0	0
Mpi-2	A	0.625	0.500	0.790	0.770	0.571	0.789	0.750
	B	0.125	0.250	0.129	0.130	0.357	0.115	0.205
	C	0.250	0.250	0.081	0.100	0.072	0.096	0.045
	D	0	0	0	0	0	0	0
Pgi-1	A	0	0	0	0	0	0	0
	B	0	0	0	0	0	0	0
Pgi-2	A	0	0	0	0	0	0	0
	B	0	0	0	0	0	0	0
	C	0	0	0	0	0	0	0
Pgm-1	A	0	0.200	0.250	0.060	0.194	0.071	0.135
	B	1	0.800	0.750	0.940	0.806	0.929	0.865
Pgm-2	A	0.125	0.050	0.287	0.320	0.333	0.268	0.250
	B	0.875	0.950	0.713	0.680	0.667	0.732	0.750
Sod-1	A	0.125	0.000	0.022	0.010	0.031	0.050	0.025
	B	0.500	0.100	0.132	0.094	0.188	0.025	0.000
	C	0.375	0.900	0.816	0.885	0.750	0.875	0.875
	D	0.000	0.000	0.000	0.000	0.031	0.025	0.100
	E	0	0	0.029	0.010	0	0.025	0
aGpd	A	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	B	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	C	0	0	0	0	0	0	0

Table 2.3 Comparison of nucleotide changes in COI

(a) Comparison of nucleotide changes in a 590-bp region of cytochrome oxidase 1 among diploid and tetraploid grey treefrogs. Summary of nucleotide substitutions at each of the 28 variable sites.

Position	15	30	78	96	99	111	114	123	135	159	199	222	235	255	287	295	298	304	319	337	349	364	418	472	514	523	535	571	574	582	588
Minnesota.4n	C	T	G	A	A	T	A	T	C	A	C	C	G	G	G	A	A	G	A	G	T	C	C	G	C	A	T	G	T	G	A
NOntario.4n	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOntario.4n	T	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	G	.	.	A	C	T	.	.	.	.	.	.	.	.	.
Michigan.4n	.	.	.	.	.	.	G	G	.	.	.	.	.	.	.	.	.	.	.	A	.	T	.	.	.	.	.	.	C	.	T
Bastrop.4n	T	.	.	.	.	.	.	.	G	.	.	.	T	.	.	.	.	.	.	A	.	.	.	A	T	.	.	.	C	.	.
Wisconsin.4n	.	C	.	.	.	C	.	.	G	G	.	T	.	A	A	.	.	A	.	A	.	.	T	A	T	G	C	A	C	A	.
Wisconsin.2n	.	C	.	.	.	C	.	.	G	G	.	T	.	A	A	.	.	A	.	A	.	.	T	A	T	G	C	A	C	A	.
Minnesota.2n	T	C	.	.	.	C	.	..	G	G	.	T	.	.	.	.	.	A	G	A	.	.	.	A	T	.	C	.	C	A	.
Michigan.2n	.	C	.	.	.	C	.	.	G	G	.	T	.	A	A	T	.	A	G	A	.	.	T	A	T	G	C	A	C	A	.
Travis.2n	.	C	.	.	.	C	.	.	G	.	.	T	.	A	A	.	.	A	.	A	.	.	T	A	T	G	C	A	C	A	.
Lamar.2n	T	C	A	G	G	C	G	.	G	G	T	T	.	A	A	C	.	A	G	A	.	.	T	A	T	G	C	A	C	A	.

(b) Comparison of nucleotide changes in a 590-bp region of cytochrome oxidase 1 among diploid and tetraploid grey treefrogs. Pairwise comparisons among taxa showing % sequence divergence (above the diagonal) and absolute number of changes (below the diagonal). Bold type indicates the closest relationships within ploidy levels.

	Minnesota 4n	NOntario 4n	SOntario 4n	Michigan 4n	Bastrop 4n	Wisconsin 4n	Wisconsin 2n	Minnesota 2n	Michigan 2n	Travis 2n	Lamar 2n
Minnesota.4n	-	**0.0**	**1.0**	**1.0**	**1.2**	2.9	2.9	3.4	2.4	2.7	4.2
NOntario.4n	0	-	**1.0**	**1.0**	**1.2**	2.9	2.9	3.4	2.4	2.7	4.2
SOntario.4n	6	6	-	**1.0**	**1.5**	3.6	3.6	3.7	3.1	3.4	4.1
Michigan.4n	6	6	6	-	**1.5**	3.2	3.2	3.7	2.9	3.1	4.1
Bastrop.4n	7	7	9	9	-	2.2	2.2	2.5	2.7	2.4	3.2
Wisconsin.4n	17	17	21	19	13	-	0	**0.5**	**1.2**	**0.2**	**1.2**
Wisconsin.2n	17	17	21	19	13	0	-	**0.5**	**1.2**	**0.2**	**1.2**
Minnesota.2n	20	20	22	22	15	3	3	-	**1.4**	**0.7**	**0.9**
Michigan.2n	14	14	18	16	17	7	7	8	-	1.4	2.0
Travis.2n	16	16	20	18	14	1	1	4	8	-	1.5
Lamar.2n	25	25	24	24	19	7	7	5	12	9	-

Figure 2.1. Populations of diploid and tetraploid grey treefrogs used in analyses. Open squares represent diploids and open circles represent tetraploids used in the mtDNA study of Ptacek et al. (1994). Solid squares represent diploids and solid circles represent tetraploids used in the mtDNA and alloyzme study presented in this chapter. Solid upright triangles represent diploids and solid inverted trianges represent tetraploids used in the studies of Romano et al. (1987) and Ralin et al. (1983). Sympatric populations (i.e., where diploids and tetraploids were collected from the same ponds) are indicated by overlapping symbols and parapatric populations (i.e., where diploids and tetraploids were collected from different ponds in the same region) are indicated by adjacent symbols.

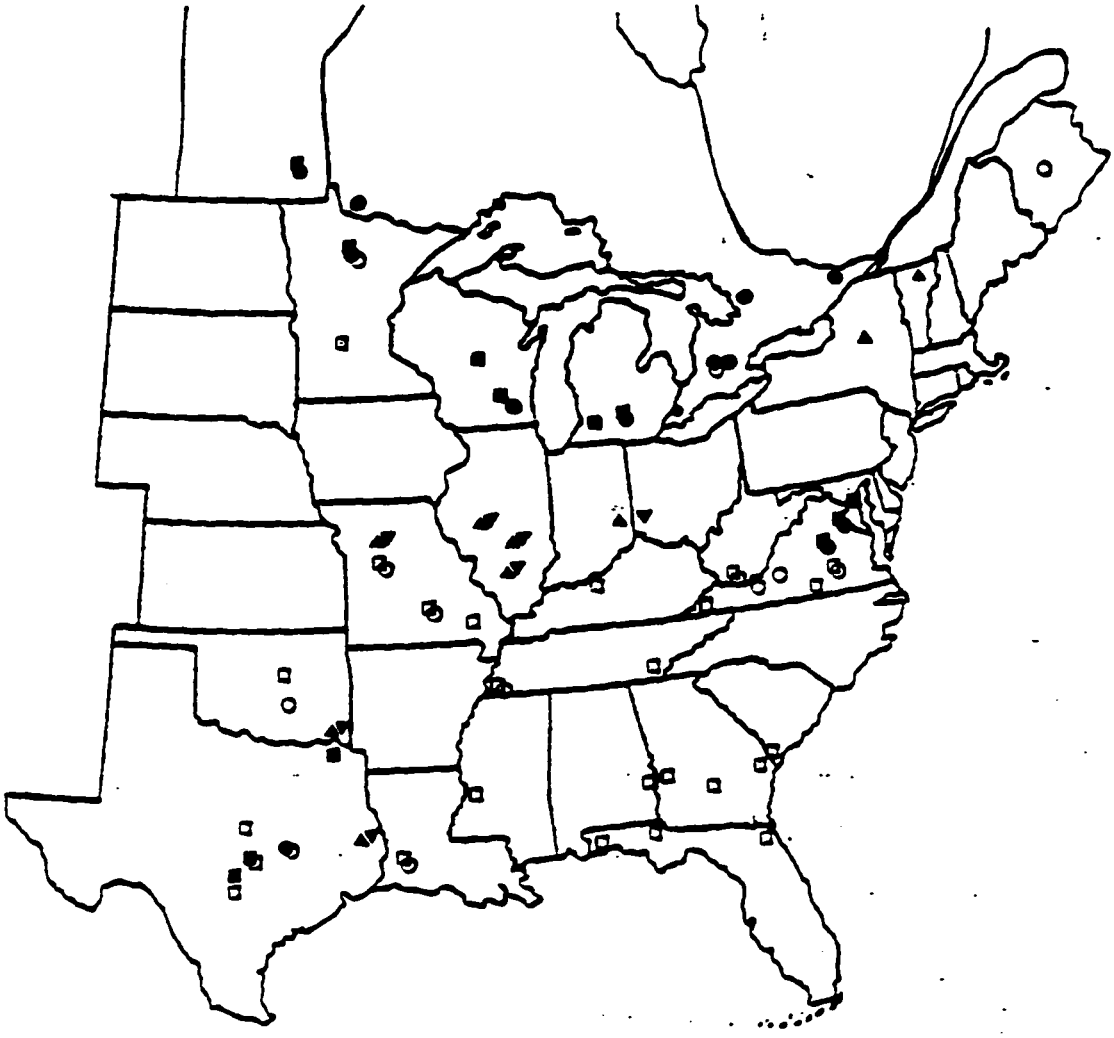
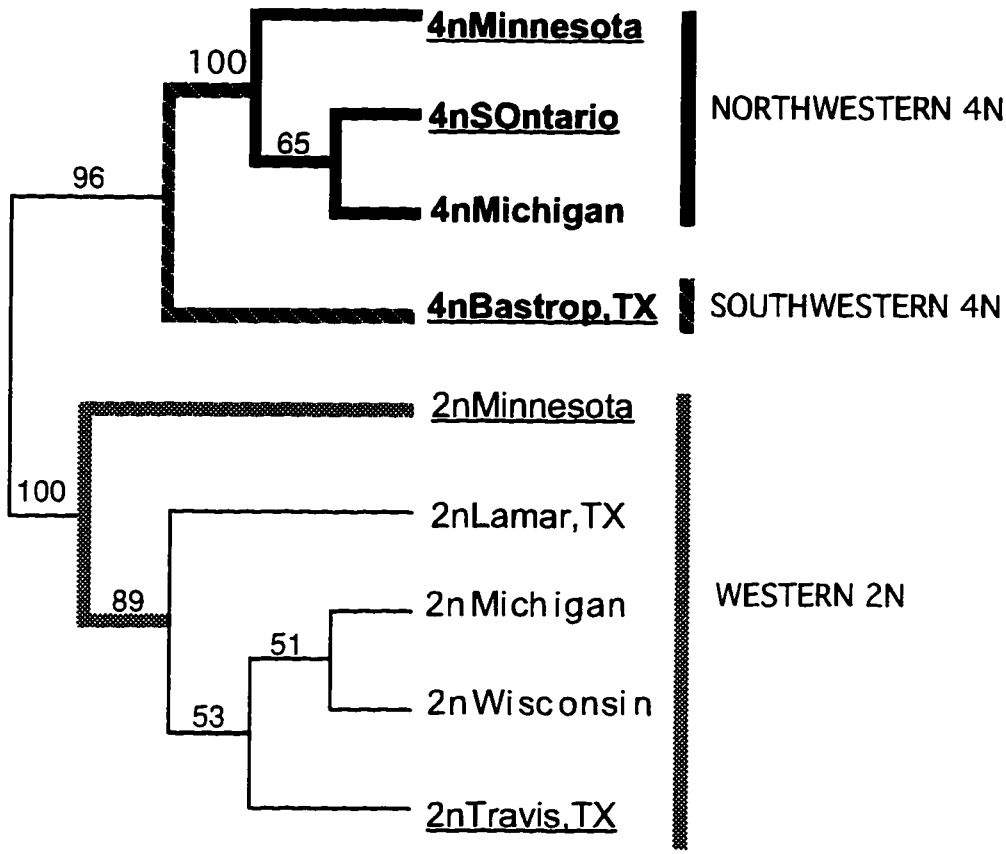


Figure 2.2. Phylogenetic reconstructions using a 590-bp region of cytochrome oxidase I for populations of diploid and tetraploid grey treefrogs. (a) Midpoint rooting tree of the single-most parsimonious trees found using parsimony with a transition:transversion weighting scheme of 7:1. The same topology was found using parsimony with any character weighting scheme and using either minimum evolution or maximum parsimony optimality criteria. (b) Unrooted phylogram of the tree shown in (a). Tetraploid taxa are indicated in bold type. Populations that were also used in the cytochrome b phylogeny presented in Ptacek *et al.* (1994) based on populations are underlined.

a)


4nMinnesota
100
4nSOntario
65
NORTHWESTERN 4N
96
4nMichigan
4nBastrop,TX
SOUTHWESTERN 4N
2nMinnesota
100
2nLamar,TX
89
2nMichigan
WESTERN 2N
51
2nWisconsin
53
2nTravis,TX

b)

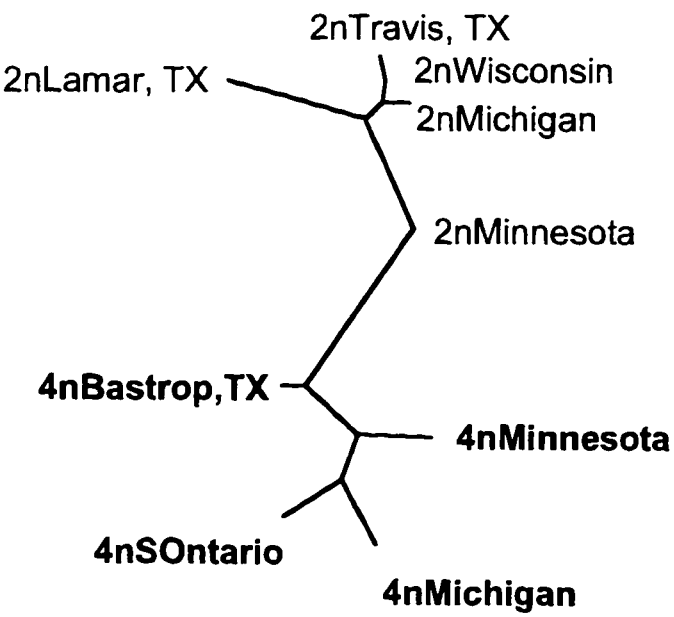

2nTravis, TX
2nLamar, TX
2nWisconsin
2nMichigan
2nMinnesota
4nBastrop,TX
4nMinnesota
4nSOntario
4nMichigan

Figure 2.3. Phylogenetic reconstructions using allozyme frequency data for northern populations of grey treefrogs collected from a field survey. (a) UPGMA tree using Rogers" Genetic distances (BIOSYS-1: Swofford and Selander 1991). (b) Phylogram of the single-most parsimonious tree found using locus specific stepmatrices determined from pairwise Manhattan distances of allozyme allele frequencies across loci. Bootstrap replications (100) did not result in much resolution but values for nodes which were supported are given. Note that both methods suggest groupings more by geographic region than by ploidy level. The main difference between the two analyses was in the relative placement of the Wisconsin diploid populations.

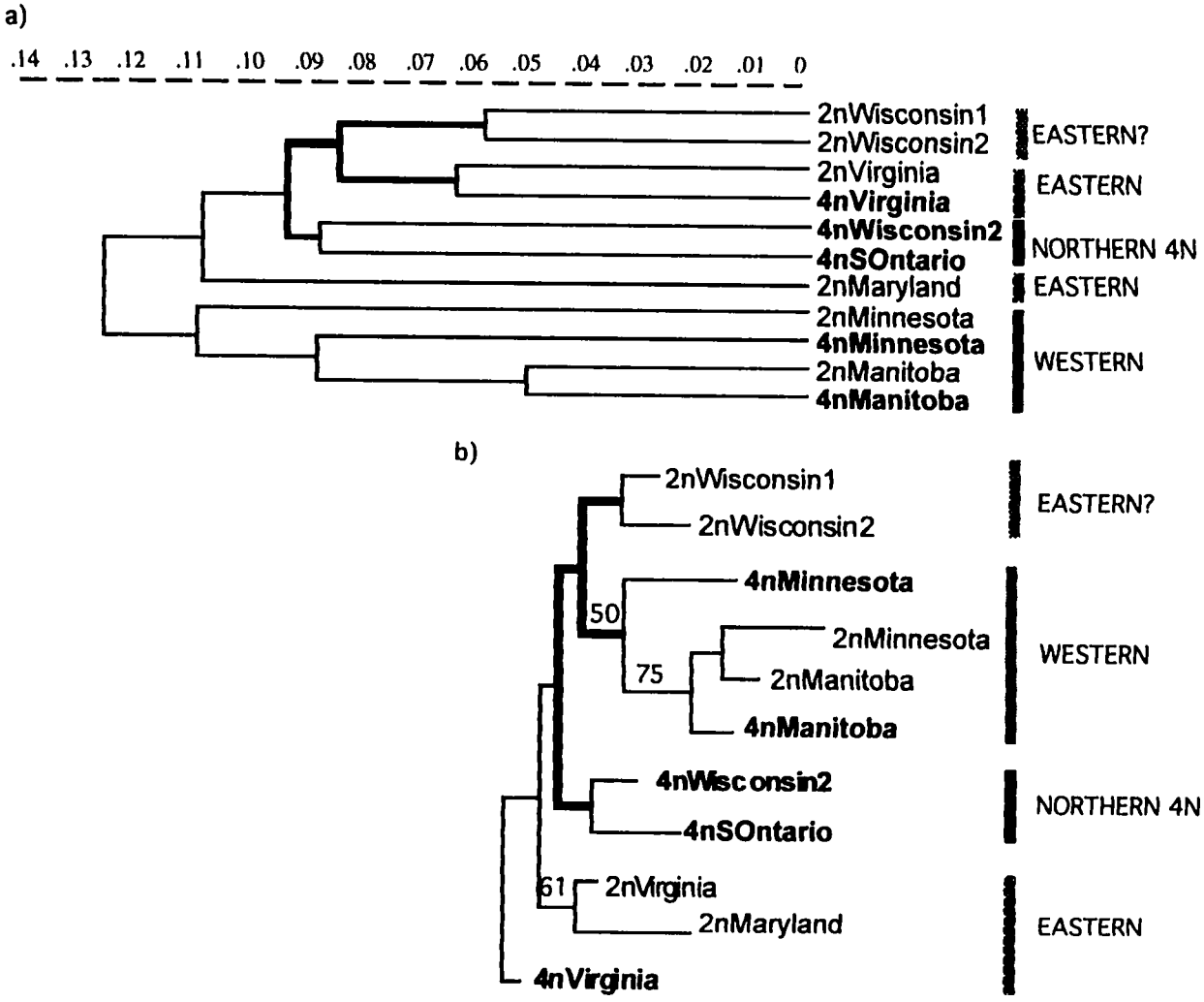

a)
.14 .13 .12 .11 .10 .09 .08 .07 .06 .05 .04 .03 .02 .01 0
2nWisconsin1
2nWisconsin2
2nVirginia
4nVirginia
4nWisconsin2
4nSOntario
2nMaryland
2nMinnesota
4nMinnesota
2nManitoba
4nManitoba
EASTERN?
EASTERN
NORTHERN 4N
EASTERN
WESTERN
b)
2nWisconsin1
2nWisconsin2
4nMinnesota
50
2nMinnesota
75
2nManitoba
4nManitoba
4nWisconsin2
4nSOntario
61
2nVirginia
2nMaryland
4nVirginia
EASTERN?
WESTERN
NORTHERN 4N
EASTERN

Figure 2.4. Phylogenetic reconstruction using parsimony and locus specific stepmatrices determined from pairwise Manhattan distances of allozyme allele frequencies across loci for all populations of grey treefrogs sampled, including those from a tissue bank collection. Phylogram of the single-most parsimonious tree found.

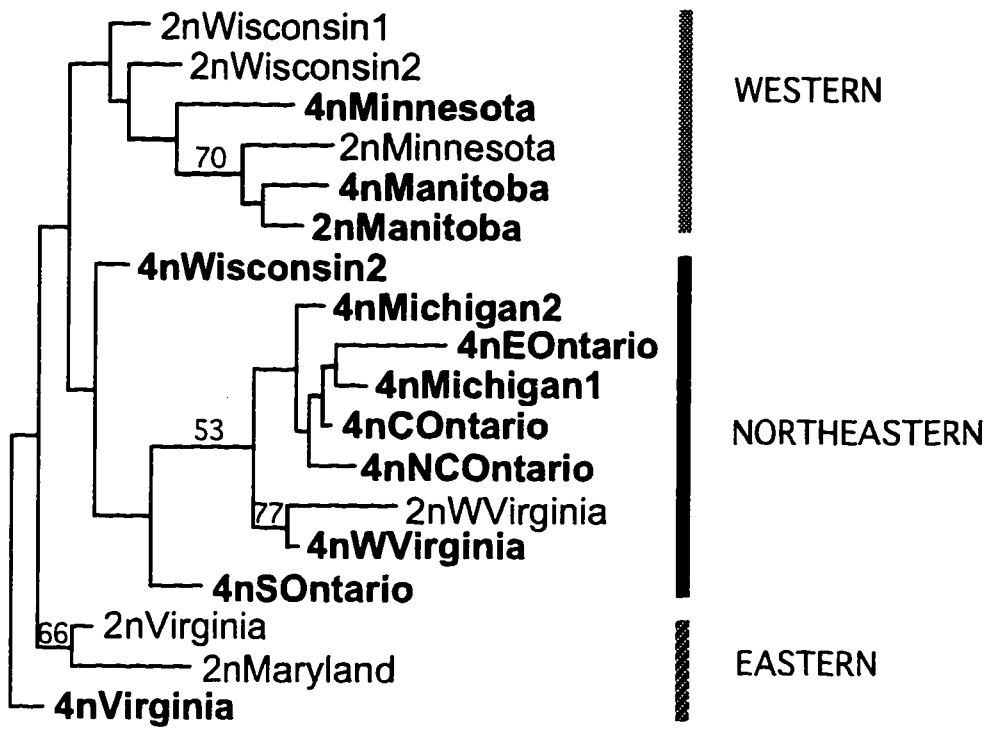
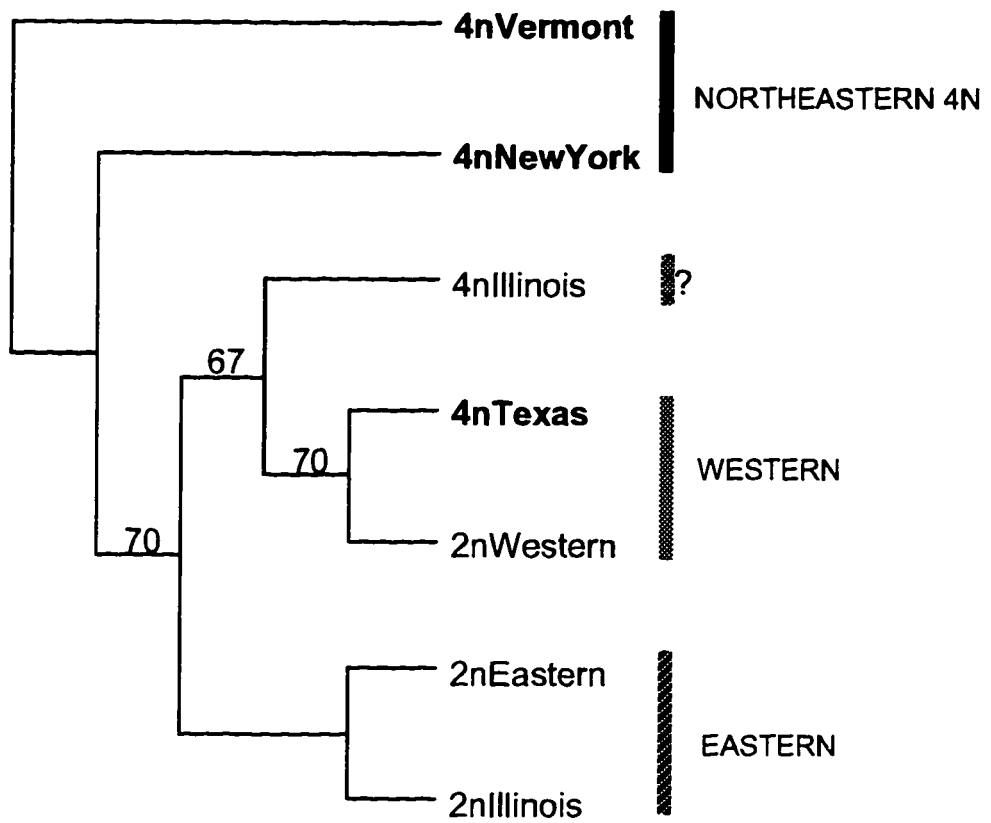

2nWisconsin1
2nWisconsin2
4nMinnesota
70
2nMinnesota
4nManitoba
2nManitoba
4nWisconsin2
4nMichigan2
4nEOntario
4nMichigan1
53
4nCOntario
4nNCOntario
77
2nWVirginia
4nWVirginia
4nSOntario
66
2nVirginia
2nMaryland
4nVirginia
WESTERN
NORTHEASTERN
EASTERN

Figure 2.5. Reanalysis of allozyme data of Ralin *et al.* (1983) using the parsimony approach with Manhattan distances in locus-specific stepmatrices. (a) Phylogram of the single most parsimonious tree found. (b) Unrooted phylogram of the tree shown in (a).

a)


4nVermont
NORTHEASTERN 4N
4nNewYork
4nIllinois
?
67
4nTexas
70
WESTERN
70
2nWestern
2nEastern
EASTERN
2nIllinois

b)

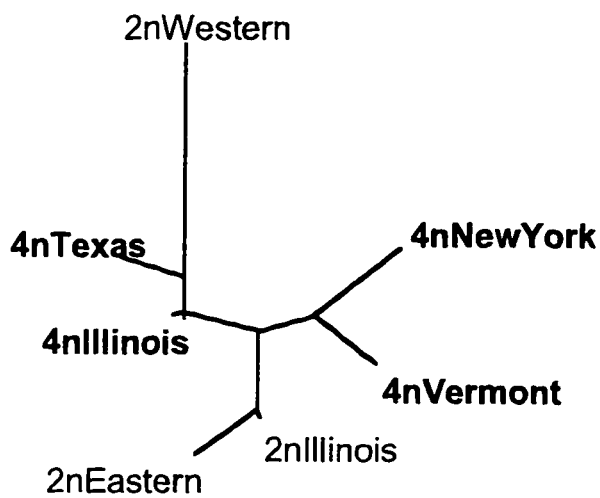
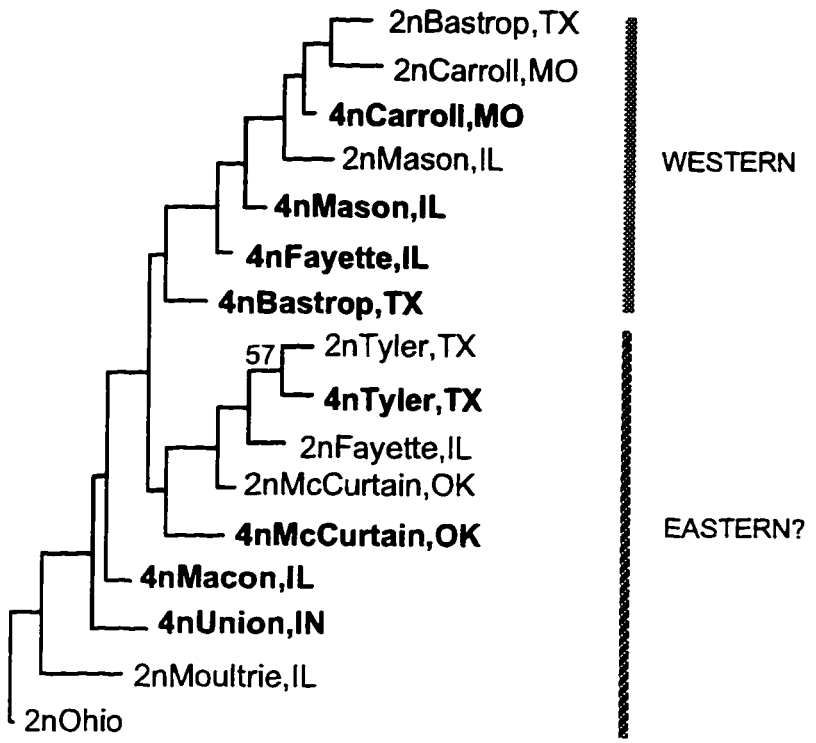

2nWestern
4nTexas
4nNewYork
4nIllinois
4nVermont
2nIllinois
2nEastern

Figure 2.6. Reanalysis of allozyme data of Romano *et al.* (1987) using the parsimony approach with Manhattan distances in locus-specific stepmatrices. (a) Representative phylogram of one of five equally parsimonious trees found in a heuristic search. (b) Strict consensus of the five equally parsimonious trees.

a)


2nBastrop,TX
2nCarroll,MO
4nCarroll,MO
2nMason,IL
4nMason,IL
4nFayette,IL
4nBastrop,TX
57
2nTyler,TX
4nTyler,TX
2nFayette,IL
2nMcCurtain,OK
4nMcCurtain,OK
4nMacon,IL
4nUnion,IN
2nMoultrie,IL
2nOhio
WESTERN
EASTERN?

b)

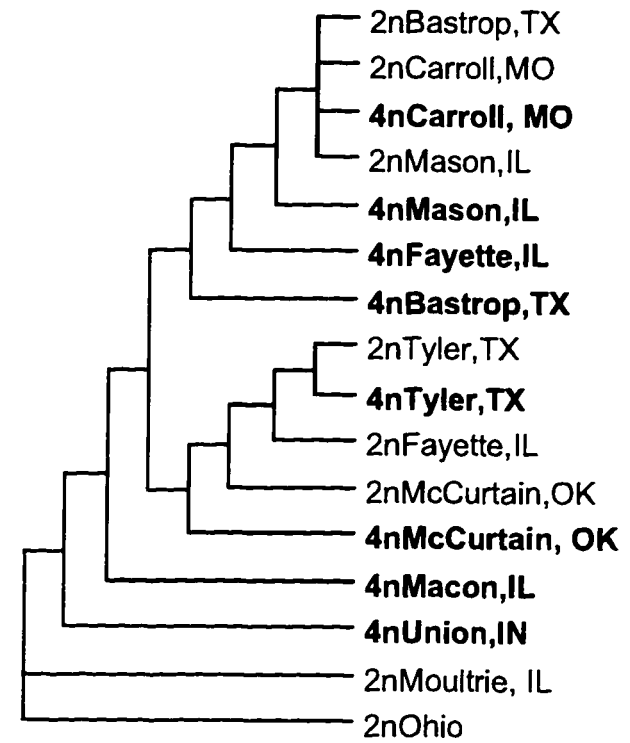
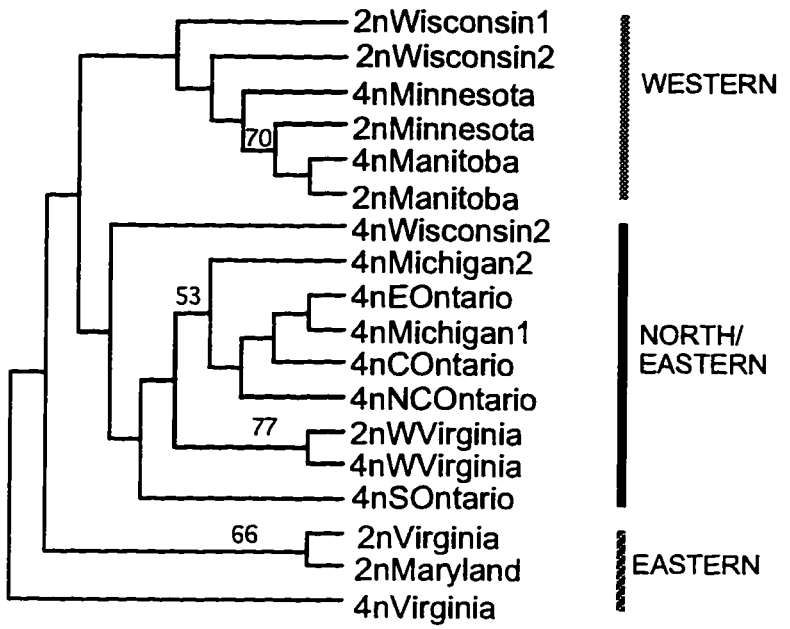

2nBastrop,TX
2nCarroll,MO
4nCarroll, MO
2nMason,IL
4nMason,IL
4nFayette,IL
4nBastrop,TX
2nTyler,TX
4nTyler,TX
2nFayette,IL
2nMcCurtain,OK
4nMcCurtain, OK
4nMacon,IL
4nUnion,IN
2nMoultrie, IL
2nOhio

Figure 2.7. Summary of allozyme analyses. (a) Cladogram of single-most parsimonious tree found in the allozyme survey of northern populations. Note that the three major clades can be grouped into "western", "eastern" and "north/eastern" divisions. (b) Combined analysis of data found in Romano *et al.* (1987) and Ralin *et al.* (1983). The authors included different loci in these analyses and appeared to change the method of scoring used so results may not be completely comparable but the "Illinois" populations used in Ralin *et al.* (1983) fell out with the corresponding populations (Moultrie and Macon) from Romano *et al.* (1987), suggesting that the two datasets were at least somewhat comparable. For the Ralin *et al.* (1983) data, the frequency data summed over populations in Texas tetraploids and Eastern and Western diploids were removed from the analysis because they represented pooled populations. Although this tree was not well resolved, a basic grouping into "western", "eastern" and "north/eastern" clades was evident. Bootstrap analysis collapsed all nodes except the clade including diploids and tetraploids from Tyler, Texas.

a)


2nWisconsin1
2nWisconsin2
4nMinnesota
2nMinnesota
4nManitoba
2nManitoba
4nWisconsin2
4nMichigan2
4nEOntario
4nMichigan1
4nCOntario
4nNCOntario
2nWVirginia
4nWVirginia
4nSOntario
2nVirginia
2nMaryland
4nVirginia
70
53
77
66
WESTERN
NORTH/
EASTERN
EASTERN

b)

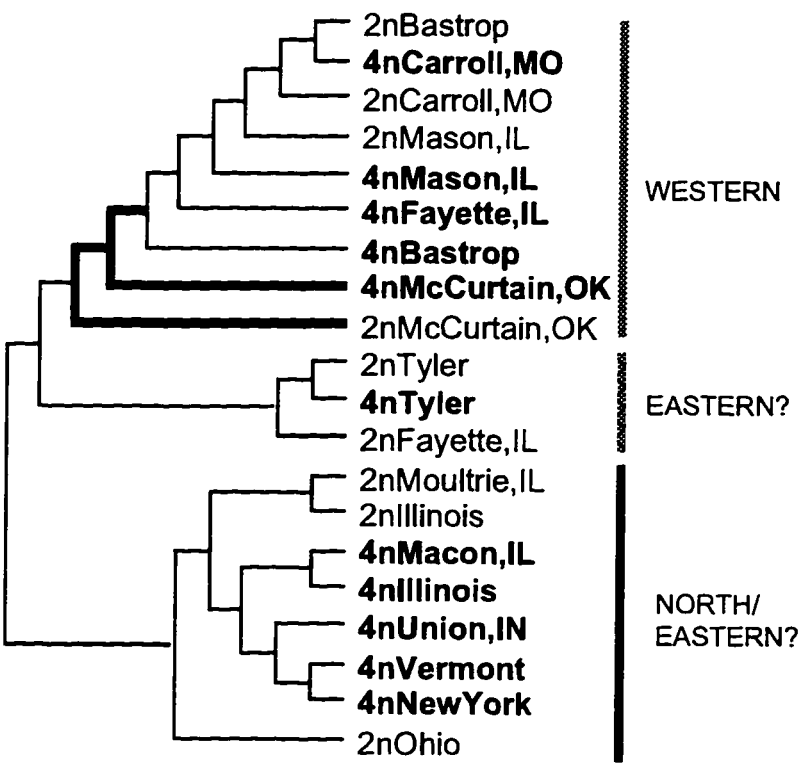
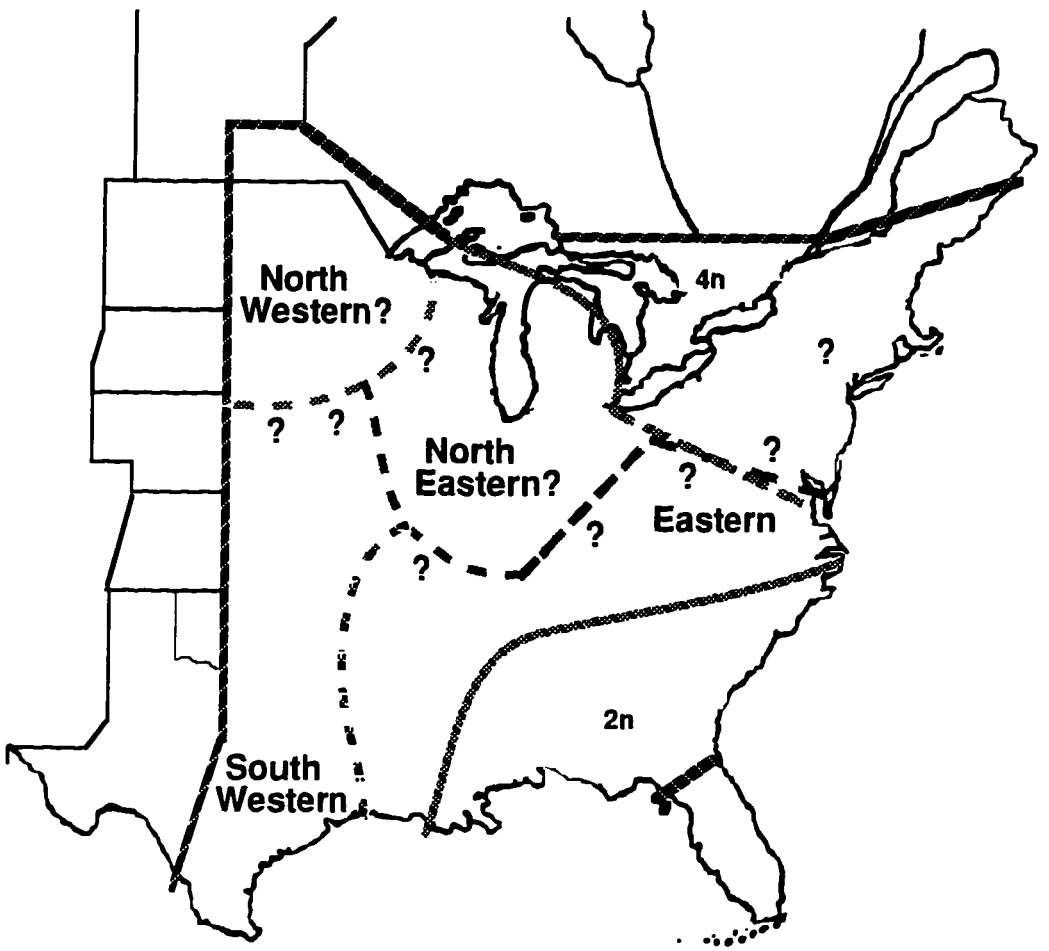

2nBastrop
4nCarroll,MO
2nCarroll,MO
2nMason,IL
4nMason,IL
4nFayette,IL
4nBastrop
4nMcCurtain,OK
2nMcCurtain,OK
2nTyler
4nTyler
2nFayette,IL
2nMoultrie,IL
2nIllinois
4nMacon,IL
4nIllinois
4nUnion,IN
4nVermont
4nNewYork
2nOhio
WESTERN
EASTERN?
NORTH/
EASTERN?

Figure 2.8 Tentative boundaries for lineages of grey treefrogs. Dark hatched lines indicate boundaries of the range of both diploids and tetraploids. Light hatched lines indicate the range of the diploids. Notice that only tetraploids are found in the northeastern part of the range and only diploids are found in the southeast. Previous data has suggested that diploids and tetraploids can each be divided into eastern and western types, with an additional tetraploid type in the northwest. Synthesis of additional information suggests that there is probably a fourth type of tetraploids in the more northeastern region and diploids in this region may also be distinctive. Diploids in the northwestern region may also be distinguishable from those in the southwest. However, insufficient information exists to identify these boundaries definitively. Additional sampling, especially in the northeastern region, is required.


North
Western?
4n
?
?
?
?
North
Eastern?
?
?
?
Eastern
?
2n
South
Western

Chapter 3

Mitochondrial DNA Evolution of Tetraploids in the Genus *Neobatrachus* (Anura: Myobatrachidae)

Table of Contents

ABSTRACT

The Australian burrowing frog genus *Neobatrachus* currently contains six diploid and four tetraploid species recognized through a combination of morphological, cytogenetic, or behavioural characters. Sequences from a coding region of the mitochondrial genome (cytochrome oxidase I) were used to reconstruct the historical relationships between diploids and tetraploids in this genus. Data were analyzed using parsimony, minimum evolution and maximum likelihood as optimality criteria using PAUP*. All analyses (including bootstrap analyses) supported the conclusion that there have been at least two origins of tetraploids in the genus *Neobatrachus*: one involving the "eastern" tetraploids (*N. sudelli* and *N. centralis*) and the other involving the more western tetraploids (*N. aquilonius* and *N. kunapalari*). The diploid *N. fulvus* was identified as the closest relative to the western tetraploids but relationships of the eastern tetraploids to extant diploids were not resolved. Minimum evolution and maximum likelihood appeared to be robust to the model of evolution and both predicted the same branching relationships within and among ploidy levels. Parsimony was found to be sensitive to the model of evolution used; different weighting schemes generated trees that varied in the relationships of the diploid taxa to each other and to the tetraploids, but showed relatively consistent relationships among the tetraploid taxa . A parametric bootstrap approach was used to assess whether clades that did not vary across analyses could be statistically supported when specific hypotheses about branching relationships were tested. It was concluded that the tetraploid lineages may be older than previously suggested, that there have been multiple origins of tetraploidy in the genus, and that some of the tetraploids may have speciated as tetraploids.

INTRODUCTION

Among vertebrates, bisexual polyploids are found in fish in the families Salmonidae and Catostomidae (Ohno *et al.* 1968), and in anurans of the families Pipidae, Leptodactylidae, Myobatrachidae, Hylidae, Bufonidae, Ranidae (reviewed by Bogart 1980; Kawamura 1984; Kuramoto 1990) and Microhylidae (Kuromoto and Allison 1989). Many of the more ancient polyploids in these groups (salmonids, catastomids, pipids) have undergone a functional rediploidization of their genomes followed by a subsequent silencing of duplicate gene expression and no longer show clear relationships to diploid ancestors (e.g. Bailey *et al.* 1978; Kobel and Pasquier 1980). Among the more recent polyploids (genus *Odontophrynus*, *Hyla*, *Tomopterna*) the degree of duplicate gene expression varies (e.g., Danzmann and Bogart 1982a,b; Danzmann and Bogart 1983), and many are found in direct association with closely related diploids. These younger groups provide more potential for establishing how polyploidy has arisen and has been maintained.

The grey treefrogs (family Hylidae) of eastern North America have been studied the most extensively but the very recent origin of tetraploidy in this group (Pleistocene glaciations) and the high degree of crypticity in the complex may make the identification of independently evolving units too difficult to establish, thereby obscuring origins. The Australian burrowing frog genus *Neobatrachus* (family Myobatrachidae) has not been studied as extensively. To date, six diploid and four tetraploid species (Mahony and Robinson 1980; Mahony 1986; Mahony and Roberts 1986; Roberts *et al.* 1991) have been named. Several of the species are morphologically similar but can be separated by a combination of behavioural, cytogenetic, geographic and morphological features. The existence of multiple

recognizable polyploid species within the same recently evolved complex provides more potential for studying the origin of tetraploids and associated isolating mechansims. Diploids and tetraploids are found across the arid and semi-arid zones of Australia, with the highest number of species in Western Australia (Figure 3.1). Tetraploids are found sympatrically with many of the diploid species and occasional hybridization does occur between both diploid and tetraploid taxa (Mahony 1986; Mahony, personal communication). The formation of tetravalents at meiosis I in the tetraploids suggests that they have not undergone rediploidization of their genome (Mahony and Robinson 1980; Mahony 1986) and immunological distance data suggest that speciation within the genus has been relatively recent, at least compared to that within closely related genera such as *Limnodynastes* and *Heleioporus* (Roberts *et al.* 1996).

The taxonomy of the tetraploid taxa has not been clearly resolved and historical relationships among diploids and tetraploids within the genus have not been predicted. Diploid species are clearly recognizable and complete distributions for most have been described. However, with the exception of *N. kunapalari* (Mahony and Roberts 1986; Roberts and Majors 1993), tetraploid taxa are more problematic. *N. centralis* was described by Parker (1940), *N. sudelli* by Lamb (1911) and *N. aquilonius* by Tyler *et al.* (1981)—before the ploidy status or distribution of any of the taxa were known. Morphological differences among these species are slight and mating call variation does not allow reliable acoustic identification, making species status of some populations uncertain. Without any formal redescription, Roberts (1978) resurrected *N. sudelli* from the synonymous of *N. pictus* (a diploid) but the names *N . centralis* and *N. sudelli* have been applied uncritically to geographically

rather than morphologically or genetically defined specimens (e.g. Roberts and Majors 1993; Cogger 1992). Similarly the name *N. aquilonius* has been applied to populations from north-western Western Australia and the Northern Territory (Tyler *et al.* 1994; Tyler and Davies 1986). Comparison of tetraploid "species" based on genetic criteria could help to resolve some of this taxonomic uncertainty.

The goals of the study were to: 1) clarify the status of notional tetraploid taxa within the genus; and 2) predict the number of origins of tetraploidy within the genus from diploid ancestors. I sequenced a region of the mitochondrial gene cytochrome oxidase I to study historical relationships within the genus *Neobatrachus*. This putatively rapidly evolving gene was chosen for comparison based on immunological evidence suggesting that speciation within the genus has been recent (Roberts *et al.* 1996). Because there were no *a priori* predictions available to assess the biological appropriateness of the resulting phylogeny, a number of analytical methods were used to assess relative confidence in hypotheses of relatedness generated from the mitochondrial DNA data set.

MATERIALS AND METHODS

Samples

Two individuals from one representative population from each of the ten currently recognized taxa within the genus *Neobatrachus* were used in the analysis (Table 3.1). All samples were obtained from a tissue bank collection at the South Australian Museum. Two additional populations of tetraploids (*N. centralis* from Western Australia and *N. aquilonius* from the Northern Territory) were included because there has been some question as to their taxonomic status. Two populations

of the diploid, *N. fulvus*, were also included because allozyme evidence suggested that these populations were heterogenous (M. Mahony, S. Donnellan, unpublished). *Notaden melanoscaphus* was chosen as an outgroup because *Notaden* has been identified as a sister genus of *Neobatrachus* in all phylogenetic analyses of relationships among Myobatrachine frogs (Heyer and Liem 1976; Farris *et al.* 1982; Maxson 1992). Sequences from *Litoria cooloolensis* (family Hylidae) from Queensland and from *Xenopus laevis* (laboratory population) were used as more distant outgroups to determine if rooting affected relationships of the ingroup taxa.

Sequences

The polymerase chain reaction (PCR) (Saiki *et al.* 1985, 1986) was used to amplify a 600-bp region of cytochrome oxidase I (COI) using primers modified to *Xenopus* sequences from the COIe-H and COIa-H primers described in Palumbi *et al.* (1991) and Palumbi (1996). Primer sequences were: 5' CCT GCA GGA GGA GGA GAY CC 3' (forward primer) and 5' TGT ATA AGC GTC TGG GTA GTC 3' (reverse primer). An annealing temperature of 48°C was used for most reactions, but certain templates were problematic and the annealing temperature was either raised or lowered to improve product quality. PCR products of the size fragment corresponding to this region were gel purified using low-melting-point agarose. Target bands were cut from the gel and DNA recovered using a commercially available purification method (Prep-a-gene® ; BioRad). Purified PCR templates were sequenced using a thermal cycle sequencing method (Craxton 1991; Murray 1989), based on the BRL dsDNA Cycle Sequencing system. This technique is based upon the dideoxy chain termination method of Sanger *et al.* (1977) but utilizes end-labeled

primers in a linear amplification procedure that results in efficient sequencing of even very small amounts of template. Sequences were generated for both strands of DNA within an individual and two individuals per population were sequenced to detect possible PCR artifacts or ambiguities in sequences generated.

Analyses

Sequences were aligned using MacVector (International Biotechnologies, Inc) and multiple alignments were optimized with the program Clustal V (Higgins and Sharp 1988, 1989; Higgins *et al.* 1992). Alignments were verified by eye based on codon positions by translating DNA sequences into protein sequences using MacClade (3.04: Maddison and Maddison 1992) and searching for stop codons. Phylogenetic reconstructions were performed using a test version of PAUP*, version 4.0d42 (Swofford 1996). Base compositional biases and average transition:transversional biases were determined using McClade 3.04 (Maddison and Maddison 1992) and/or likelihood estimates (PAUP* 4.0d42: Swofford 1996) from the most-parsimonious tree found in an initial search using unordered characters. These values were used to define stepmatrices for subsequent weighting of particular types of nucleotide substitutions. Where more than one equally parsimonious tree was found, the tree with the highest likelihood was used to estimate parameters. Likelihood estimates were also compared for trees found using different weighting schemes and/or different optimality criterion. Differences in codon usage among taxa were also assessed using the base composition and dinucleotide comparison options in PAUP*. The average ratio of synonymous: nonsynonymous changes was

estimated using the method of Nei and Gobojori (1984) using a program written by D. Haydon at Oxford University.

Three sets of phylogenetic analyses were performed: using maximum parsimony, maximum likelihood, and minimum evolution as optimality criteria (PAUP* 4.0d42: Swofford 1996). For each of the three optimality criteria a number of searches were conducted using different models of evolution (i.e., ranging from simple to complex) to assess relative sensitivity to violations in assumptions about underlying model of nucleotide substitution. In each case, heuristic searches were conducted with initial trees obtained by simple stepwise addition, followed by branch swapping using the TBR (tree bisection-reconnection) routine implemented in PAUP*. Only minimal length trees were saved and zero length branches were collapsed. Multistate characters within taxa were interpreted as uncertainties and an accelerated transformation procedure (ACCTRAN in PAUP*) was used for character optimization. To further assess sampling variation inherent in the data, nonparametric bootstrap analyses (100 replicates) were performed on each data set using each weighting scheme. Trees obtained using the various methods were compared using likelihood estimates of fit of the data using an HKY85 model of nucleotide substitution (Hasegawa *et al.* 1985). Likelihoods were evaluated using test version 4.0d45 of PAUP*.

Parsimony

Analyses using parsimony were conducted initially with characters unweighted. Subsequently, a number of different weighting schemes were used which were based on predictions about biases that may be apparent in nucleotide data

sets from coding regions in general and on specific patterns observed in the data set. For example, although there are twice as many types of transversions as transitions possible, transitions tend to occur more frequently in the mitochondrial genome (Brown *et al.* 1982; Moritz *et al.* 1987). Analyses were conducted using the most appropriate weighting based on estimates of these parameters as described above and using inflated weightings to assess how an unreasonably high estimates of bias might affect the results. Analyses with third base pairs removed and differential weighting of first and second versus third base pair positions were conducted to determine if high frequency of changes at third base positions (i.e., larger number of synonymous changes) might obscure resolution of relationships, especially among deeper splits in the phylogeny.

To evaluate the relative signal to noise ratio in the data set, the length of the most parsimonious tree found from the combined data set using unweighted parsimony was compared to the treelength distribution generated from searches based on random permutations of the data using the PTP test option in PAUP* (version 4.0d42: Swofford 1996). This permutation-tailed-probability test can be used to assess whether significant signal is present in the data compared to the 95% limits of a random distribution.

Minimum Evolution

Analyses using minimum evolution (PAUP*, 4.0d42; Distance) were conducted using a variety of distance estimates implemented in PAUP*. Distance measures were chosen to include simple, i.e., absolute distances, unweighted *p*; specific, i.e., Jukes-Cantor (Jukes and Cantor 1969) and HKY85 (Hasegawa *et al*

1985); and more general models, i.e., general-time reversible (Lanave *et al.* 1984; Rodríguez *et al.* 1990) and LogDet transformation (Lockhart *et al* 1994) and likelihood distances.

Maximum Likelihood

Analyses using maximum likelihood (PAUP*, 4.0d42; Likelihood) were conducted, again based on a variety of models of evolution. To evaluate relative change in the statistic under a hypothesis of constant branching relationships, the log likelihoods for each model were compared initially based on the tree found using minimum evolution with a LogDet transformation. Independent heuristic searches using maximum likelihood based on each model were also conducted to determine if the particular model used would affect determination of the branching relationships.

Parametric Bootstrap

A parametric bootstrap approach also was used to determine if conflicting hypotheses of branching relationships could be distinguished. This method compares the observed estimates in a parameter under alternative hypotheses to a null distribution based on simulations using the null hypothesis as a model (Hillis *et al.* 1996b). To test the null hypothesis that tetraploids were more closely related to each other than to the diploid taxa (i.e., monophyly of tetraploids), the best tree found under the constraint of tetraploid monophyly was compared to the best tree found without the constraint. The difference in the constrained and unconstrained trees found using parsimony (comparing difference in tree length) and maximum likelihood (comparing difference in likelihood) was computed. These real differences were then

compared to those found from a set of 100 data sets of the same size as the original that were simulated based on the branch lengths, base frequencies and transition:transversion ratios of the best trees found under the constraint of monophyly using a program written by John Huelsenbeck (The Siminator). Branch lengths in terms of numbers of substitutions per site, base frequencies, and transition:transversion ratios of the model tree were calculated for both parsimony and likelihood trees using likelihood estimates (PAUP* 4.0d42: Swofford 1996). The distribution of differences determined from the simulated data sets was used to establish a null distribution for comparison with the difference found from the original data set. Rejection of the null hypothesis would occur if the real difference fell outside the 95% limits of this distribution.

RESULTS

Sequences

Comparison of sequences from a 572-bp region of COI (Appendix C) yielded 178 variable sites with an average transition: transversion ratio of approximately 3.6:1 based on the tree with the highest likelihood obtained from an initial unweighted analysis using all taxa (using MacClade 3.01). Transition: transversion biases also varied widely among different pairs of taxa, ranging from 3.90 to 21.00, suggesting that averaging across branches could be misleading (Table 3.2). The likelihood estimates for parameters based on the tree with the highest likelihood (Ln Likelihood=–3368.09833) were: transition/transversion ratio=3.4:1; frequency of A=0.24286, frequency of C=0.28413, frequency of G=0.1888, frequency of T=0.28413. Most of the variation among sequences resulted from synonymous

amino acid substitutions and third base pair substitutions. The ratio of synonymous to nonsynymous changes was 18:1 when outgroup (*Notaden*) sequences were included in the analysis but 60:1 when only the ingroup taxa were included. Sequences from individuals within populations were identical for *N. wilsmorei* populations and for *N. sutor* populations, so only a single sequence was used for each in the phylogenetic analysis. Sequences from *N. aquilonius*-NT and *N. sudelli*-QLD were only obtained from a single individual for each because PCR amplifications were unsuccessful for the second sample of each.

Analyses

Phylogenetic Signal

The PTP test showed a significant difference ($p \leq 0.01$) between the trees found from an initial search using unweighted parsimony compared to a distribution generated from random permutations of the data, suggesting that sufficient signal was present in the data to distinguish phylogenetic signal from random noise (Figure 3.2).

Parsimony

An heuristic search (as described in the methods section) using unweighted maximum parsimony (PAUP* 4.0d42; Swofford 1996) yielded 42 equally parsimonious trees, each with a consistency index (C.I.) of 0.532. The strict consensus tree is shown in Figure 3.3a. The tree with the lowest likelihood (-3345.144) differed from the highest likelihood tree (-3340.158) only in relationships among some of the diploid taxa. The most notable feature in these trees is the relationship of the tetraploid taxa to each other. Two tetraploids from South Australia

(*N. centralis* -SA and *N. sudelli*-SA), one from Queensland (*N. sudelli*-QLD), one from central Western Australia (*N. centralis* -WA), and one from the Northern Territory (*N. aquilonius*-NT) formed a clade ("eastern tetraploids") that was separated from any of the diploid taxa and from the remaining tetraploid taxa. Individuals from within populations (e.g., *N. centralis*-SA1 and *N. centralis*-SA2) tended to be each other's closest relatives but precise relationships within this clade varied across trees. One of the tetraploids from Western Australia (*N. aquilonius*-WA: Figure 3.1) appeared to be most closely related to the diploid *N. fulvus* , which is found in the same general region. The two tetraploid *N. kunapalari*-WA samples used were from different populations in midwestern Western Australian and these did not appear to be each other's closest relatives. *N. kunapalari*-WA2 appeared to be most closely related to the diploid, *N. albipes*-WA, with which it has an overlapping range. However, the *N. kunapalari*-WA1 population did not appear to be closely related to any of the other taxa. The relationships of the rest of the diploids (*N. wilsmorei, N. sutor, N. pictus* and *N. pelobatoides*) to the tetraploids and to each other varied across trees but *N. wilsmorei* and *N. sutor* usually were placed as sister taxa.

A reanalysis of the data weighting transversions four times as heavily as transitions (based on the average transition bias estimated) produced 24 equally most parsimonious trees that again varied in relationships mainly within the clades identified in the unweighted search (Figure 3.3b). The only major difference between these trees and the trees from the unweighted search was in the placement of the tetraploid *N. kunapalari*-WA1; in this analysis it was included in the clade with tetraploid *N. aquilonius*-WA and diploid *N. fulvus*. Weighting transversions three times as heavily as transitions produced 48 equally parsimonious trees that placed

both populations of *N. kunapalari*-WA together with *N. albipes* and resolved the diploid relationships better than in the 4:1 weighting scheme (Figure 3.3c). However, the 4:1 weighting scheme produced a higher likelihood tree than the 3:1 scheme (trees with highest likelihood: -3340.08973 vs -3351.68109, respectively). Using an asymmetric stepmatrix based on average frequency of specific types of changes (e.g , AG transitions vs CT transitions) determined using McClade (3.04) or increasing the transversion weighting to 10 times the transition rate (i.e., representing an inflated bias), significantly altered branching relationships. Weighting first and second base pair positions three times as heavily as third base pair positions or removing third base pair positions from the analysis appeared to remove most of the signal as the trees found (>4000 equally parsimonious trees) were largely unresolved.

Nonparametric bootstrap analyses (100 replicates) collapsed all relationships except the tetraploid relationships, regardless of the weighting scheme used. Figure 3.3d shows the consensus tree and bootstrap proportions generated using the 4:1 weighting scheme of transversions to transitions. Following Hillis and Bull (1992), bootstrap proportions >70% may be interpreted as indications of strong support for a particular clade given equal rates of change among taxa and relatively low levels of internodal change. Using this criterion, support was limited for resolution of the diploid taxa but strong for the "eastern tetraploid" clade (bootstrap proportion = 100%) and for the clade including one of the tetraploid *N. kunapalari* populations (WA2) and the diploid *N. albipes*. Support was moderate for the clade including the tetraploid *N. aquilonius* -WA and the diploid *N. fulvus* (bootstrap proportion = 55%). The only other support obtained was for within population groupings of some of the taxa (Figure 3.3d).

Minimum Evolution

Although specific branch length estimates varied using various minimum evolution models, tree topology and relative divergence estimates were the same for all (Figure 4a). The minimum evolution tree found the same tetraploid clades as the parsimony search using a 3:1 weighting of transversions to transitions (i.e., most notably the grouping of both populations of *N. kunapalari* in a clade with *N. albipes*), but the relationship of the tetraploid clades to the other diploids was not quite the same. Phylograms of the highest likelihood parsimony tree found using a 3:1 transition to transversion weighting scheme (Figure 3.4b) and using a 4:1 transition to transversion weighting scheme are shown for comparision (Figure 3.4c). The likelihood of the minimum evolution tree (-3367.349) was lower than for any of the trees found using parsimony. Bootstrap analysis using minimum evolution again supported only the tetraploid clades.

Maximum Likelihood

All of the searches using maximum likelihood gave the same basic tree as in the minimum evolution searches (Figure 3.4d) but again, specific branch length estimates varied slightly depending on the model used. The placement of the *N. kunapalari*-WA1 sample also varied across trees. When the tree was kept constant, likelihood estimates improved with complexity of the model used from a log likelihood of -3608.070 using a Jukes-Cantor model (Jukes and Cantor 1969) to -3367.349 using an HKY85 model (Hasegawa *et al.* 1985) to -3354.402 using a general-time reversible 6-parameter model (see Swofford *et al.* 1996). Bootstrap

analysis based on maximum likelihood showed similar results to those using parsimony or minimum evolution criteria.

Parametric Bootstrap

To reduce the complexity of the data set to allow comparisons of trees using both parsimony and likelihood in a computationally realistic time frame, a reduced data set was used for parametric bootstrap tests (Figure 3.5a). Within the "eastern" tetraploid clade, all duplicate samples within populations were removed because they caused variation within the clade but did not affect relationship of this clade to the other taxa in the analysis. Duplicate samples within populations for all of the diploid taxa except *N. fulvus* were also removed because these taxa varied in relation to each other but did not affect diploid-tetraploid relationships in general. For *N. fulvus*, the two samples that were most different from each other were used. Both populations of *N. kunapalari*-WA were used, as were both individuals from the *N. aquilonius*-WA population because of the variation in relationships of these taxa across trees. Based on an unweighted search on this set of taxa, two equally parsimonious tree were found for which the average transition:transversion ratio was 3.6:1. Base frequencies were estimated to be A=0.24286; C=0.28413; G=0.18888; and T=0.28413. The data were reanalyzed using a 4:1 weighting of tranversions to transitions and the difference in tree length of the single most-parsimonious tree found in this search (Figure 3.5a) compared to that of the best tree found under a constraint of monophyly of the tetraploids (Figure 3.5b) was found to be 23 steps. The transition:transversion ratio, base frequencies and branch lengths of the best tree found under the constraint of monophyly (determined by likelihood distance

estimates) were used to reconstruct 100 data sets of 569 bp using The Siminator (J. Huelsenbeck). These simulated data sets were then analyzed using parsimony with a 4:1 transversion:transition weighting. For each data set, the difference in treelength for the best tree and the best tree under the constraint of tetraploid monophyly was computed. The largest difference found for the simulated data sets was only 2 steps, so even without producing a null distribution it could be concluded that the real difference (23 steps) was highly significant and the hypothesis of monophyly of the tetraploids was rejected.

The analysis was repeated using likelihood with HKY85 used as an underlying model of evolution (Hasegawa *et al.* 1985). The real difference between the unconstrained and constrained trees was 36.84 likelihood units while the largest difference found from the simulated data sets was 2.90. Although likelihood produced a broader distribution of differences, the analysis took 10 days of real time and conclusions did not change from the parsimony analysis, which took less than an hour of computation time.

DISCUSSION

Evaluation of Trees Obtained

Results from this study identify at least two apparently independent origins of tetraploids in the genus *Neobatrachus*, but make few definitive suggestions on relationships among diploids or between tetraploids and diploids. While there was some resolution of diploid relationships, the variation in maximum parsimony and the limited resolution in nonparametric bootstrap analyses using any of the optimality criterion suggests that only the tetraploid relationships have been resolved with

confidence using this mitochondrial data set. While this may be a reflection of homoplasy, the results do emphasize some considerations that should not be overlooked when using sequence data for phylogenetic reconstruction.

The main precautionary note is that our limited current understanding of molecular evolution may inhibit the ability to extract unambiguous information from sequence data, especially when large discrepancies exist among branch lengths. Problems identified using the four taxon case (e.g., Huelsenbeck 1995) could become even more of a problem when using larger numbers of taxa for which the possible number of alternative topologies increases astronomically. The difficulty in using methods of phylogenetic reconstruction to establish a hypothesis of historical relationships within a group for which little other information exists is that the true underlying phylogeny cannot be verified and it is therefore nearly impossible to assess how the model of evolution affects the ability to extract meaningful phylogenetic signal.

My results support recent findings that phylogenetic analyses tend to be sensitive to biases in underlying patterns of nucleotide substitution (e.g., Hasegawa *et al.* 1985; Thomas and Beckenbach 1989; Moritz *et al.* 1992; Goldman 1990; Wakely 1994; Yang *et al.* 1994; Zhu *et al.* 1994; Swofford *et al.* 1996). Analysis of pairwise comparisons of dinucleotide frequencies across taxa in my study showed large differences in transition:transversion bias among pairs of taxa, suggesting that averaging across all taxa may not adequately account for biases in patterns of nucleotide substitution. This may explain why differences in trees were found even with a small change in transition:transversion weighting schemes (i.e., 4:1 vs 3:1 using parsimony). Although conclusions based on the 3:1 weighting scheme

predicted a more consistent grouping of the *N. kunapalari* populations and resolved diploid relationships better, the likelihood of the best tree found with the 4:1 weighting scheme was higher, suggesting that this transition bias fits the data slightly better. It may be that the lower ratio resolved deeper splits while the higher ratio was better able to resolve more shallow splits. Other factors, such as rate heterogeneity (e.g., Yang *et al.*, 1994; Wakely, 1994), and rates of change for synonymous and nonsynonymous nucleotide substitutions (e.g., Li *et al.* 1985; Nei and Gojobori 1986; Li 1990; Muse and Gaut 1994; Yang *et al.* 1994) also may be important in establishing a realistic model of nucleotide change; ignoring these effects could result in false conclusions. Because it is nearly impossible to evaluate when the "best" model for the data has been chosen using parsimony, it may not be the most appropriate optimality criterion to use when evidence suggests that patterns of nucleotide change may obscure resolution. Estimating the likelihoods for trees found using parsimony may be one way of generally choosing the most appropriate analysis, but if patterns of nucleotide change are not consistent across taxa, as was apparent for transitional biases, this may not reflect when the best resolution has been obtained across all parts of the tree.

Using a model-based optimality criterion such as maximum likelihood provides a statistical criterion for establishing how well the model of evolution fits the data set compared to an alternative model, but there is still no guarantee that improvement in the model will lead to a "better" tree. It has been suggested that maximum likelihood tends to be relatively robust to the model of evolution used (e.g., Yang *et al.*, 1994). It would be tempting in my study to ignore conclusions based on parsimony and place confidence in the tree produced by these more model-

based analyses. However, the collapse of resolution in bootstrap analyses using both parsimony and minimum evolution introduces a reasonable doubt that these analyses actually do better than parsimony at reconstructing "true" relationships.

I would therefore place confidence only in the branching relationships supported by all analyses and consider those that vary to be unresolved by this dataset. Analytical programs, such as PAUP*, now make it easy to perform a variety of analyses on the same data set and my results emphasize that this type of approach should become the standard, rather than the exception.

Origins of Tetraploids

The relationships that were strongly supported by my data were the relationships of the tetraploids to each other. The most strongly supported result is that there have been at least two independent origins of tetraploids—one resulting in the "eastern" tetraploids (*N. centralis N. sudelli* and *N. aquilonius* from the Northern Territory) and one in the tetraploids from Western Australia (*N. aquilonius* and *N. kunapalari*). There was some evidence that the two "western" tetraploids may have had separate origins, but this result varied across analyes. *N. fulvus* was identified as the sister taxon to *N. aquilonius* from Western Australia in most analyses and to one population of *N. kunapalari* in some analyses. *N. albipes* was identified as the most closely related taxon to *N. kunapalari*, but again, this was not a completely consistent result. None of the diploids were identified as more closely related to the "eastern" clade and it is not clear how the two groups of tetraploids are related to each other. Relationships within the "eastern" tetraploid clade were not well resolved but

divergence among them was not high. In fact, divergence within populations was almost as high as among nominal species.

Among the tetraploids, morphology, mating call variation and distribution are well defined only for *N. kunapalari*. The karyotype of *N. kunapalari* is also unique, with a shift in NOR position from subterminal on the long arm of chromosome 5 (as in all other tetraploids and all diploids except *N. fulvus*, for which the NOR is terminal on chromosome 5) to subterminal on the long arm of chromosome 7. Although the mating call of *N. kunapalari* can be distinguished from those of the other tetraploids, calls are broadly similar with a low pulse number and low pulse rate (Mahony and Roberts 1986; Figure 3.6). These call features are also shared with *N. fulvus* (Roberts and Majors 1993) but not with any other diploid *Neobatrachus* species (Littlejohn 1965; Roberts 1978; Mahony and Roberts 1986). The range of *N. kunapalari* is restricted to the southwestern corner of Western Australia and it has not been found north of Menzies or Payne's Find or east of the western edge of the Nullarbor Plain (Roberts and Majors 1993). Although it overlaps the range of several of the diploids (*N. wilsmorei*, *N. sutor*, *N. albipes*, *N. pelobatoides*) it does not overlap the range of *N. fulvus*, which is restricted to the Exmouth peninsula and Shark's Bay. Current biogeography would therefore support the finding that *N. kunapalari* is more closely related to *N. albipes* than to *N. fulvus*. While the finding that the two populations of *N. kunapalari* did not appear to be more closely related to each other than to other taxa in many of the analyses may reflect real divergence since polyploid origin, it unfortunately could be due to contamination of one or either of the sequences. Sequences from both individuals used were very difficult to obtain and both are incomplete in the middle part of the sequence. An attempt to resequence the

individual designated *N. kunapalari*-WA1 resulted in the generation of a different sequence, which was rejected as a contaminant. However, it is also possible that the original sequence was the contaminant. Because of this doubt, conclusions on origins of this species cannot be made conclusively.

The distribution of *N. aquilonius* is less well defined but it has not been found south or west of Karratha. The distinction between *N. kunapalari* and *N. aquilonius* is therefore clear: no geographic overlap and distinct chromosome morphology. This implies either separate origins of these two "western" tetraploids or post-origin differentiation from a common progenitor in Western Australia. Although *N. fulvus* was identified as the most closely related diploid to *N. aquilonius* (and sometimes to *N. kunapalari*), has a similar mating call to the tetraploids, and thus could be identified as a potential ancestor, the terminal NOR position on chromosome 5 in *N. fulvus* is not shared with either *N. kunapalari* or *N. aquilonius*. It is also possible, therefore, that *N. fulvus* diverged from a common ancestor with these western tetraploids.

While mitochondrial data reflect only maternal inheritance and provide gene rather than species level phylogenies (e.g., Moritz *et al.* 1987), my results do suggest some problems with the current taxonomic status and geographic distributions of some of the named tetraploid species. A consistent result from my analyses is that frogs referred to as *N. aquilonius* in central Australia (Northern Territory) have a distinctively different mitochondrial haplotype than *N. aquilonius* from Western Australia (Port Headland) and appeared to be more closely related to *N. centralis* from South Australia and *N. sudelli* from Queensland (i.e., in the eastern tetraploid clade). Within the eastern tetraploid clade *N. sudelli* from Queensland appeared to be

more closely related to *N. centralis* from South Australia than to *N. sudelli* from South Australia. *N. centralis* from Western Australia appears to have a different haplotype than *N. centralis* from South Australia. However, *N. centralis* from Western Australia currently divides the ranges of the two western tetraploids (*N. kunapalari* and *N. aquilonius*; Roberts and Majors 1993; Figure 3.1) and a closer relationship to *N. sudelli* from South Australia than to *N. centralis* would require a complicated crossing of ranges.

Because identification of tetraploids (except *N. kunapalari*) has been based largely on geographic location rather than morphological or genetic criterion, it is possible that these individuals have not been identified correctly or that the three "species" actually represent a single lineage that has diverged in areas of isolation. Further, branch lengths within the eastern clade are relatively small compared to those among the western tetraploids, suggesting that little genetic divergence has occurred (even in this rapidly evolving mitochondrial gene). The mating calls of *N. sudelli* from South Australia and *N. centralis* are recognizably distinct. As well, *N. centralis* from Western Australia produces calls that can be distinguished from those of *N. centralis* from South Australia (J.D. Roberts, unpublished data).

In summary, my results do not resolve the taxonomic status of tetraploids within the eastern clade with any confidence but do suggest reservations with the names currently applied. The important points are 1) that the eastern tetraploids are genetically removed from the western tetraploids, 2) their high degree of similarity in mitochondrial haplotype suggests that what are currently referred to as *N. sudelli* and *N. centralis* have arisen from a common ancestor and 3) that *N. aquilonius* from the Northern Territory is genetically distinct from *N. aquilonius* from Western Australia.

If tetraploids within this clade do represent distinct species, the high degree of mitochondrial similarity among them compared with the large divergence from other species in the genus suggests that they may have speciated as tetraploids following the original polyploidization event. Although relationships were not well resolved, results generally suggest sequence divergence by geographic region, which would support a hypothesis of speciation by geographic isolation. If this is true, it would be the first evidence of speciation within a tetraploid lineage in anurans.

Origins of the eastern tetraploids from a diploid ancestor also remain unresolved. The most consistent result from my study is that the eastern tetraploids arose from a common ancestor with all of the other extant species in the genus and are thus older than previously predicted. The inability to define ancestry of tetraploids using mitochondrial genes is consistent with results using other groups of polyploids. The North American grey treefrogs are have been divided into a tetraploid, *Hyla versicolor* (4n=48), and a diploid, *Hyla chrysoscelis* (2n=24) (Wasserman 1970; Bogart and Wasserman 1972). Various data sets—allozymes (Ralin 1978; Ralin and Selander 1979; Ralin *et al.* 1983; Romano *et al.* 1987); immunological distance data (Maxson *et al.* 1977) morphometric analyses (Ralin and Rogers 1979); mating call variation (Gerhardt 1974; Ralin 1977); and mtDNA sequences (Ptacek *et al.* 1994) have not resolved the number (single or multiple) or, especially, type (allopolyploid or autopolyploid) of origin of the tetraploids (e.g. Maxson *et al.* 1977; Ralin 1978; Ralin and Rogers 1979; Ptacek *et al.* 1994). Ptacek *et al.* (1994) predicted multiple origins of tetraploids within the North American grey treefrog complex but were unable to identify an originating diploid population for at least some of the tetraploid "lineages". Chapter 2 of this dissertation suggests that tetraploids in the northern

parts of the range are not closely related to the diploids with which they are currently in association and may have arisen from an ancestral population that no longer exists, possibly because it was outcompeted by the pioneering tetraploid species.

Even if this study had generated a clearly resolved picture of historical relationships, maternal phylogenies cannot distinguish between autopolyploid (i.e., from a single population) and allopolyploid (i.e., hybrid) origins. The presence of multivalents at meiosis I in the tetraploid taxa examined (Mahony 1986) indirectly suggests an autopolyploid origin, but the high degree of conservation in chromosome morphology may make this an unreliable criterion. Comparison with nuclear genes, or preferably with paternally inherited genes would be necessary to address questions on which type of origin is most likely. Attempts to sequence intron regions of single-copy nuclear genes for comparison with the mitochondrial results are discussed in Chapter 4.

Relationships among the diploid taxa to each other and to the tetraploids were not well resolved but species distinctions for the diploids have been better defined by behavioural, morphological and ecological factors and their relationships were not as critical to this study. One result that was consistent among the diploid taxa was that *N. wilsmorei* and *N. sutor* were sister taxa and never appeared as close relatives to the tetraploid taxa. This result is consistent with the most parsimonious explanation for origins of mating call type as these two taxa are the only ones that produce single-pulsed calls (see Figure 3.6).

CONCLUSIONS

I conclude that there have been at least two, and possibly three origins of tetraploids in the genus *Neobatrachus*. Relationships to existing diploids are poorly defined. If there have been only two origins, there must have been post-origin differentiation in the "western" clade resulting in NOR rearrangements and changes in morphology between western *N. aquilonius* and *N. kunapalari*. In the "eastern" tetraploid clade sequence differences were small and may reflect geographic and temporal isolation of populations of explosive breeders in the arid zone (Roberts *et al.* 1996). Species status cannot be determined from an analysis of mitochondrial DNA data alone, but this could be the first evidence of speciation of tetraploids from other tetraploids following the original polyploidization event. My results suggest a preliminary framework in which to address further questions on the origins of tetraploids and their subsequent speciation patterns, such as reproductive compatibility among the identified lineages.

I also conclude that phylogenetic hypotheses based on mitochondrial (or other rapidly evolving) DNA sequence data be treated with caution. Incorporation of complex models of evolution may improve the ability to extract meaningful phylogenetic signal when rates of nucleotide substitution differ among branches or potential levels of homoplasy are high. However, the lack of a completely reliable method for assessing when model complexity adequately describes the data does not allow a criterion for choosing between alternative trees when predictions of branching relationships vary using different optimality criteria or different models of evolution. I therefore suggest that confidence be placed in results that remain stable over a range of models and more than one optimality criterion, while results that vary should be interpreted as poorly resolved.

Table 3.1. Tissue samples used in species level comparisons of *Neobatrachus* species. All samples were obtained from the Evolutionary Biology Unit tissue collection at the South Australian Museum in Adelaide. Tissue sample numbers are given in the left hand column. Specimen registration numbers (where known) are given in the second column.

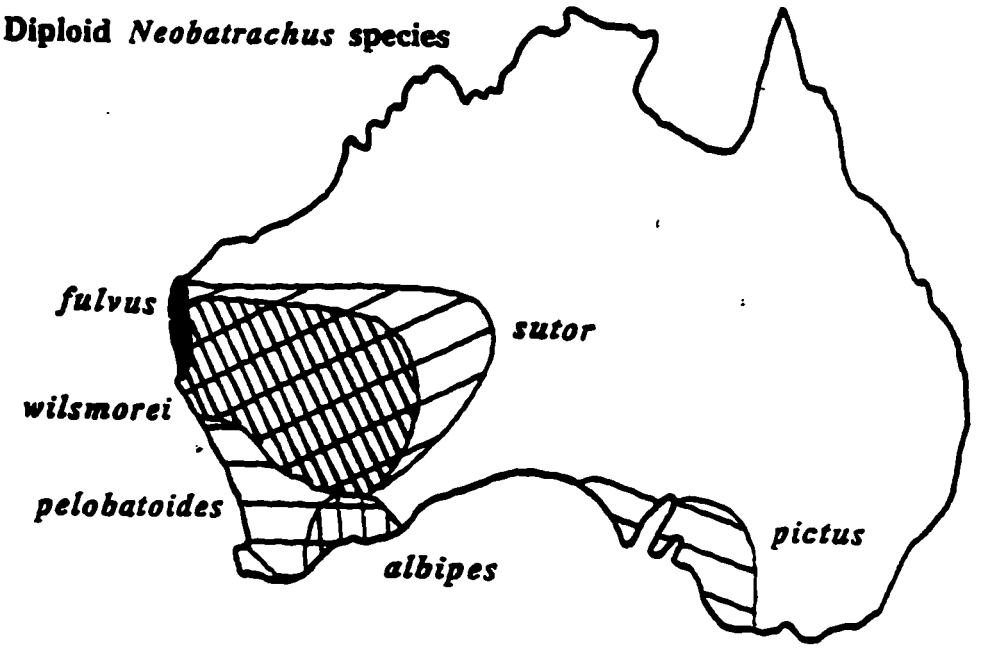
CAT#	REG#	SPECIES	LOCALITY
A/ Tetraploids			
C132	SAMA R24186	*Neobatrachus centralis*	Gawler Ranges, SA
I20		*Neobatrachus centralis*	Woomera, SA
K314	WAM R103596	*Neobatrachus "centralis"*	Mt. Magnet, WA
K319	WAM R103601	*Neobatrachus "centralis"*	Mt. Magnet, WA
H519		*Neobatrachus sudelli*	East of Burra, SA
H520		*Neobatrachus sudelli*	East of Burra, SA
K246	SAMA R43075	*Neobatrachus sudelli*	QLD
K247	SAMA R43076	*Neobatrachus sudelli*	QLD
K336	SAMA R37874	*Neobatrachus aquilonius*	Port Headland, WA
K338	SAMA R43077	*Neobatrachus aquilonius*	Port Headland, WA
K504	WAM R103620	*Neobatrachus kunapalari*	Ninghan Sta, WA
J672		*Neobatrachus kunapalari*	Coolgardie, WA
B/ Diploids			
K124	SAMA R29186	*Neobatrachus pictus*	Whartook, VIC
K125	SAMA R43073	*Neobatrachus pictus*	Whartook, Vic
K92	SAMA R43070	*Neobatrachus pelobatoides*	Ravensthorpe, WA
K93		*Neobatrachus pelobatoides*	Ravensthorpe, WA
K509	WAM R103625	*Neobatrachus fulvus*	Carnarvon, WA
K510	WAM R13626	*Neobatrachus fulvus*	Carnarvon, WA
K805		*Neobatrachu fulvus*	Carnarvon, WA
K808		*Neobatrachu fulvus*	Carnarvon, WA
K496	WAM R103612	*Neobatrachus sutor*	Carnarvon, WA
K498	WAM R103614	*Neobatrachus sutor*	Carnarvon, WA
K493	WAM R103609	*Neobatrachus wilsmorei*	Carnarvon, WA
K494	WAM R103610	*Neobatrachus wilsmorei*	Carnarvon, WA
J699	WAM R103633	*Neobatrachus albipes*	Quairading, WA
J752		*Neobatrachus albipes*	Coolgardie, WA
C/ "Outgroups"			
A43		*Notaden melanoscaphus*	Townsville, QLD
A44		*Notaden melanoscaphus*	Townsville, QLD

Table 3.2. Pair-wise comparisons of relative sequence divergence in cytochrome oxidase I (COI) using *N. sudelli* from South Australia as a reference taxon. The COI sequence had a total of 769 sites, of which 178 were variable.

Taxon	Transitions	Transversions	Proportion Difference	Ti/Tv Ratio
N. sudelli-QLD	21	1	0.04	21.0
N. centralis-SA	25	1	0.05	25.0
N. centralis-WA	24	0	0.04	infinite
N. aquilonius-NT	11	0	0.04	infinite
N. aquilonius-WA	52	10	0.11	5.20
N. kunapalari-WA	39	10	0.11	3.90
N. fulvus-WA	58	17	0.13	3.41
N. albipes-WA	62	10	0.13	6.20
N. pictus-Vic	66	9	0.13	7.33
N. pelobatoides-WA	64	9	0.13	7.11
N. wilsmorei-WA	66	14	0.14	4..71
N. sutor-Wa	65	13	0.14	5.00
Notaden	71	53	0.22	1.34

Figure 3.1. Distribution of named species of *Neobatrachus* across Australia. (a) Diploid distributions—note that *N. pictus* is the only diploid found in eastern Australia and the limited range of *N. fulvus*. (b) Tetraploid distributions—note that tetraploid taxa extend across the entire central desert region. The range of *N. kunapalari* is well defined but the ranges of the other tetraploids are not known, especially in the more inaccessible central parts of the range. The black bar between the ranges of *N. centralis* and *N. kunapalari* in Western Australia indicates a region of hybridization.

a) **Diploid *Neobatrachus* species**



b) **Tetraploid *Neobatrachus* species**

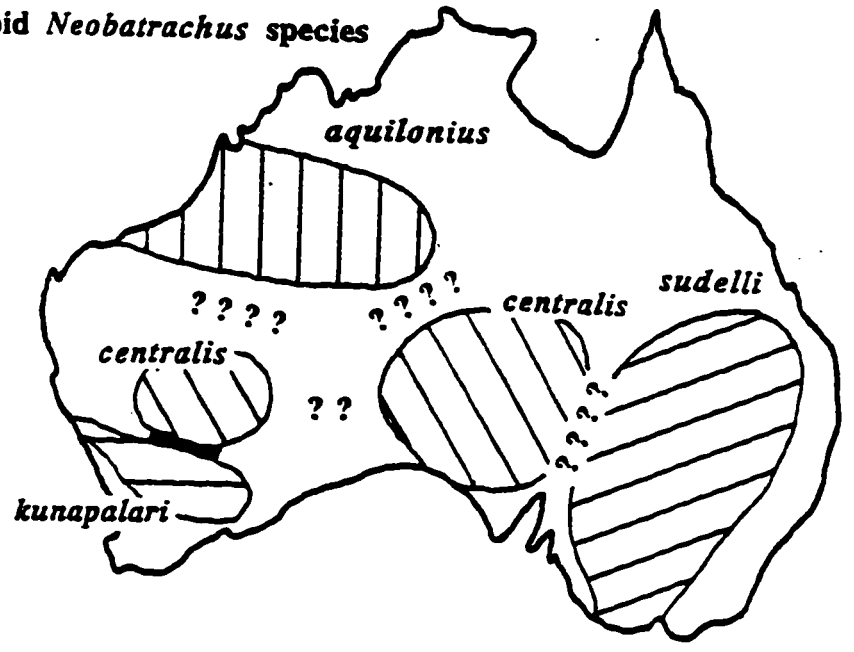
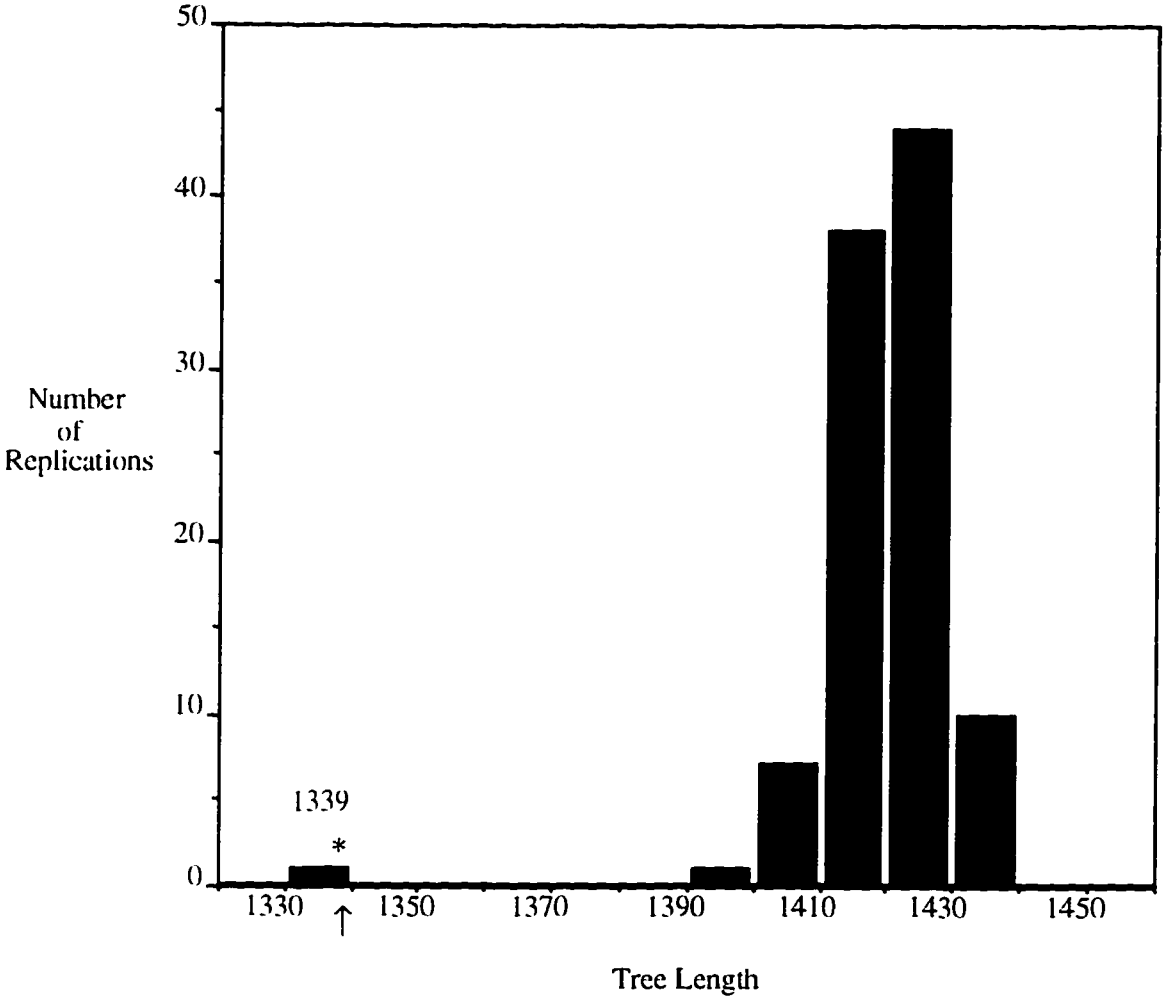


Figure 3.2. Distribution of random trees generated from random permutations of the data using the PTP test (permutation tailed probablility test) implemented in PAUP* 4.0d42 (Swofford 1996). The treelengths of the actual trees found in an unweighted search fell outside the 99% limits of this distribution ($p<0.01$), suggesting that sufficient signal was present in the data to distinguish phylogenetic signal from random noise.

PTP Test



* = length for original (unpermuted) data

PTP = 0.010000

Figure 3.3. Comparison of trees generated using parsimony (PAUP*4.0d42; Swofford 1996) with different character weighting schemes. For each tree, branches that are shaded indicate ones that vary within a clade across analyses; branches in bold indicate nodes that change radically across analyses; and tetraploid taxa are indicated in bold type. (a) Strict consensus of 42 equally parsimonious trees using characters as unordered and unweighted. (b) Strict consensus of 24 equally parsimonious trees when transversions are weighted 4 times as heavily as transitions. (c) Strict consensus of 48 equally parsimonious tree when transversions are weighted 3 times as heavily as transitions. (d) Strict consensus of 100 bootstrap replications using the 4:1 transversion to transition weighting scheme. Note that variability in trees is mainly due to variation within major clades and in the relative placement of the two tetraploid *N. kunapalari* populations from Western Australia. Note especially that small differences in weighting result in large differences in branching relationships for some of the taxa. Diploid relationships are not well resolved in any of the analyses.

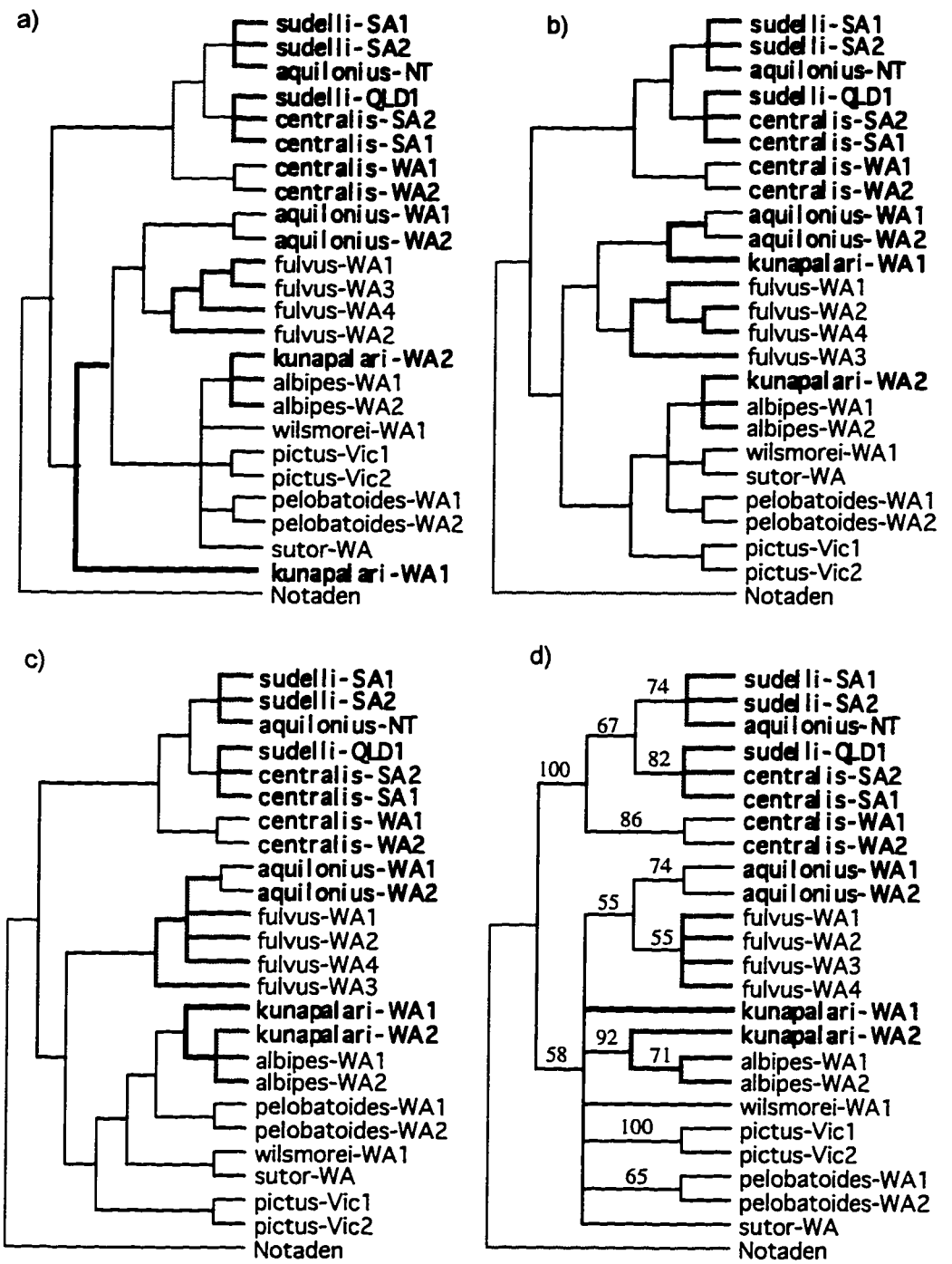

a)
sudelli-SA1
sudelli-SA2
aquilonius-NT
sudelli-QLD1
centralis-SA2
centralis-SA1
centralis-WA1
centralis-WA2
aquilonius-WA1
aquilonius-WA2
fulvus-WA1
fulvus-WA3
fulvus-WA4
fulvus-WA2
kunapalari-WA2
albipes-WA1
albipes-WA2
wilsmorei-WA1
pictus-Vic1
pictus-Vic2
pelobatoides-WA1
pelobatoides-WA2
sutor-WA
kunapalari-WA1
Notaden
b)
sudelli-SA1
sudelli-SA2
aquilonius-NT
sudelli-QLD1
centralis-SA2
centralis-SA1
centralis-WA1
centralis-WA2
aquilonius-WA1
aquilonius-WA2
kunapalari-WA1
fulvus-WA1
fulvus-WA2
fulvus-WA4
fulvus-WA3
kunapalari-WA2
albipes-WA1
albipes-WA2
wilsmorei-WA1
sutor-WA
pelobatoides-WA1
pelobatoides-WA2
pictus-Vic1
pictus-Vic2
Notaden
c)
sudelli-SA1
sudelli-SA2
aquilonius-NT
sudelli-QLD1
centralis-SA2
centralis-SA1
centralis-WA1
centralis-WA2
aquilonius-WA1
aquilonius-WA2
fulvus-WA1
fulvus-WA2
fulvus-WA4
fulvus-WA3
kunapalari-WA1
kunapalari-WA2
albipes-WA1
albipes-WA2
pelobatoides-WA1
pelobatoides-WA2
wilsmorei-WA1
sutor-WA
pictus-Vic1
pictus-Vic2
Notaden
d)
74
67
100
82
86
74
55
55
92
58
71
100
65
sudelli-SA1
sudelli-SA2
aquilonius-NT
sudelli-QLD1
centralis-SA2
centralis-SA1
centralis-WA1
centralis-WA2
aquilonius-WA1
aquilonius-WA2
fulvus-WA1
fulvus-WA2
fulvus-WA3
fulvus-WA4
kunapalari-WA1
kunapalari-WA2
albipes-WA1
albipes-WA2
wilsmorei-WA1
pictus-Vic1
pictus-Vic2
pelobatoides-WA1
pelobatoides-WA2
sutor-WA
Notaden

Figure 3.4. Comparison of trees found using parsimony, minimum evolution, and maximum likelihood. (a) Minimum evolution tree using a Jukes-Cantor (Jukes and Cantor 1969) model of evolution (i.e, no transition:transversion bias; equal base frequencies). Although specific branch lengths varied, using simpler or more complex models did not change the branching relationships. This tree was similar to those found using parsimony with a 3:1 transversion to transition weighting scheme in the relative placement of the *N. kunapalari* populations but differed in relationships among some of the diploid taxa. (b) Phylogram of the highest likelihood tree found using parsimony with transversions weighted 3 times as heavily as transitions. (c) Phylogram of the highest likelihood tree found using parsimony with transversions weighted 4 times as heavily as transitions. (d) Maximum likelihood tree using an HKY model (Hasegawa *et al.* 1985) of evolution (i.e., base frequencies, transition:transversion ratio and branch lengths estimated).

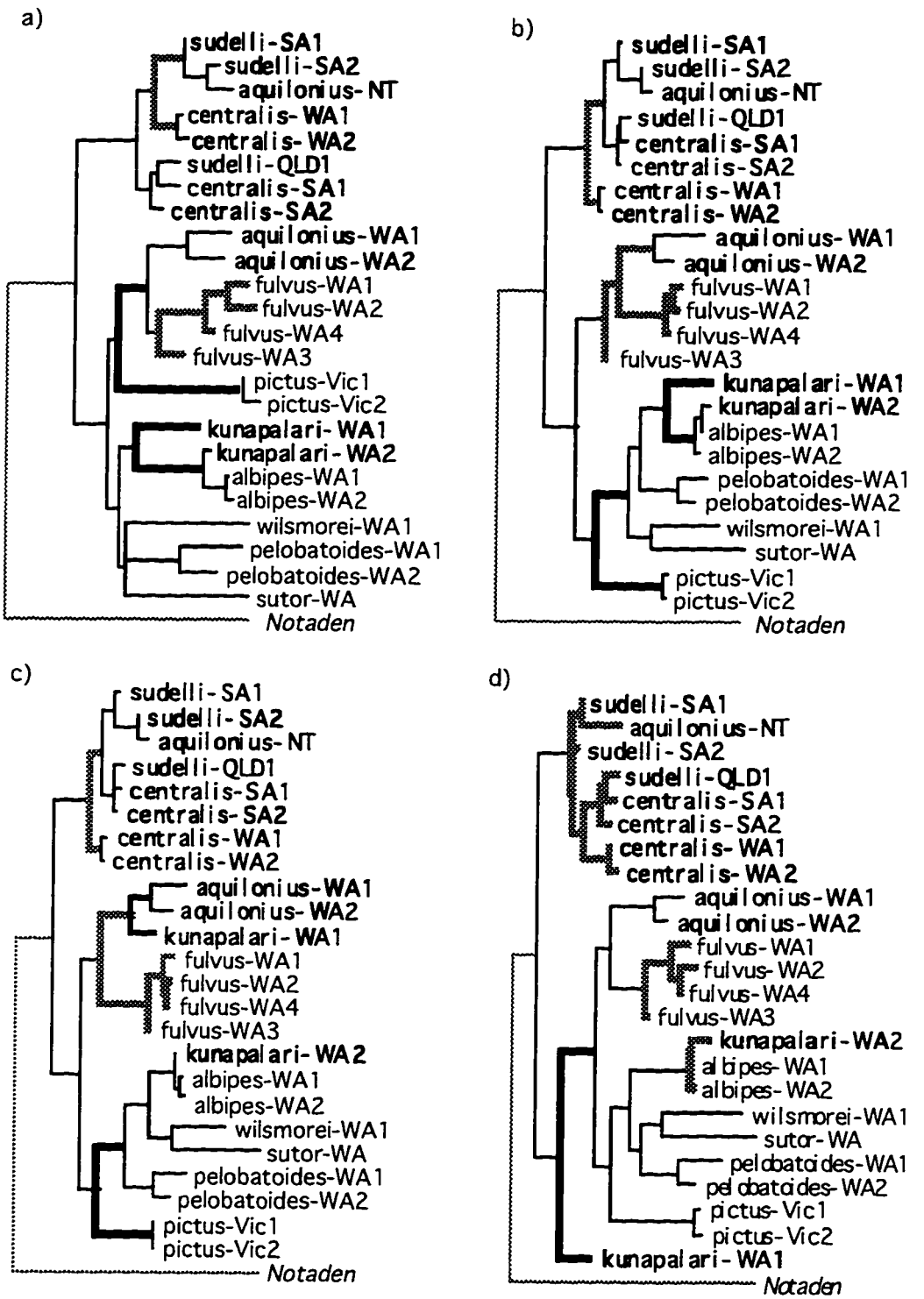
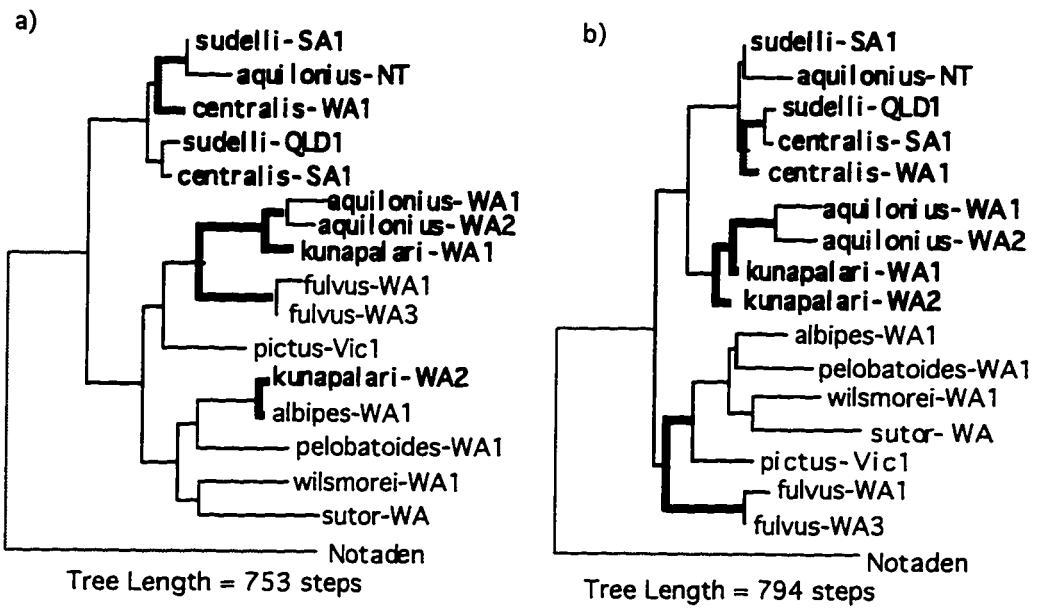

a)
sudelli-SA1
sudelli-SA2
aquilonius-NT
centralis-WA1
centralis-WA2
sudelli-QLD1
centralis-SA1
centralis-SA2
aquilonius-WA1
aquilonius-WA2
fulvus-WA1
fulvus-WA2
fulvus-WA4
fulvus-WA3
pictus-Vic1
pictus-Vic2
kunapalari-WA1
kunapalari-WA2
albipes-WA1
albipes-WA2
wilsmorei-WA1
pelobatoides-WA1
pelobatoides-WA2
sutor-WA
Notaden
b)
sudelli-SA1
sudelli-SA2
aquilonius-NT
sudelli-QLD1
centralis-SA1
centralis-SA2
centralis-WA1
centralis-WA2
aquilonius-WA1
aquilonius-WA2
fulvus-WA1
fulvus-WA2
fulvus-WA4
fulvus-WA3
kunapalari-WA1
kunapalari-WA2
albipes-WA1
albipes-WA2
pelobatoides-WA1
pelobatoides-WA2
wilsmorei-WA1
sutor-WA
pictus-Vic1
pictus-Vic2
Notaden
c)
sudelli-SA1
sudelli-SA2
aquilonius-NT
sudelli-QLD1
centralis-SA1
centralis-SA2
centralis-WA1
centralis-WA2
aquilonius-WA1
aquilonius-WA2
kunapalari-WA1
fulvus-WA1
fulvus-WA2
fulvus-WA4
fulvus-WA3
kunapalari-WA2
albipes-WA1
albipes-WA2
wilsmorei-WA1
sutor-WA
pelobatoides-WA1
pelobatoides-WA2
pictus-Vic1
pictus-Vic2
Notaden
d)
sudelli-SA1
aquilonius-NT
sudelli-SA2
sudelli-QLD1
centralis-SA1
centralis-SA2
centralis-WA1
centralis-WA2
aquilonius-WA1
aquilonius-WA2
fulvus-WA1
fulvus-WA2
fulvus-WA4
fulvus-WA3
kunapalari-WA2
albipes-WA1
albipes-WA2
wilsmorei-WA1
sutor-WA
pelobatoides-WA1
pelobatoides-WA2
pictus-Vic1
pictus-Vic2
kunapalari-WA1
Notaden

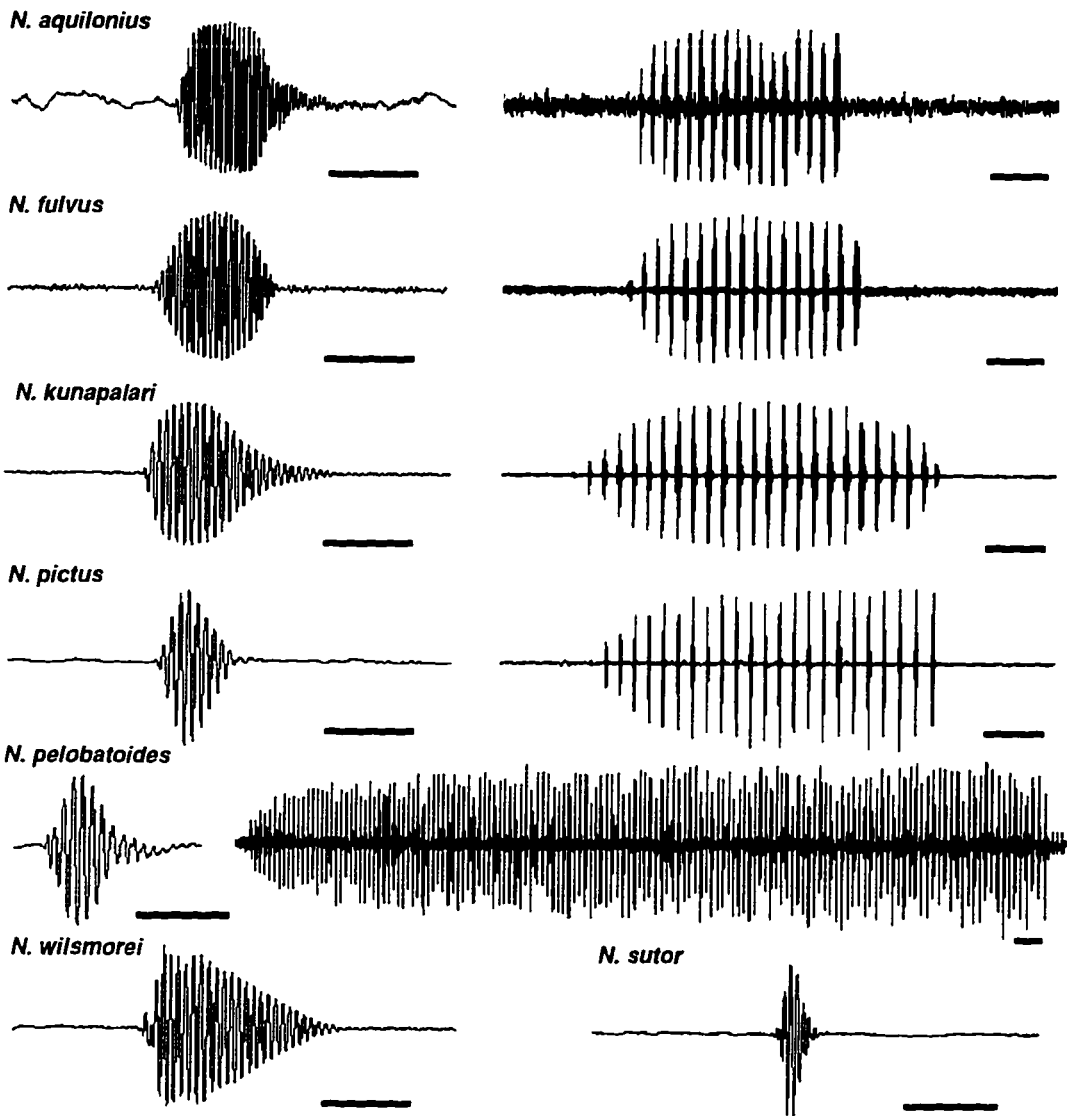
Figure 3.5. Model trees used in a parametric bootstrap analysis (using parsimony) to test the null hypothesis that tetraploids are monophyletic and have thus experienced a single origin. (a) The "best" tree (single most parsimonious tree) found using parsimony on a reduced data set (tree length=753 steps). (b) The "model" tree, found under the constraint of tetraploid monophyly (tree length=794). The difference in the length of these trees (41 steps) was compared to a null distribution of differences generated by comparing the best trees and the best trees under the constraint of monophyly for 100 data sets simulated based on the likelihood estimates for branch lengths, base frequencies, and transition:transversion ratios of the model tree. The largest difference found for the simulated data sets was only 2 steps, so the real difference was interpreted as highly significant, and the hypothesis of monophyly was rejected.



Difference = 41 steps

Conclude: Reject H_0 of tetraploid monophyly

Figure 3.6. Oscillograms of the major mating call types produced by species of *Neobatrachus*. *N. sutor* and *N. wilsmorei* produce single note calls while all other species produce pulsed calls. *N. pelobatoides*, *N. pictus* and *N. albipes* produce calls with a more rapid pulse rate than any of the tetraploids at the same temperature (all species show significant regression of pulse rate on temperature). *N. fulvus* produces calls with a pulse rate more similar to the tetraploids than to any of the diploids (c.f. *N. pelobatoides* and *N. aquilonius*). Although the calls of all tetraploids are similar, *N. kunapalari* produces a call with a slightly faster pulse rate than any of the other tetraploids (c.f. *N. Aquilonius*). Bars represent 10 and 200 ms respectively for pulse and whole call. *N. aquilonius*, 27.0°C; *N. fulvus*, 26.5°C; *N. kunapalari*, 18.0°C; *N. pictus*, 17.9°C; *N. pelobatoides*, 19.0°C; *N. wilsmorei*, 18.0°C; *N. sutor*, 20.4°C. Used with permission of J.D. Roberts.


N. aquilonius
N. fulvus
N. kunapalari
N. pictus
N. pelobatoides
N. wilsmorei
N. sutor

Chapter 4

Nuclear Genes and Polyploid Frogs

Table of Contents

ABSTRACT

The grey treefrogs of North America are a complex of diploid and tetraploid frogs whose evolutionary histories have not been completely resolved. While phylogenies based on mitochondrial gene sequences predict multiple origins of tetraploids and identify at least two independent lineages of diploids in the complex, mechanisms of origin (allopolyploid or autopolyploid) are difficult to resolve without knowledge of paternal contributions. Australian burrowing frogs in the genus *Neobatrachus* also represent a complex of diploid and tetraploid species. Mitochondrial DNA data for this group also suggests multiple origins of the tetraploids but does not provide definitive information on which diploid species were most likely involved in the origination or whether this was a result of hybridization or endoreduplication within a species. Allozyme data have not resolved these questions because there are no fixed differences between diploids and tetraploids in either group. To investigate whether sequences from nuclear genes can provide more information on relationships in these groups of closely related "species", several single-copy nuclear genes were targeted as potential candidates. One gene, EF-1α, gave consistent multiple banding patterns in PCR-mediated amplifications. Gel isolation and sequencing of some of these products revealed that the strongest band was that of a processed pseudogene that did not include the intron sequence. Pseudogenes sequenced from diploid and tetraploid treefrogs were identical and were only 8% divergent from functional gene sequences in *Xenopus*. Comparison with several known pseudogene sequences from *Xenopus* revealed approximately 16% sequence divergence compared to the *Hyla* sequences. Therefore, comparison of the pseudogenes did not appear to be useful for population level analyses. Isolation of

the functional gene was complicated by preferential amplification of the pseudogene and problems with heteroduplex formation. Sequences were obtained from the putative functional gene from the most divergent diploid and tetraploid *Hyla* (based on the mitochondrial DNA analysis) and from the most divergent diploid and tetraploid *Neobatrachus* but revealed no variation within the genus and produced identical sequences within species pairs. Similar problems with multiple banding were also found in another conserved gene, NCAM, that is not known to have pseudogenes in *Xenopus*. However, attempts to isolate the functional gene were unsuccessful. A third gene that is less conserved but is known to vary in a phylogenetically informative way at the protein level was also used (SOD-1). Sequences from this gene revealed evidence of heterogeneity within samples, probably as the result of amplification of different allelic copies of the gene. However, sequencing of clones from this gene revealed that the target intron was not actually amplified. Whether this is the result of pseudogene amplification or absence of the intron in these taxa could not be determined. Finally, published primers were used to amplify an intron region of a member of the aldolase gene family (ALD-1 or ALD-C). Sequences were obtained for several of the *Hyla* taxa and all of the *Neobatrachus* taxa used in an earlier mitochondrial DNA study (one population from each of the nominal species in the group). However, there was no significant variation within the ingroup for either of these taxa but alignment (within the intron) of ingroup and outgroup sequences was effectively a random process. Potential benefits and limitations of population level comparisons using intronic regions of nuclear genes are discussed in light of these results as well as general problems with the use of nuclear genes for phylogenetic analysis.

INTRODUCTION

Studies using segments of mitochondrial DNA as a phylogenetic marker have the potential to identify differences between populations that may provide an indication of degree of relatedness of species or structuring of populations within species. However, because it is maternally inherited (most of the time, in most organisms) it provides a phylogeny of haplotypes of genes rather than a complete picture of genetic relatedness (e.g., Brown *et al.* 1982) . Because of this and the fact that it is nonrecombining, all parts of the molecule share the same historical pattern of common descent (Wilson *et al.* 1985). In addition, the mitochondrial genome, thought to have arisen originally as a prokaryotic parasite of eukaryotic cells, evolves at a much faster rate (Brown *et al.* 1979; 1982) and may have experienced a different evolutionary history than the nuclear genome.

The most common method for comparing nuclear genomes has been the use of gel electrophoresis to compare allozymes of specifically stained proteins. While this method still provides great power for comparisons at the population and closely related species level, it has some limitations. Because it compares gene products rather than actual genes, different allozymes with the same charge or that migrate through the gel at the same rate due to structural configurations may be missed and the degree of divergence may be underestimated (e.g., Murphy *et al.* 1996). It also relies on absolute standardization of conditions so that the same alleles are consistently coded and may not be useful beyond a certain threshold of species differentiation (e.g., Murphy *et al.* 1996). There may also be a problem in interpretation of duplicated gene loci, especially if differences in expression are tied to unseen regulatory loci that affect which products are detectable: apparently silent

alleles may be expressed differentially during development or in different tissue types (e.g., Murphy *et al.* 1996).

The grey treefrogs of North America are a complex of diploid and tetraploid frogs whose evolutionary histories have not been completely resolved. While phylogenies based on mitochondrial gene sequences predict multiple origins of tetraploids and identify at least two independent lineages of diploids in the complex, mechanisms of origin (allopolyploid or autopolyploid) are difficult to resolve without knowledge of paternal contributions. Allozyme data are relatively uninformative in this respect because there are few fixed allelic differences between populations, and diploids and tetraploids share most of the same alleles and patterns of variation could have multiple explanations (see Chapter 2). Australian burrowing frogs in the genus *Neobatrachus* also represent a complex of diploid and tetraploid species. Mitochondrial DNA data for this group also suggest multiple origins of the tetraploids but do not provide definitive information on which diploid species were most likely involved in the origin of the various tetraploid species or whether this was a result of hybridization or endoreduplication within a species. Allozyme data have not resolved these questions because there are no fixed differences between diploids and tetraploids and frequency data have not been collected.

For phylogenetic comparisons to identify ancestry and evolution of tetraploids in each of these groups, it would therefore be desirable to obtain nuclear markers at the actual gene coding level rather than the gene product level. Although ideally comparison of the maternally inherited mitochondrial genome with a paternally inherited gene would be most useful for studies designed to distinguish hybrid origins or to detect current hybridization events, reliable markers for this purpose

have not been identified. Introns in coding sequences of sex chromosomes (in species in which the male is the heterogametic sex), such as the Y-chromosome of mammals (Dorit *et al.* 1995) have been identified as possibilities. However, in humans, no variation was found in a sample of 38 males in an intron of the *ZFY* gene (Dorit *et al.* 1995). Hillis and Dixon (1991) have documented the usefulness of rDNA sequences for phylogenetic comparisons at various levels of divergence. Spacer regions of rRNA genes (e.g., Kambhampati and Rai 1991; Phillips and Pleyte 1991; Lee and Taylor 1992; Pleyte *et al.* 1992; Ritland *et al.* 1993; Schlotterer *et al.* 1994; Vogler and DeSalle 1994; Vilgalys and Sun 1994) have been used recently for closely related species comparisons in several taxonomic groups. In polyploid plants, areas of repeated nucleotide sequences have been used to predict the origins of tetraploids (Zhang and Dvorak 1991). However, use of genes from the ribosomal gene array for reconstruction of species histories relies on the assumption that concerted evolution occurs to homogenize sequences among copies within individuals so that homology of repeat units is not questioned.

H. versicolor has less than twice as much DNA as its diploid progenitor (1.9 X) (Bachman and Bogart 1975) but twice as much rRNA based both on nucleolar counts (Cash and Bogart 1978) and absolute number of rRNA genes (Toivonen *et al.* 1983). There is also evidence that considerably more base substitutions have occurred in the tetraploid rDNA genes than in the diploid (based on thermal stability of homologous and heterologous hybrids) due to either a relaxation of the gene regulatory "correction" mechanism hypothesized to be responsible for the maintainance of identical tandem rRNA genes in the tetraploid or a release of one gene set from the normal selective constraints (Toivonen *et al.* 1983). Restriction site

mapping of regions of the 18s and 28s repeat units has suggested that tetraploids may contain elements of two distinct types of diploids (eastern and western) that may reflect a hybrid origin. However, the appearance of unique restriction sites within individuals suggested that complete homogenization of copies may not have occurred. The utility of repeated DNA elements for the study of polyploids may therefore be inhibited by potential risk of comparing paralogous rather than orthologous gene copies.

Although comparisons using coding regions of nuclear genes have been used for deep level phylogenetic approaches, genes useful for the comparison of closely related species have not been as abundant. Intron-containing regions of single-copy nuclear genes have been used only recently to compare closely related groups of taxa (Lessa 1992; Lessa and Applebaum 1993; Slade *et al.* 1993, 1994; Palumbi and Baker 1994) but an extensive data base has not yet accumulated. The neutral theory (reviewed in Kimura 1983) or nearly neutral theory (Ohta 1992) predicts that introns should be highly variable because they are nonfunctional and thus free of selective constraints that would be apparent in the flanking exons. Comparative databases using this approach have been slow to accumulate but some literature now exists (see Palumbi 1996).

Lessa (1994), Lessa and Applebaum (1994), Palumbi and Baker (1994) and Slade *et al.* (1993; 1994) have recently described an approach for using multiple single copy nuclear genes for comparison of closely related taxa that has been called exon-priming intron crossing (EPIC) PCR (see also Palumbi 1996; Dowling *et al.* 1996). The principal approach is that intronic regions of even highly conserved genes should have sufficient variability for the study of closely related taxa. The

advantage of using conserved genes is that primer design based on published cDNA (i.e., transcribed coding regions) should enable construction of primers that would be useful across a broad range of taxonomic groups.

The purpose of this chapter is to outline attempts to identify nuclear gene markers that might be useful at the population and closely related species level in diploids and tetraploids of the grey treefrogs of North America and Australian burrowing frogs in the genus *Neobatrachus*.

METHODS AND RESULTS

Details for PCR, cloning and sequencing protocols used are described in Hillis *et al.* (1996a). A general outline of the techniques employed will be discussed in this chapter along with specific results and detailed protocols where variable techniques were used.

Single-copy Nuclear Genes: EF1α

A genebank search identified elongation factor 1-α (EF-1α), an important gene in the DNA replication process, as one of the most conserved genes for which sequences of the coding region are available for a diverse array of organisms (i.e., from *Artemia* through primates). Primers were constructed by choosing the most conserved 20-bp segment within 50-100 bp on either side of an intron region. Primers were designed to be complementary to regions in exons 3 and 4 which flank the mammalian intron c. Primer sequences were: 5'-GGA AAT TTG AGA CCA GCA ART ACT ATG T-3' and 5'-TCA AAT TCA CCA ACA CCA GCA GCA AC-3'.

A preliminary PCR screen using these primers resulted in amplification of gene products from all taxa attempted from urochordates through fish, amphibians and mammals. However, each species type produced multiple size PCR products that differed between groups, indicating possible differences in intron length and multiple copies within individuals (Figure 4.1). Sequencing of asymmetric PCR products of the smallest size band in diploid and tetraploid grey treefrogs suggested that a small 150 bp band that was common to all species amplified was probably a processed pseudogene lacking intron sequences. EF-1α is known to have several pseudogenes associated with it in *Xenopus*. An 88-bp fragment of this presumptive pseudogene was sequenced using the forward primer. The sequence was unreadable 3 bp upstream from the start of the intron in the functional gene and readable sequences were not obtained using the reverse primer. The degree of divergence between one of the published *Xenopus* pseudogene sequences (XELEFPSUD1) and the sequenced region of the presumed *Hyla* pseudogene was approximately 16% and the degree of divergence of the *Hyla* pseudogene from the published *Xenopus* functional copy was approximately 8% based only on nucleotide substitions (Figure 4.2). A 7-bp region at the end of the *Hyla* pseudogene sequence appeared to contain a paired insertion and deletion; inclusion of these changes would raise the estimated sequence divergence to 10% and 18%, respectively. The *Xenopus* pseudogene sequence used was 10% divergent from the *Xenopus* functional gene over the region compared for the *Hyla*. Comparison of more extensive regions between these two gene copies indicated that insertions and deletions were common. It is not possible to determine which of the published *Xenopus* pseudogene sequences might be homologous to the putative pseudogene found in *Hyla* so this is probably not a very

realistic comparison but it does give a general idea of relative rates of change between functional and putatively nonfunctional copies. Sequences of pseudogenes from diploid and tetraploid grey treefrogs were identical across the region sequenced.

Gel isolation and reamplification of the largest band (approximately 300 bp) produced in PCR amplifications (thought to represent the intron-containing gene) lead to reformation of the pseudogene and several intermediate bands. This was interpreted to be due to heteroduplex formation between the pseudogene and the protein coding gene as has been found for such genes as actin and FcRIIb (Zorn and Krieg 1991). Alkaline treatment of PCR products followed by gel isolation from denaturing gels using either 6% urea polyacrylamide gels (Zorn and Krieg 1991) or alkaline agarose gels (Valentine *et al.* 1992) has been found to circumvent problems associated with heteroduplex formation. Both of these methods were attempted to isolate the functional gene copy (e.g., Figure 4.3a). The amounts of DNA that could be produced by these methods was insufficient to allow asymmetric amplification of products to be used for dideoxy sequencing (Figure 4.3b), partly because of preferential amplification of the pseudogene compared to the larger sized amplification products.

Sequences from the presumptive functional gene were finally obtained by using cycle sequencing, which uses end-labeled primers in a linear amplification procedure of double-stranded DNA that can result in readable sequences of even small amounts of template (Murray 1989; Craxton 1991). With this technique, there is no need for generation of a single-stranded product prior to sequencing, and enough template can be generated from a gel-isolated band from a single 25-µl PCR reaction to result in readable sequences (see Hillis *et al.* 1996a). Readable sequences

were obtained only from the forward primer and extended approximately 75 bp into the intron region from the 5' direction. No variation was found within this region of EF-1α between diploid and tetraploid *Hyla* but intron regions were nearly impossible to align between the *Hyla* and *Xenopus* sequences. Coding regions of the *Hyla* pseudogene sequence appeared to be identical to the intron-containing copy except for the region at the end of the pseudogene containing the insertion and deletion mentioned above. The coding region of the *Hyla* functional gene was 5.5% divergent from the *Xenopus* functional gene, 13% divergent from the *Xenopus* pseudogene, and 2% divergent from the corresponding region in the *Hyla* pseudogene (Figure 4.4). Partial sequences were obtained from a tetraploid *Neobatrachus* (*N. centralis* from Western Australia) that appeared to be the right size to contain an intron but sequences were only readable over a 60 bp region in the coding region. This segment had three nucleotide substitutions (5%) compared to the *Xenopus* functional gene and four nucleotide substitions (7%) compared to the *Hyla* functional gene. Additional sequencing was not attempted given the difficulty of obtaining sequences from the *Hyla* functional genes and the limited variation present.

Single-copy Nuclear Genes: NCAM, SOD

A second set of primers was designed to amplify another highly conserved single-copy nuclear gene that is not thought to have associated pseudogenes in *Xenopus* (Neural Cell Adhesion Molecule--NCAM) (Paul Krieg, personal communication). Primers were designed initially with degenerate sequences to allow application to a wide variety of species groups. However, attempts to amplify products from *Rana*, *Xenopus* and *Hyla* were unsuccessful. A second set of primers

were constructed with less degeneracy. Amplification products were produced but were of variable size and did not amplify consistently in any of the taxonomic groups attempted. Although a variety of conditions and taxa were attempted, sequences were not obtained from this region.

A third set of primers was obtained from Leslie Rye at the University of Guelph, which were designed to amplify a region of superoxide dismutase (SOD) that spanned two introns (3 and 4) of the human Cu-Zn form of the gene (e.g., Levanon *et al.* 1985). Primer sequences were: SOD-a: 5'-GCT GTA CCA GTG CAG GTC CTC ACT TTA AT CC-3' and SOD-b: 5'-TCA AAT TCA CCA ACA CCAGCA GCA AC-3'. Two major size classes were amplified in most taxa for which PCR was attempted: a 350-bp fragment and an 800-bp fragment. Fragments of each size were gel-isolated and sequenced using both automated sequencing (following fluorescent end-labelling using cycle sequencing) and direct cycle sequencing. Sequences from the 350-bp fragment resulted in ambiguous sequences, suggesting that more than one allele or locus was amplified within individuals. As well, sequences from the 350-bp fragment appeared to be completely different from the 800-bp fragment and did not appear to correspond to the Cu-Zn SOD for which the primers had been designed (Figure 4.5). Application of individual samples from gel-isolated bands to denaturing gradient gel electrophoresis (DGGE), a technique which is capable of resolving single base-pair differences between samples in fragments of up to 600 bp (see Hillis *et al.* 1996a; Dowling *et al.* 1996 and references therein), confirmed that the 350-bp PCR fragments contained at least two different sequences while the 800-bp fragment appeared to be homogeneous (Figure 4.6).

Both the large and small fragments from several taxa were cloned in an attempt to separate individual sequences and to get a clearer idea of how primers might be redesigned to result in more specific amplification of particular gene copies that would include introns. It was thought that designing primers to cross the intron-exon boundary might result in less ambiguous amplification. An internal primer between the two introns was also designed (SOD-3: 5' GCC CAG GTC ATC TGG TTT TTC ATG GAC CT 3') but did not result in consistent amplification. Cloning was performed using the blunt-end cloning method described in Hillis *et al.* (1996a) using a BlueScript vector and colonies initially were screened for presence of inserts using X-gal colour selection (Berger and Kimmell 1987). White (positive) colonies were characterized for insert size by extracting the plasmid + insert DNA using a standard alkaline-lysis miniprep procedure (e.g., Hillis *et al.* 1996a), excising the insert with an appropriate combination of restriction enzymes (e.g., *Eco*RI and *Hin*dIII) and comparing migration on minigels next to a standard size ladder and uncut bluescript DNA. An alternative method to mini-preps utilizing PCR to amplify and screen for multiple size fragments was also used as a quick screening procedure. This method allows much better separation of size fragments on agarose minigels and does not require the use of restriction enzymes to screen colonies. Colonies were picked and amplified using Universal (5' CAT TTT GCT GCC GGT CA 3') and Reverse (5' AAC AGC TAT GAC CAT G 3') primers designed for BlueScript and fragments of the correct length were sequenced using either internal vector primers (e.g., T3: 5' ATT AAC CCT CAC TAA AG 3' and T7: 5' TTA TGC TGA GTG ATA TC 3'), original PCR product primers, or the Universal and Reverse vector primers.

Positive clones obtained from both size classes were sequenced but the 350-bp band still could not be aligned to the human Cu-Zn SOD sequence and the 800-bp band did not appear to contain introns. Further sequencing of this gene was therefore abandoned.

Single-copy Nuclear Genes: Anonymous single copy genes

The anonymous single copy nuclear DNA (scnDNA) approach of Karl and Avise (1992) was also attempted as a means of isolating single-copy genes for comparison. Approximately 1 µg of total genomic DNA from a diploid grey treefrog from Bastrop County, Texas, was cut to completion with *Eco*RI (4.5 hours at 37°C). It was found that the best results were obtained if the DNA was digested for 2-4 hours and then more enzyme added for an overnight digestion. A small volume (5 µl) of cut DNA was run on a 0.5 % agarose gel using a 1X TAE buffer next to an equal volume of uncut DNA to ensure that digestion was complete. The remaining digestion was precipitated with 2 *M* NaCl and 2.5 volumes of 95% EtOH and rehydrated to a concentration of approximately 0.1 µg/µl. A total genomic library was constructed by ligation of the cut DNA to Lambda Zap prepared arms that had been precut with EcoRI (Lamda Zap II/EcoRI/CIAP cloning Kit: Stratagene) according to the kit instructions. An approximately equal molar ratio of insert to vector arms was used in the ligation based on the average size predicted for *Eco*RI fragments (e.g., based on a predicted average fragment size of 5 kb, approximately 0.15 µg of digested DNA was used in the ligation). Ligation was performed overnight at 14°C. Packaging of the ligation was performed according to the manufacturer's protocol and packaged total genomic libraries were stored at 4°C in

500 µl SM buffer and 20 µl chloroform. XL1Blue MRF- cells obtained from Stratagene were grown to log phase in LB supplemented with 0.2% maltose and 10*mM* MgSO4 for 6-8 hours with shaking at 37°C. Cells were sedimented by centrifuging at 4000 x g for 10 min at 4°C and resuspended in 10 *mM* $MgSO_4$ to a final optical density of 0.5. Cells (200 µl) were infected with packaged phage (2 µl) and placed at 37°C for 15 minutes. The mixture was then added to melted top agar and poured onto unsupplemented LB plates with X-Gal. Plates were incubated at 37°C overnight. Colonies were screened through IPTG/X-Gal colour selection. For screening Lambda Zap plaques, a higher concentration of IPTG/X-Gal is required (25 µl of 10 mg/ml IPTG and 500 µl 25 mg/ml X-Gal) than for plasmid screening (Stratagene) and colour selection is not as efficient. Positive clones (white plaques) were cored with a sterilized Pasteur pipette and plasmids excised according to the manufacturer's directions using XL-1 Blue cells and ExAssist helper phage to produce a plasmid packaged as filamentous phage particle, which could be stored for up to 2 months at 4°C. To plate the rescued phagemid, SOLR cells (200 µl) were transformed with 1-15 µl (depending on desired cell density) of phage stock and grown overnight at 37°C. Colonies were screened using IPTG/X-Gal (15 µl of 10 mg/ml IPTG; 40 µl of 25 mg/ml X-Gal) and "positive" colonies were screened for inserts using alkaline/lysis minipreps and restriction digests. Miniprep DNA containing genomic inserts were then screened for relative copy number using a Slot-Blot apparatus (BioRAD Dot blot) which allows standardization of DNA quantity applied to a hybridization filter. The principle was that hybridization to a random primed genomic probe should indicate the relative copy number of genes through the intensity of the autoradiograph. The apparatus uses a vacuum manifold to suck

samples directly onto a nylon filter (in 2 volumes of 20X SSC). Filters were dried, denatured for 2 minutes in 0.5 *N* NaOH/1.5 *M* NaCl, neutralized for 3 minutes in 0.5 *M* Tris/3 *M* Na Cl and rinsed in 5X SSC prior to cross-linking and baking at 80°C for 2 hours. Filters were hybridized (in 5X SSC; 0.5% SDS; 5X Denhardt's and 25mM $KP0_4$) with a random primed probe constructed by labelling 50-100 ng of total genomic DNA (from the sample used to construct the library) with α^{32}P-dCTP using Klenow fragment and random primers (see Hillis *et al.* 1996a). Hybridization was performed at 65°C overnight, after which the filter was rinsed in 5% SSC/1% SDS until background radiation dissipated. Filters were wrapped in Saran wrap and placed on an autoradiograph film (with intensifier screens), which was exposed for 24-48 hr at -80°C. Initial attempts were limited by low efficiency of insertion into the plasmid, and several reattempts at library construction were made. Although several apparently low copy number clones were identified, none appeared to be present in single copy and sequencing attempts (using plasmid primers) from miniprep DNA were unsuccessful. Given these difficulties, and the labour intensive screening involved in this procedure, the project was abandoned before the final stage: sequencing clones for primer design to be used in PCR amplifications. There was also a concern that problems with establishing homology would complicate results, especially when using polyploid individual.

Aldolase

Sequences were obtained from diploid and tetraploid *Hyla*, diploid and tetraploid *Neobatrachus* and *Xenopus* using published primers (Lessa 1992; Lessa and Applebaum 1993) designed to amplify a region surrounding intron 5 of aldolase

genes. These primers have resulted in amplification of two to three different forms of aldolase genes in other taxonomic groups (Slade *et al.* 1993). A 400-bp fragment that aligned to the ALD-1 fragment in lizards (ALD-C of humans) amplified most consistently but a 650-bp fragment appeared in some of the taxa. Figure 4.7 shows an alignment of sequences obtained from *Xenopus*, a diploid and tetraploid *Hyla*, a diploid and tetraploid *Neobatrachus*, and published skink sequences for the smaller band (Slade *et al.* 1993). The sequence for the diploid hylid appeared to be either an unidentifiable contaminant (a comparable sequence was not found in a search of genebank) or a different gene copy than was sequenced for the other taxa. In the coding region for aldolase there appeared to be about 15% divergence in a 113-bp sequence between a tetraploid *Hyla* and *Xenopus* but 42% sequence divergence within the first 36 bp of the intron, suggesting that similarities were more likely due to random chance in this region. Sequence variation among *Neobatrachus* samples in the coding region was also small but 20% divergent compared to *Xenopus*. Variation within the intron was characterized by insertions and deletions between genera but mainly nucleotide substitutions within genera.

For the Australian frogs, sequences were obtained from the smaller of the two amplified sequences (ALD-C) from a single representative from most of the *Neobatrachus* taxa and from two closely related genera chosen as "outgroups" (*Notaden melanoscaphus* and *Heleioporus psammophilus*). Sequence divergence was low and a heuristic search (simple stepwise addition followed by TBR branch swapping) using unweighted parsimony (PAUP* 4.0d24: Swofford 1996) did not result in much resolution within *Neobatrachus* but the introns of the two outgroups were distinctly different from the ingroup sequences (Figure 4.8). The two

Heleioporus samples appeared to be much more similar to those of *Notaden* than to *Neobatrachus*. A bootstrap analysis gave strong support for this grouping. This finding is of interest because sequences from mitochondrial genes (cytochrome oxidase 1) suggested that inclusion of these individuals resulted in nonmonophyly within *Neobatrachus* (Figure 4.9). The implications of these results are either that the mitochondrial findings were an artifact of the high levels of sequence divergence exhibited (i.e., homoplasy or parallel evolution) or that tetraploids within *Neobatrachus* may have arisen through hybridization between *Heleioporus* and one of the *Neobatrachus* diploids. This result needs to be confirmed by verification of specimens used (i.e., examination of voucher specimens, sequencing of further populations of *Heleioporus pssamophilus* and other species of *Heleioporus*) and by further aldolase sequencing.

Attempts were made to isolate and sequence the second aldolase locus to provide more characters for analysis (see Slade *et al.* 1994) but attempts to amplify this region were unsuccessful. Long-range PCR (Barnes 1994) was also used to try to improve amplification of the larger bands. The principle was that by combining thermostable polymerases with and without proofreading activity (i.e., 3'→5' exonuclease activity), larger amplification products might be less prone to incomplete formation resulting from early replication errors. Combinations of an enzyme without proofreading activity (i.e., *Taq*) and one with proofreading activity (i.e., one of *Vent*, *Deep Vent* or *Pfu*) were evaluated for effectiveness at varying relative ratios. Enzyme buffers were also varied. However, none of the combinations used resulted in improved amplification of larger sized products.

DISCUSSION

Rates of Evolution in Intron Regions

Sequencing of the pseudogene for EF-1α between diploid and tetraploid *Hyla* and diploid and tetraploid *Neobatrachus* revealed no differences within groups and only 10% sequence divergence between *Xenopus* and *Hyla*, suggesting a much higher degree of conservation than would be predicted by models of neutral evolution (e.g., Kimura 1983; Ohta 1992), which assume that pseudogenes should be functionally neutral (e.g., Nei 1987). However, other studies have suggested nonrandomness of point mutations in pseudogenes (e.g., Li *et al* 1984) suggesting that some type of selection or constraint is acting. Deviations from expected neutrality have also been described for a pseudogene of ADH in *Drosophila* (Sullivan *et al.* 1994). In certain groups of parasites, such as trypanosomes, it has been found that intragenic recombination leading to the formation of expressed mosaics (which may alter epitopes as a means of hiding from the host's immune system) can occur not only within and between functional gene copies of multigene families, but can also involve pseudogene sequences (e.g., Donelson *et al.* 1985; Kamper and Barbet 1992). In fact, some variable surface antigen genes appear to be entirely constructed of pseudogenes in some species (Roth *et al.* 1989). Similarly, in chickens, immunoglobulin diversity is generated by gene conversion of the single functional gene by many variable region pseudogenes (Ota and Nei 1995). It may be, therefore, that pseudogenes are not simply "junk" but may provide additional sources of genetic material for increasing diversity. Predictions of neutrality should not be assumed without evidence that selection does not control patterns of evolution. Comparison of the rates of synonymous versus nonsynonymous nucleotide substitutions has been

used to search for evidence of selection (e.g., Hughes, 1993; Hughes and Hughes 1993; Ota and Nei 1995) and may provide a useful starting point for establishing whether the genes of interest might be constrained in a phylogenetically uninformative way.

Sequences from an intron-containing copy of EF-1α also revealed 100% sequence similarity within the *Hyla* group but extensive sequence divergence to the point of random similarity between *Hyla* and *Xenopus* introns. This could suggest that intron structure may also be more important than previously realized. Recent literature has identified specific functions for introns (e.g., Ableson 1992; Baldi *et al.* 1992; Sollner-Webb 1993) that may result in constraints that would also cause deviations from a purely neutral rate of evolution. The finding that intron sequences of human X-linked colour vision genes were found to be much less divergent than the exon sequences for two gene copies was explained by gene conversion, because neutrality was assumed (Shyue *et al.* 1994). The intron region sequenced from aldolase also showed less variation than might be expected within groups (although relatively more variation was found than for EF1α) and large jumps between groups. It is interesting that in a gene for which more variation existed in the coding region between taxa (15% vs 5.5% between *Hyla* and *Xenopus* for Aldolase and EF1α, respectively), there also appeared to be more variation in the intron region. If intron regions were completely neutral, this type of difference would not be expected: all introns should evolve at a rate equivalent to the rate of mutation in the genome. It will be interesting to examine divergence of more intron-containing regions as the database of comparative studies increases.

Homology

An important problem with the study of nuclear genes that was consistently made apparent in this study is the difficulty in establishing homology when multiple gene copies or multiple alleles at a single locus might be present (e.g., Hillis 1994; Moritz and Hillis 1996; Dowling *et al.* 1996). When additional copies are of a distinct size class from the target sequence, these problems may be overcome without too much difficulty. For example, pseudogenes of EF-1α completely lacked introns and therefore could be isolated from the putatively functional copy that contained introns. In addition, heteroduplex bands forming as a result of amplification of the two "gene" types could also be identified and avoided. For aldolase, Slade *et al.* (1993) were able to separate individual copies of the gene family by isolating unique size class bands and did not experience major difficulties with heterogeneity within size classes. However, all studies that have used single-copy nuclear genes have stressed that sequencing of multiple genes is a very labour intensive undertaking and some sort of screening measure (i.e., cloning or a rapid screening method such as DGGE or heteroduplex analysis) is necessary to establish homology (Lessa 1992; Lessa and Applebaum 1993; Slade *et al.* 1993; Slade *et al.* 1994; Palumbi and Baker 1994).

Sequences from aldolase for the frogs used in this study suggested that one of the samples sequenced (diploid hylid) was different enough from those of the other taxa that it either represented a contaminant or an additional copy of the gene that was the same size as the target sequence. Because there is only a limited database on this gene and no other frog sequences were available in gene banks, these alternatives could not be distinguished. This finding questions the validity of the rest of the

sequences compared even though variation seemed to be low enough that it was likely that the same gene copy was sequenced. The extent to which multiple copies might be a problem could be investigated by sequencing of multiple clones per individual but this may not be worth the effort considering the limited variation present within the taxa of interest. However, without this information, conclusions based on phylogenetic analyses using these data can only be made tentatively because it cannot be determined if the necessary assumption that orthologous genes are being compared among taxa (see Hillis 1994; Moritz and Hillis 1996; Dowling *et al.* 1996; Hillis *et al.* 1996a) has been violated. Unfortunately, economic feasibility is also a factor in the utility of nuclear genes.

Problems with homology were more clearly identified in the study of SOD sequences. Heterogeneity of sequences within size classes was apparent from ambiguities in sequences and was confirmed using a rapid screening method (DGGE). Sequences from the two size classes isolated may have represented completely different gene copies and there was evidence of heterogeneity within size classes. The finding that the region sequenced apparently lacked intron regions also introduces another question of homology. In the limited studies that have used a comparative approach in the characterization of intron-exon boundaries, extensive variation in the size and structure of introns has been found. For example, for the Cu/Zn form of SOD, Kwiatowski *et al.* (1992) suggested that intron insertions and deletions have occurred in the evolution of the gene in various animal phyla, plants and fungi. In many cases it cannot be predicted a priori whether a particular taxon of interest will actually contain an intron in the target region and if sequences are consistently amplified without introns, it could be difficult to establish whether the

sequence amplified represents a "nonfunctional" pseudogene or the functional gene with reduced or excised intron regions compared to the taxa from which primers were designed. In either case, it is probably best not to use the particular gene for phylogenetic analysis rather than risk violating assumptions of homology.

Pseudogenes technically could be compared among taxa but there is a risk that independent duplication of pseudogenes or independent insertion of pseudogenes from reverse transcription of the ancestral functional gene product could result in the comparison of paralogous rather than orthologous sequences for these as well. It is becoming increasingly evident that pseudogenes, rather than being due to rare duplications or insertion events, may be a significant component of the genomes of many organisms and Ohno (1985) suggested that their abundance may provide evidence of the efficacy of gene duplication as a means of acquiring new genes with novel functions. Recent evidence suggests that nuclear copies of mitochondrial genes are much more common than originally assumed and can provide significant problems for the study of mitochondrial genes if sequences obtained from total genomic DNA are not compared with purified mitochondrial copies (e.g., Zevering *et al.* 1991). In studies of immune invasion by parasitic organisms it has been suggested that the involvement of pseudogene copies in multigene families in the formation of mosaic genes may be an important mechanism to increase the epitope repertoire of certain protein coat genes (e.g., trypanosomes: Kamper and Barbet 1992). The genetic information contained in pseudogenes in these cases may not be the evolutionary dead ends that they have been assumed to be. The extent to which pseudogene copies may be used in recombination in "higher" eukaryotes is not known but pseudogenes have been implicated as a means of generating

immunoglobulin diversity in chickens (Ota and Nei 1995). Caution should therefore be exercised in the use of these genetic elements with unknown evolutionary histories for phylogenetic comparisons.

The ribosomal gene array has been used as a nuclear marker for phylogenetic comparison at various levels of divergence (e.g., Hillis and Davis 1987; Hillis and Dixon 1991). Although there are multiple copies of the gene repeated in tandem as well as on different chromosomes, there appears to be homogenization of sequences across copies that has been termed "concerted evolution" (Hillis and Davis 1988; Hillis *et al.* 1991). In the study of the more slowly evolving regions in the rDNA array, this homogenization has been documented to occur relatively completely, even on rRNA genes concentrated in different nucleolar organizing regions (NORs). However, evidence from studies of tetraploid plants (Delseny *et al.* 1990) and amphibians (Romano and Vaughn 1986) suggests that heterogeneity may exist between repeat units within individuals and that concerted evolution may not result in 100% identity of repeats of every set of rDNA genes in these instances. In addition, amphidiploid plants have been found to contain rDNA repeats from both parental species, probably maintained at NOR positions unique to each parental type for *Brassica* (Delseny *et al.* 1990). However, in other plant species, artificial hybridization studies have indicated that homogenization occurs to one or the other parental species (Wendel *et al.* 1995). In grey treefrogs, restriction site polymorphisms across the rDNA array suggested that eastern and western forms of the diploid (*Hyla chrysoscelis*) were identical in coding regions but had several unique restriction sites in the nontranscribed spacer region (NTS). Tetraploids (*Hyla versicolor*) from a different part of the range than either diploid showed

approximately 50% of the restriction sites that characterized each diploid group, suggesting that they had formed through hybridization (Romano and Vaughn 1985, 1986). In addition, tetraploids apparently showed restriction site polymorphisms within individuals in the coding regions of the 18s gene. While variation among repeat units could be due to a relaxation of normal evolutionary constraints and higher rates of nucleotide substitutions in tetraploids (e.g. Toivonen *et al.* 1983), these results suggest that, even in genes for which homogenization of gene sequences occurs in most organisms, the assumption of homology of repeat units should still be tested through screening of clones. Because most studies using ribosomal genes initially were restricted to cloning studies, this should not have been as much of a problem as the current reliance on PCR, which may not allow immediate detection of deviants present in low copy numbers.

Techniques such as RAPDs and the anonymous single-copy nuclear gene approach attempted here may be even more prone to violations of assumptions of homology. Hare *et al.* (1996) suggested that PCR-based DNA-level assays such as the restriction site analysis of anonymous single-copy nuclear gene approach of Karl and Avise (1993) can result in locus-specific "artifacts" that may be due to polymorphisms at PCR priming sites and resultant amplification from multiple paralogous loci or from incomplete digestion when performing restriction site screening. They suggest that altered PCR conditions or redesigning primers can improve this situation but that it may not be possible to completely bypass the problem of paralogous gene amplification. Careful screening is therefore required to reduce the risk of nonorthologous comparisions.

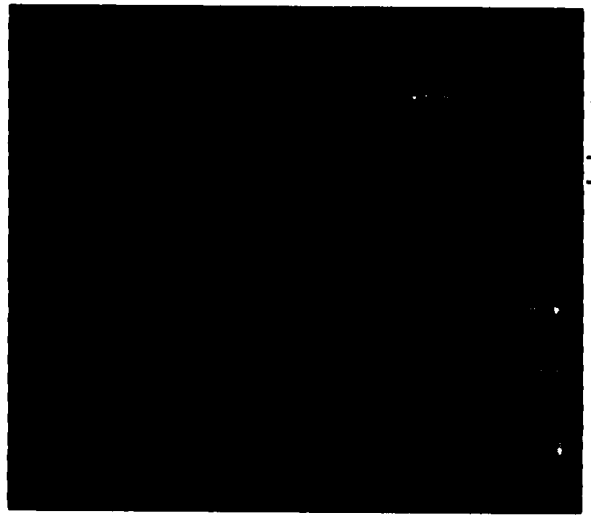
CONCLUSIONS

Attempts to identify a nuclear gene region that to compare populations and closely related species of two groups of polyploid frogs were unsuccessful. However, several general conclusions can be drawn about the approach that must be taken if nuclear genes are to be used for phylogenetic comparisons at this level of taxonomic divergence. The most important conclusion is that homology should not be assumed but requires careful screening to ensure that sequences compared are orthologous. Rapid screening methods such as heteroduplex analysis, SSCP (Orita *et al.* 1989; Hayashi, 1991a, b), DGGE (Myers *et al.* 1986, 1989), and TGGE (Campbell *et al.* 1995) to assess sequence variation within and between individuals prior to sequencing (reviewed by Lessa and Applebaum 1993; Slade *et al.* 1993; Dowling *et al.* 1996; Hillis *et al.* 1996a) are the most efficient way to reduce wasted effort. However, while these techniques are relatively inexpensive and are designed to detect up to single base pair changes, optimization for each species examined and for each gene region is still required. Therefore, the study of nuclear genes can be expected to require a much more labour intensive approach than an equivalent study of mitochondrial genes, especially until the currently small comparative database on nuclear genes increases to an extent that proven primers are readily available.

Finally, studies of this type, in addition to providing markers for phylogenetic comparison, should contribute to the general body of knowledge on how evolution at the molecular level proceeds.

Figure 4.1. Demonstration of amplification fragments produced using EF-1α primers on a range of taxa. (a) Note that while all taxa demonstrated multiple banding patterns, band sizes varied among taxa. (b) Using low melting point gels, the degree of resolution of bands was improved. Notice that the *Hyla* samples each produced three predominant bands. The smallest band was thought to represent a processed pseudogene. The two larger bands represent the functional gene and a heteroduplex band formed between the functional gene and the pseudogene.

1% HMP


154
200
394
516
Hyla 2n
Hyla 4n
H. cinerea
Rana
Eurycea
Bufo
Amphipod
Fish
Control

1.2% LMP

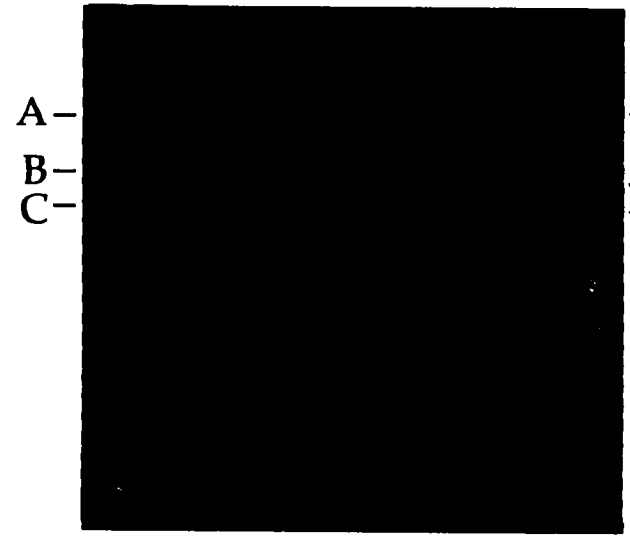

A
B
C
220
344
516
Hyla 2n
Hyla 4n

Figure 4.2. Alignment of putative EF-1α pseudogene sequences from *Hyla* with published sequences from the *Xenopus* functional gene and one of the several known *Xenopus* pseudogenes.

EF-1α Pseudogene

	2260			2270			2280			2290			2300			
Xenopus	ATT	GAC	ATC	TCC	CTG	TGG	AAA	TTT	GAG	ACC	AGC	AAA	TAC	TAT	GTC	ACT
Xensud1	...	A..	...G	...	...	...	...	...	...	...	...	...	...	...	.AA	...
HYLA 2n	---	---	---	---	---	--.	...	...	...	...	...	..G	...	...	...	..C
HYLA 4n	---	---	---	---	---	--.	...	...	...	...	...	..G	...	...	...	..C
	ile	asp	ile	ser	leu	trp	lys	phe	glu	thr	ser	lys	tyr	tyr	val	thr

	2310			2320			2330			2340				2350		
Xenopus	ATC	ATT	GAT	GCT	CCT	GGA	CAC	AGA	GAT	TTC	ATC	AAG	AAC	ATG	ATC	ACT
Xensud1	...	...	...	.A.	..A	...	...	...	..C.		..T..	..T	...	...	...	...
HYLA 2n	...	...	...	..C	...	...	...	...	..C	...	...	...	...	...	..T	...
HYLA 4n	...	...	...	..C	...	...	...	...	..C	...	...	...	...	...	..T	...
	ile	ile	asp	ala	pro	gly	his	arg	asp	phe	ile	lys	asn	met	ile	thr

	2360	*INTRON C*					
Xenopus	GGT	ACT	TCT	CAG	GTAGGG	TTTTT	TTCTT
Xensud1	.C.	...	.A..	..	.----	------	-----
HYLA 2n	...	..A	CTA	GCT	-----	------	-----
HYLA 4n	...	..A	CTA	GCT	-----	------	-----
	gly	thr	ser	gln			

Figure 4.3. Improved isolation of bands produced in EF-1α amplification using denaturing PAGE. (a) Note that three bands can now very clearly be distinguished. Bands were cut from these gels and used in PCR reamplifications. (b) Denaturing PAGE when larger concentrations of DNA were added to the lanes to allow better extraction of DNA from gel bands. Unfortunately, while resolution was much improved with the concentration used in (a), insufficient amounts of DNA were present to allow reamplification from these bands.

Xenopus.

Denaturing PAGE

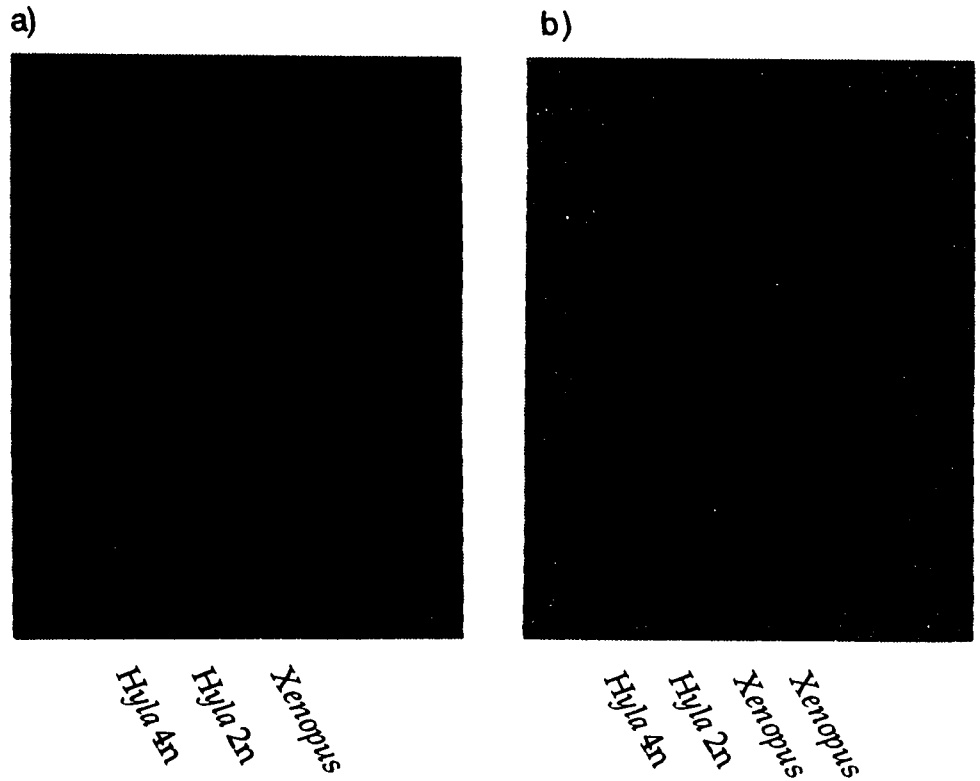


Figure 4.4. Alignment of EF-1α sequences. Alignments are shown for the *Hyla* pseudogene with the functional *Xenopus* gene, part of the *Neobatrachus* coding region, and a short section of the *Hyla* functional gene. In the intron, diploid and tetraploid *Hyla* sequences were identical but were nearly impossible to align to *Xenopus* sequences. Intronic sequences for *Neobatrachus* were not obtained.

```
                2260            2270            2280            2290            2300
Xenopus         ATT GAC ATC TCC CTG TGG AAA TTT GAG ACC AGC AAA TAC TAT GTC ACT
Hyla P          --- --- --- --- --- --. ... ... ... ... ... ..G ... ... ... ..C
Hyla.F          --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C
Neo F           --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C
                ile asp ile ser leu trp lys phe glu thr ser lys tyr tyr val thr

                2310            2320            2330            2340            2350
Xenopus         ATC ATT GAT GCT CCT GGA CAC AGA GAT TTC ATC AAG AAC ATG ATC ACT
Hyla P          ... ... ... ..C ... ... ... ... ..C ... ... ... ... ... ..T ...
Hyla F          ... ... ... ..C ... ... ... ... ..C ... ... ... ... ... ..T ...
Neo F           ..T ... ... ... ... ..C ... ... ..C ... ... ... ... ... ... ...
                ile ile asp ala pro gly his arg asp phe ile lys asn met ile thr

                2360            INTRON C        2380            2390            2400
Xenopus         GGT ACT TCT CAG GTAGG GTTTT TTTCT TATTG GTTTT AGCAT AAACA TCCAT
Hyla P          ..A CTA GCT --- ----- ----- ----- ----- ----- ----- ----- -----
Hyla F          ... ... ... ... ...A. C AGA .G.TA A..CT T.A.A .AT.. ..T.T ATGGA
                gly thr ser gln

                2410            2420            2430
Xenopus         AAAAT AAAAG GGGAA AAAAT AAGTA AAATA
Hyla F          TC.CA CTGT. T..CT TG.G. ..TGT CT---
```

Figure 4.5. Autoradiograph of SOD sequences. Sequencing of the two size class fragments amplified using SOD primers showed evidence of significant sequence heterogeneity within size classes. The larger of the two size fragments (Band C=800 bp) was relatively unambiguous but did not appear to contain an intron region. The smaller of the two size fragments (Band B=350 bp) was heterogenous in some samples but sequences from the less ambiguous samples could not be aligned to published SOD sequences.

SOD

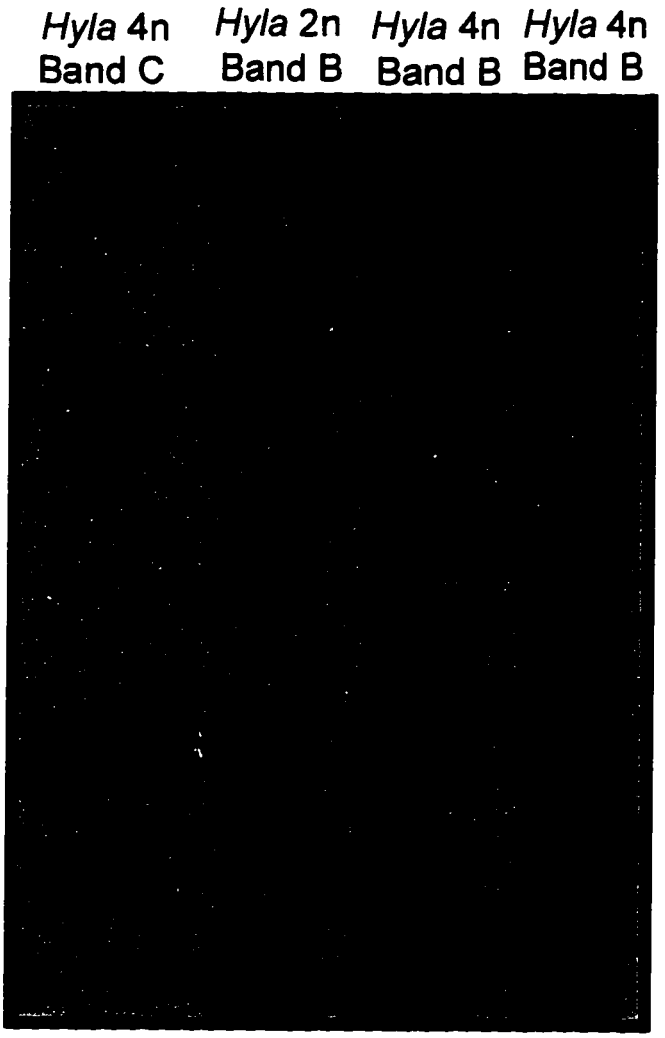
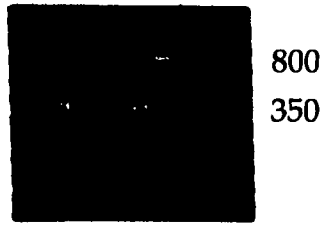


Figure 4.6. Agarose and denaturing gradient gels to show heterogeneity of SOD sequences. (a) Demonstration of the bands amplified using SOD primers on *Hyla* samples. Two predominant size classes were isolated (350 bp and 800 bp). (b) The sample indicated by the arrow (lane 5) was applied to a 7% denaturing gradient gel (DGGE) to determine if heterogeneity within the sample could be determined by this method. Note that the sample in lane 8 shows three distinctive bands, corresponding to the two size fragments and another band indicating that at least one of the size fragments was heterogeneous.

SOD-PCR Products

1.2%Agarose



DGGE

7% Denaturing Gradient Gel

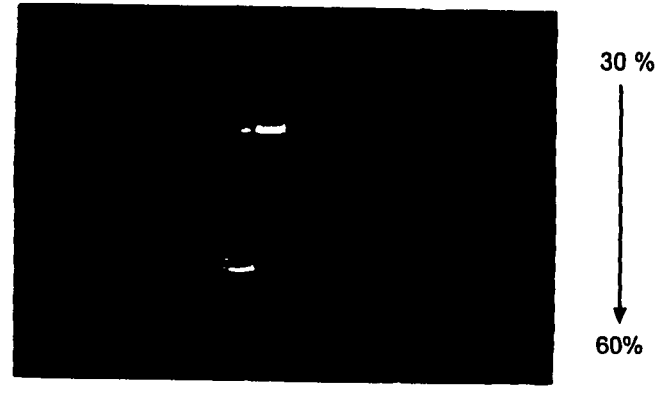


Figure 4.7. Alignment of a partial fragment of Aldolase C from a diploid and tetraploid *Hyla*, a diploid and tetraploid *Neobatrachus* and the published Skink sequence. Frog sequences are compared to the tetraploid *Hyla* sequence for reference. Matching characters are indicated by a dot. Nucleotides that differ between the Skink sequence and the tetraploid *Hyla* sequence are represented in lowercase. Notice that the diploid *Hyla* sequence appears to be distinctive from those of the other taxa, suggesting that it may represent a contaminant or an additional copy of the aldolase gene.

ALDOLASE 1

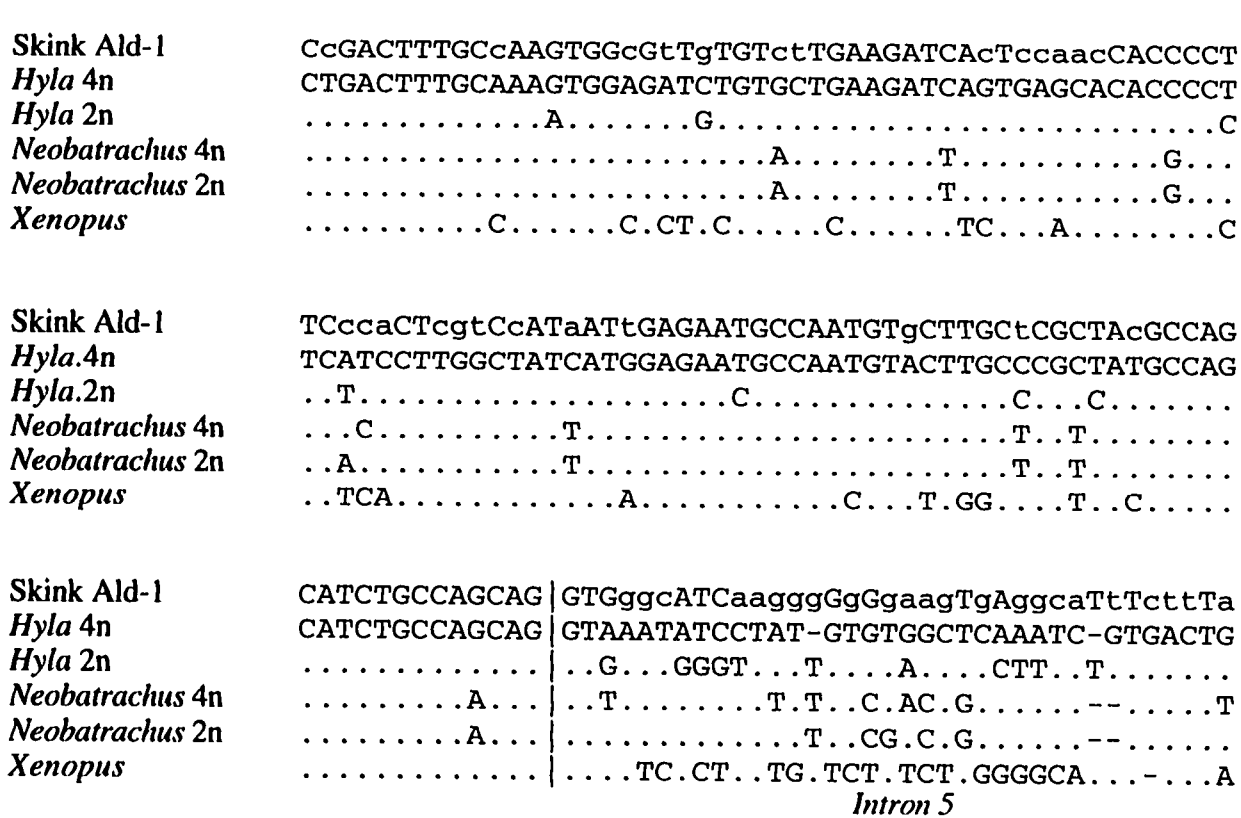
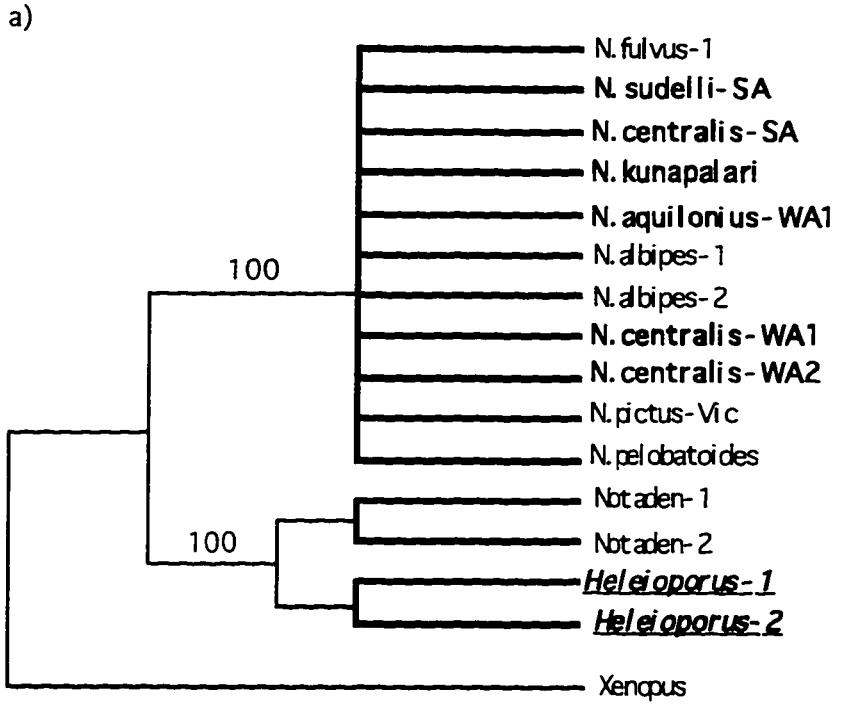

50
Skink Ald-1 CcGACTTTGCcAAGTGGcGtTgTGTctTGAAGATCAcTccaacCACCCCT
Hyla 4n CTGACTTTGCAAAGTGGAGATCTGTGCTGAAGATCAGTGAGCACACCCCT
Hyla 2n .....A.....G.....C
Neobatrachus 4n .....A.....T.....G...
Neobatrachus 2n .....A.....T.....G...
Xenopus .....C.....C.CT.C.....C.....TC...A.....C
100
Skink Ald-1 TCccaCTcgtCcATaATtGAGAATGCCAATGTgCTTGCtCGCTAcGCCAG
Hyla.4n TCATCCTTGGCTATCATGGAGAATGCCAATGTACTTGCCCGCTATGCCAG
Hyla.2n ..T.....C.....C...C.....
Neobatrachus 4n ...C.....T.....T..T.....
Neobatrachus 2n ..A.....T.....T..T.....
Xenopus ..TCA.....A.....C...T.GG....T..C.....
150
Skink Ald-1 CATCTGCCAGCAG|GTGggcATCaagggGgGgaagTgAggcaTtTcttTa
Hyla 4n CATCTGCCAGCAG|GTAAATATCCTAT-GTGTGGCTCAAATC-GTGACTG
Hyla 2n .....|..G...GGGT...T....A....CTT..T.....
Neobatrachus 4n .....A...|..T.....T.T..C.AC.G.....--.....T
Neobatrachus 2n .....A...|.....T..CG.C.G.....--.....
Xenopus .....|....TC.CT..TG.TCT.TCT.GGGGCA...-...A
Intron 5

Figure 4.8. Phylogenetic reconstructions using aldolase sequences in *Neobatrachus* and closely related genera. Note that *Heleioporus* (*H. psammophilus*) and *Notaden* (*N. melanoscaphus*) have distinctive sequences from all of the *Neobatrachus* taxa but that few changes occur within the ingroup. (a) Strict consensus of the 2812 most parsimonious trees found in an unweighted search using all taxa sequenced. (b) Representative phylogram of one of the 7 equally parsimonious trees found in a search with a reduced data set. Bootstrap proportions (100 replications) for nodes supported are indicated.

a)


N. fulvus-1
N. sudelli-SA
N. centralis-SA
N. kunapalari
N. aquilonius-WA1
N. albipes-1
N. albipes-2
N. centralis-WA1
N. centralis-WA2
N. pictus-Vic
N. pelobatoides
Notaden-1
Notaden-2
Heleioporus-1
Heleioporus-2
Xenopus
100
100

b)

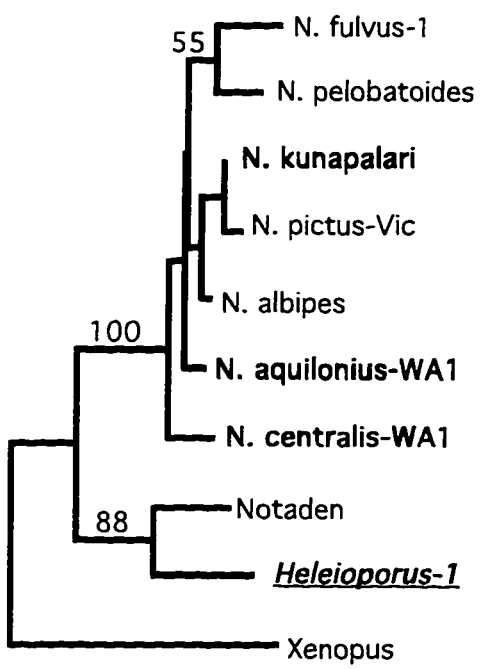
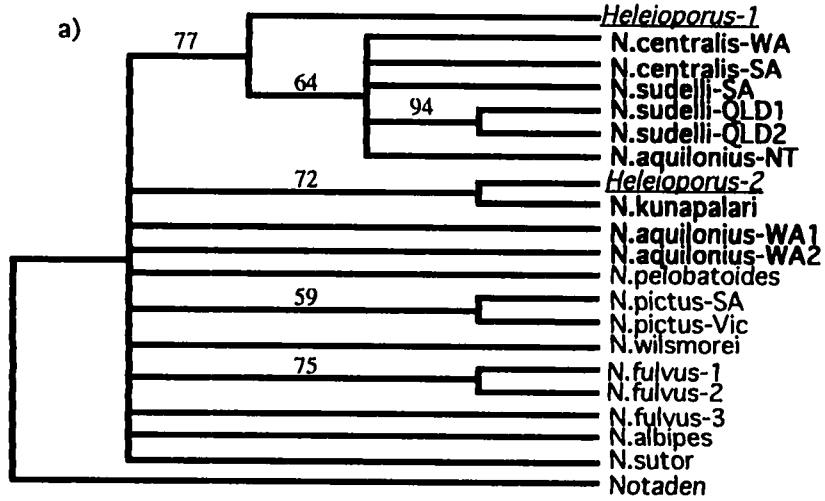
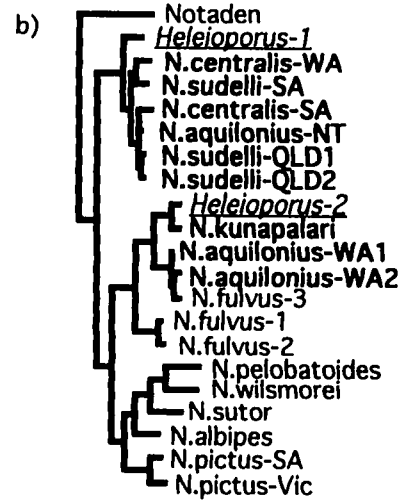
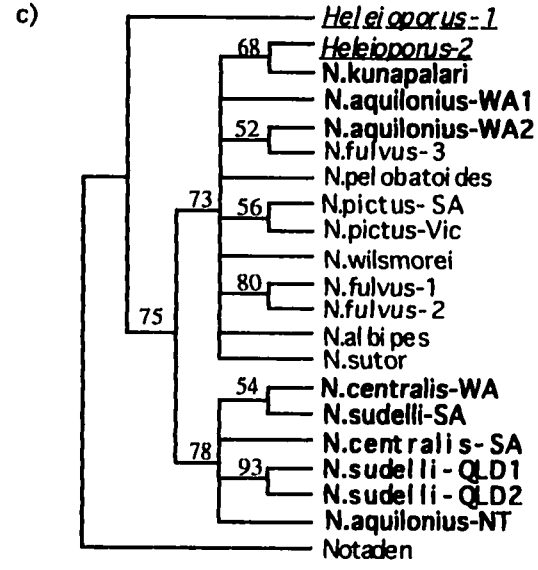
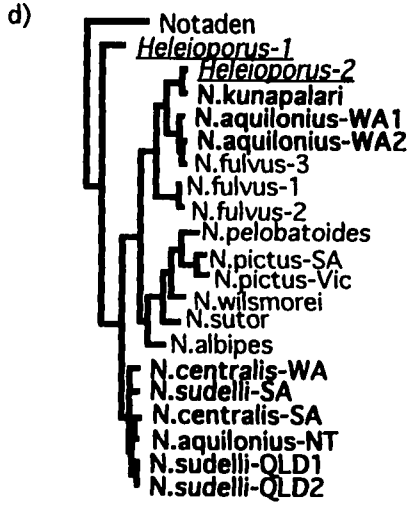

N. fulvus-1
N. pelobatoides
N. kunapalari
N. pictus-Vic
N. albipes
N. aquilonius-WA1
N. centralis-WA1
Notaden
Heleioporus-1
Xenopus
55
100
88

Figure 4.9. Comparison of parsimony trees found using sequences from mitochondrial DNA and aldolase. (a) Strict consensus tree of 100 bootstrap replications and (b) a phylogram of the single most parsimonious tree found in a heuristic search with characters unweighted using sequences obtained from the mitochondrial genes cytochrome oxidase I and cytochome b using a representative sample from each of the ten nominal *Neobatrachus* species and the two most closely related genera, *Notaden* and *Heleioporus*. Note that inclusion of the *Heleioporus* samples resulted in nonmonophyly within the genus *Neobatrachus* in the mitochondrial genes. (c) Strict consensus of 100 bootstrap replications and (d) representative phylogram from one of four equally most parsimonious trees found when aldolase and the two mitochondrial data sets were combined.


a)
77
64
94
72
59
75
Heleioporus-1
N.centralis-WA
N.centralis-SA
N.sudelli-SA
N.sudelli-QLD1
N.sudelli-QLD2
N.aquilonius-NT
Heleioporus-2
N.kunapalari
N.aquilonius-WA1
N.aquilonius-WA2
N.pelobatoides
N.pictus-SA
N.pictus-Vic
N.wilsmorei
N.fulvus-1
N.fulvus-2
N.fulvus-3
N.albipes
N.sutor
Notaden


b)
Notaden
Heleioporus-1
N.centralis-WA
N.sudelli-SA
N.centralis-SA
N.aquilonius-NT
N.sudelli-QLD1
N.sudelli-QLD2
Heleioporus-2
N.kunapalari
N.aquilonius-WA1
N.aquilonius-WA2
N.fulvus-3
N.fulvus-1
N.fulvus-2
N.pelobatoides
N.wilsmorei
N.sutor
N.albipes
N.pictus-SA
N.pictus-Vic


c)
68
52
73
56
80
75
54
78
93
Heleioporus-1
Heleioporus-2
N.kunapalari
N.aquilonius-WA1
N.aquilonius-WA2
N.fulvus-3
N.pelobatoides
N.pictus-SA
N.pictus-Vic
N.wilsmorei
N.fulvus-1
N.fulvus-2
N.albipes
N.sutor
N.centralis-WA
N.sudelli-SA
N.centralis-SA
N.sudelli-QLD1
N.sudelli-QLD2
N.aquilonius-NT
Notaden


d)
Notaden
Heleioporus-1
Heleioporus-2
N.kunapalari
N.aquilonius-WA1
N.aquilonius-WA2
N.fulvus-3
N.fulvus-1
N.fulvus-2
N.pelobatoides
N.pictus-SA
N.pictus-Vic
N.wilsmorei
N.sutor
N.albipes
N.centralis-WA
N.sudelli-SA
N.centralis-SA
N.aquilonius-NT
N.sudelli-QLD1
N.sudelli-QLD2

Chapter 5 : Discussion

Table of Contents

APPROPRIATENESS OF MOLECULAR MARKERS

Allozymes

While allozymes remain a very powerful tool for certain types of questions (see Hillis *et al.* 1996b), they may not be the most effective method for the questions posed in this dissertation. Allozyme data could potentially provide evidence for a hybrid origin of tetraploids or of current or historical gene flow between diploids and tetraploids. However, interpretation of banding patterns is complicated (especially when dealing with polyploids) by the fact that this type of analysis compares gene products and thus relies on the assumption that all variants present are equally expressed. Considering that one of the interesting questions about polyploidy is how patterns of gene expression might change over time, this could be a very serious obstacle. If silencing of duplicate genes could be detected reliably, information could still be gained on the fate of duplicated gene sequences following polyploidization, if not on phylogenetic relationships. However, since electrophoretic detectability requires a net charge difference, not all duplicate genes will be detectable and the proportion of loci at which duplicate genes have been silenced can only be estimated based on rates of divergence in other organisms (e.g., Hughes and Hughes 1993).

Kimura and Ota (1974) predicted that "gene duplication must always precede the evolution of a gene having a new function". However, Nei (1969) suggested that the most likely outcome for gene duplication is for one copy to become nonfunctional as a result of the accumulation of deleterious mutations, which eventually results in silencing of the redundant gene. Most of the information on duplicated genes in polyploids comes from studies on tetraploid fish (e.g., Ferris and Whitt 1977a,b; Li 1980; Larhammar and Risinger 1994) and polyploid frogs in the *Xenopus laevis*

species group (e.g., Kobel and Pasquier 1986; Vriz *et al.* 1989; Principaud and Spohr 1991; Schuldiner *et al.* 1991; Hughes and Hughes 1993). Given the predicted age of tetraploids in these groups (35-50 MYA), along with evidence that substitution rates are often high after gene duplication (Czelusniak *et al.* 1982), and that a single gene is sufficient for normal function in every enzyme system studied in catostomids (Li 1980), one might predict that duplicated genes would experience rapid divergence to either one of these fates. These studies have found a surprisingly low rate of gene silencing (Li 1980; Hughes and Hughes 1993; Larhammar and Risinger 1994). They conclude that a long period of tetrasomy following tetraploidization slows the rate of reduction of duplicate gene expression, but that slow accumulation of mutations in the gene pair may result in radical amino acid changes which might eventually lead to new gene functions or different regulatory pathways so that the now divergent gene copies are expressed at different times or in different places (Hughes and Hughes 1993; Larhammar and Risinger 1994).

Inheritance studies using the North American tetraploid grey treefrogs suggest that gene expression in hybrids may be altered and may exhibit tissue and developmental-stage specific expression (e.g., Mable 1989; Mable and Bogart 1995), but it is not known whether this reflects actual regulatory differences between the parental types related to polyploidy or if differences in expression are revealed as a result of hybridization itself (e.g., "hybrizymes": Woodruff 1989; Bradley et al. 1993). Ohno (1985) suggested that "dispensable genes" such as nonfunctional duplicate gene copies may last a long time in vertebrates but may exist only transiently in bacteria or viruses because there is an inverse relationship between the spontaneous mutation rate and the total number of gene loci in a species, which may be dependent

on the relative accuracy of polymerases and repair mechanisms. A study of a chaperonin gene duplicated in the common ancestor of *Mycobacterium* and *Streptomyces* species suggested that duplicated genes have evolved at different rates in the former but have evolved at the same rate in the latter (Hughes 1993). This emphasizes that broad generalizations in rate comparisons may not be valid even within similar taxonomic groups let alone between very different taxonomic groups because selective constraints may play an important role in the rate of change of duplicate genes. Therefore, using rates of evolution of duplicate genes as a means of predicting times of divergence of the speciation event that led to the emergence of polyploid species seems like a futile exercise, especially using allozyme data alone.

Mitochondrial DNA

Allozymes also may not have diverged enough to detect the same degree of differentiation between diploids and tetraploids as seen in mtDNA. In tetraploid plants, autotetraploids tend to share the same alleles as their diploid counterparts (Soltis and Soltis 1989) and multiple origins of tetraploids predicted from cpDNA sequence differences were not detectable in an allozyme study on the same populations (Soltis *et al.* 1989). In this dissertation, diploids and tetraploids within both groups of frogs studied shared most of the same alleles and in the North American group, tetraploids tended to cluster with their sympatric diploids. However, sequences from the mitochondrial gene COI suggested that tetraploids did not share similar haplotypes with sympatric diploids and tended to be more similar to each other than to any of the diploids examined. Tetraploids in the Australian group also tended to have more similar mitochondrial haplotypes to the other tetraploids in their clade

(eastern or western) than to any of the diploids, but these relationships were not as clear. Rates of evolution in mitochondrial genes are thought to be 5-10 times greater than in single-copy nuclear genes for mammals, possibly due to a higher mutation rate (Brown *et al.* 1979) or less efficient repair or editing mechanisms (e.g., Martin *et al.* 1992). This discrepancy between conclusions drawn from allozyme and mitochondrial data thus could be due to a difference in resolution since allozymes only detect amino acid differences that result in a charge change, and so divergence in mitochondrial genes might be seen before the accumulation of protein changes. For the grey treefrog complex, pairwise comparisons of COI sequences showed mainly third base pair synonymous nucleotide and showed few amino acid substitutions (82:1 ratio of synonymous:nonsynonymous nucleotide substitutions), so changes in proteins would not be expected even for the potentially more rapidly evolving mitochondrial genes.

This pattern of change in the mitochondrial genes (i.e., rapid evolution of silent differences and slow accumulation of replacement differences) could partially explain the limited resolution found between diploid and tetraploid *Neobatrachus*. Cytochrome oxidase I had relatively high sequence divergence (4 to 20%) within ingroups and transition:transversion ratios varied widely across pairs of taxa (18:1 to 0.5:1). It is well known that transitions often outnumber transversions in comparisons of sequences from closely related taxa and that this difference is reduced with degree of divergence until transition:transversion ratios approach unity or less (e.g., Brown *et al.* 1982; Graybeal 1993, 1994). The COI data show this entire range: tetraploids within the "eastern" clade are separated by transitional changes only but they are separated from all other taxa by a mixture of transitional and

transversional changes. Averaging across such a broad range within a data set may not be particularly meaningful and weighted corrections may not improve the ability to extract meaningful phylogenetic signal. The saturation of transitional changes in deeper phylogenetic splits is probably due to the obliteration of the record of multiple substitions at a given site and corrective measures can also be applied to account for this (e.g., log det transformation: Lockhart *et al.* 1994). However, these measures all assume both neutrality and relatively equal rates of change among taxa (see Swofford *et al.* 1996). Bufonid toads show an "unusually low" rate of amino acid replacement in cytochrome b (cytb) compared to other vertebrates, but silent differences are saturated and the author suggests that this could be due to physiological constraints on cytb in toads (Graybeal 1993). The rationale is that toads are thought to have high endurance capacities compared to other amphibians (Bennett and Licht 1974) and cytb might perform an important energetic function that would limit the types of changes that are allowable in the gene (Graybeal 1993). In diploid and tetraploid *Neobatrachus*, the much lower rate of nonsynonymous changes compared to the high rate of synonymous changes could be evidence of similar selective constraints acting. In addition, there appeared to be an excess of C:T transitions compared to A:G transversions that appeared to be relatively uniform across taxa. If codon usage is restricted, it seems that the potential for bouncing back and forth between a very restricted number of allowable changes is high and multiple hits could result in extreme homoplasy. Comparisons of synonymous changes per synonymous site and nonsynonymous changes per nonsynonymous site have been used to find evidence of positive Darwinian selection acting on immunologically important genes (e.g., Nei and Gobojori 1986; Tanaka and Nei 1989; Ota and Nei 1995; Hughes 1993). A

comparison of synonymous and nonsynonymous changes among various combinations of *Neobatrachus* for both cytb (using a data set provided by Steve Donnellan at the South Australian Museum) and COI suggested that this ratio varied in similar manner to transition:transversion ratios across the data sets for COI but changes appeared to be saturated in cytb (Table 5.1). This suggests that the two genes might be expected to show very different patterns of divergence but the high ratio of synonymous to nonsynonymous changes in each could be an indication that both are under strong selective constraints. If different taxa within the group of study experienced different selection pressure within these genes, it could make the genes inappropriate for phylogenetic analysis.

Even if the possibility of selection is not considered, these mitochondrial genes may not have been the most appropriate for comparison, especially for *Neobatrachus*. Characters are phylogenetically informative when their rate of evolution matches the divergence times of interest: if they evolve too slowly they will be uninformative and if they evolve too quickly homoplasy will obscure phylogenetic signal (Graybeal 1994). However, this matchmaking scheme is a very difficult one for taxonomic groups such as frogs for which the fossil record is not very informative. Average percent sequence divergence at different taxonomic levels has tended to be the major criterion used in choice of genes for phylogenetic analysis of a particular taxonomic group. However, this measure averages over several rates that could be important in phylogenetic comparisons and ignoring them could seriously reduce the ability to extract meaningful relationships among the taxa (Graybeal 1994). These include the rate of substitution per site, the proportion of sites which are allowed to vary, and the kinds of substitutions accepted at the varying sites (Fitch

1986; Graybeal 1994). Rate heterogeneity among sites (e.g., Yang *et al.*, 1994; Wakely, 1994), and different rates of change for synonymous vs nonsynonymous nucleotide substitutions due to selective constraints (e.g., Li *et al.* 1985; Nei and Gojobori 1986; Li 1990; Muse and Gaut 1994; Yang *et al.* 1994) as discussed above may also reduce the probability that percent sequence divergence alone will reflect the usefulness of the genes chosen. Establishment of these parameters is also taxon specific and it is not obvious how *a priori* predictions could be made to incorporate these types of complicating factors in choice of a gene. The saturation effect discussed above results in a limit to phylogenetic resolution: sequences that have diverged for only 20 million years tend to differ by almost as much as those which have diverged for 80 million years (Brown *et al.* 1982).

Nuclear Genes

The analysis of nuclear genes did not provide much additional information to the resolution of phylogenetic relationships in either group. However, it did point out some serious considerations for the use of nuclear genes in phylogenetic comparison of genes that evolve rapidly enough for comparison of closely related taxa. The key problem can be summarized in one word: homology. In phylogenetic analysis, it is essential that orthologous sequences are compared (e.g., Hillis 1994; Moritz and Hillis 1996; Dowling *et al.* 1996). The presence of pseudogenes and duplicated gene copies that are amplified with the target sequence introduces a serious risk of comparing paralogous or xenologous gene copies. When using gene families that are known to undergo concerted evolution, it is also important to establish that a

consensus sequence does not obscure within-individual variability that may be present.

The rate of evolution in the nuclear gene introns examined was less than predicted by neutral theory models. It may be that putatively functionless gene regions or gene copies may have purposes that have not been identified. Therefore, rates of evolution cannot be predicted *a priori* but need to be properly tested by pilot screening studies. For application to questions of polyploid evolution, the use of nuclear genes may complicate analyses beyond interpretation. In polyploid plants, areas of repeated nucleotide divergence have been compared and evidence suggests that concerted evolution may preserve both parental sequences in hybrids, may randomly "choose" one or the other or may be unique to the hybrid taxa (e.g., Zhang and Dvorak 1991; Wendel *et al.* 1995). Studies of patterns of nucleotide changes in duplicate genes in both tetraploid animals (Hughes and Hughes 1993) and within diploid taxa (e.g., Hughes 1993) also point out another potential concern: that duplicated genes within an animal might be subject to different selective pressures and therefore might evolve differently. These factors should be considered when using nuclear genes for study of diploid or polyploid organisms. Study of nuclear genes for use in phylogenetic comparisons could, however, have the added benefit of contributing significantly to an increase in the overall understanding of evolution at the molecular level.

Evolutionary Origins of Tetraploids

Discrepancies between data sets used and reservations about the appropriateness of the particular markers chosen for comparison limit the potential for

confident conclusions on evolution of tetraploids in either group examined. However, there were some common conclusions that may bear on the evolution of polyploidy in general. The conclusion that was most strongly supported in both analyses is that there have been multiple origins of tetraploidy in both *Hyla* and *Neobatrachus*. Independent mitochondrial lineages of tetraploids is the only result that has been supported strongly by both allozymes and mtDNA sequences. This suggests that polyploidy may have been a more common phenomenon than previously considered, with tetraploids possibly being successful in times of environmental upheaval. Tetraploids in the plant genus *Heuchera* appeared to have arisen from morphologically indistinguishable diploids through multiple autopolyploidization events: a process previously thought to be much less common in plants than allopolyploidization (Soltis *et al.* 1989).

In artificial laboratory crosses using amphibians, polyploid eggs are produced in about 5% of normal crosses (Richards and Nace 1977). With cold or pressure shock treatment of eggs just prior to fertilization, this number can be increased to 80 or 90% (Briggs 1947; Nishioka and Ueda 1983; Kawamura 1984). Therefore, during times of environmental change (or fluctuating temperature conditions in general), the production of polyploid offspring could be quite high. Under stable conditions, it may be that polyploid individuals may not be able to establish themselves in the midst of an already well adapted diploid population. This could change under times of environmental change, if, as has been frequently suggested, polyploids are more genetically plastic due to the increased number of gene copies available to draw on and increased heterozygosity that may be able to balance or hide deleterious mutations. In plants, it has been suggested that enzyme multiplicity and

high heterozygosity could give both allopolyploids (e.g., Barber 1970) and autopolyploids (Soltis and Soltis 1989) greater biochemical and genetic flexibility relative to diploid progenitors. The rationale is that multiple enzyme forms may minimize the effect of changed reaction conditions, and thus expand the range over which physiological coordination of metabolic processes is possible (Johnson 1974, 1976). This could permit polyploids to maintain sufficient flux through metabolic pathways in the event of altered reaction conditions when the environment changes (Levin 1983). However, under more stable conditions this effect might be balanced by genetic load effects.

It is not yet known how many polyploid species of frogs actually exist. Bisexual diploid-polyploid cryptic species are known to exist in the South American genera *Odontophrynus* (family Leptodactylidae), *Ceratophrys* (family Leptodactylidae), and *Phyllomedusa* (family Hylidae); North American frogs in the genus *Hyla* (family Hylidae); Australian frogs in the genus *Neobatrachus* (family Myobatrachidae, formerly Leptodactylidae); African frogs in the genera *Xenopus* (family Pipidae), *Dicroglossus* (family Ranidae), *Tomopterna* (formerly *Pyxicephalus*) (family Ranidae) and *Bufo* (family Bufonidae)(reviewed by Bogart 1980; Kawamura 1984); and more recently discovered, European frogs in the genus *Bufo* (family Bufonidae)(Borkin *et al.* 1995); and Papua New Guinean frogs in the genus *Cophixalus* (family Microhylidae: Genyophryninae) (Kuramoto and Allison 1989). There is some evidence that polyploidy may also exist in North American representatives of the families Microhylidae and Pelobatidae (Bogart, personal communication). These groups encompass quite a divergent array of frog families and ecological types that do not obviously provide a clue as to why tetraploidy exists

in these groups but not in others. However, this diversity emphasizes that polyploidy may have been a significant mode of speciation throughout anuran evolution.

One consequence of polyploidy that has been proposed as a reason for the success of polyploid species is that they may be more tolerant to environmental stress due to hybrid vigour effects in allopolyploids or the increased capacity for genetic flexibility provided by the "extra" genome in both autopolyploids or allopolyploids (e.g., Bogart 1980). However, no studies have conclusively demonstrated differences in ecology, physiology or reproductive output that might indicate this in anurans. In North America, tetraploids extend across the Northern border of the Great Lakes while diploids are restricted to the southern limits of the Great Lakes and extend much further south and east than the tetraploids. One reason that has been invoked for the current distribution has been that tetraploids may have higher freeze tolerance than diploids and may have been able to extend further north during periods of glacial recession than diploids. While anecdotal evidence (i.e., tetraploids from northern populations can survive complete freezing at -20°C without damage while diploids from southern populations cannot) supports this, it has not been confirmed experimentally. Tetraploid grey treefrogs are known to use glucose in freeze tolerance and it may be that tetraploids produce a greater absolute quantity, or have more receptor sites that prevent ice crystal formation upon thawing. Diploids have not been as carefully examined. It would be interesting to determine if differences in freeze tolerance can be verified by biochemical evidence.

Another factor that has been suggested is that polyploids might be more tolerant to dessication than their diploid counterparts (Bogart 1980). This would be a tempting speculation given that at least four groups of polyploids (*Odontophrynus*,

Ceratophrys, *Neobatrachus*, microhylids) are burrowing species. However, comparison of diploid and tetraploid *Neobatrachus* has not revealed any identifiable differences in water retention or dessication tolerance (Withers 1995). Further study comparing the different groups would be interesting, to see if tetraploids do survive long periods of time underground better than diploids.

This study did not identify an extant diploid that was likely to have given rise to either the northwestern tetraploid grey treefrogs in North America or to most of the tetraploids in the Australian *Neobatrachus* complex. While it is possible that this is due to insufficient resolution of phylogenetic relationships, it is possible that this is an evolutionarily meaningful result. In fact, most studies of polyploid frogs have been unable to identify a definite diploid ancestor. Because polyploids are readily produced artificially, it is surprising that more polyploid species do not occur naturally. It may be that the establishment of polyploid species is only reproductively feasible under conditions for which polyploids are able to outcompete the diploids with which they are in closest association. In times or regions where polyploidy is favoured, diploids may not be as competitive and tetraploids may actually replace their diploid progenitors. It would be interesting to compete diploids and tetraploids under stable and fluctuating conditions and compare relative survivorship ability. One prediction of this hypothesis would be that tetraploids would be favoured under variable conditions, while diploids would be favoured under stable conditions. If this were true, the present sympatry of tetraploids and diploids could result from secondary contact rather than continual sympatry since conception.

The data presented here also did not suggest whether it was more likely that tetraploids in either group arose through hybridization or through autopolyploidy. A

recent study on South African burrowing frogs in the genus *Tomopterna* (formerly *Pyxicephalus*, family Ranidae) suggests that a tetraploid species (*Tomopterna tandyi*) arose through hybridization between two diploid species in the same genus (*Tomopterna cryptosis* and *Tomopterna delalandii*), neither of which are currently sympatric with it (Channing and Bogart, unpublished). This was suggested based on allozyme alleles that were shared uniquely between the tetraploid and each of the diploid species but that were not shared by the two diploids. It was predicted from mitochondrial data (cytb) that tetraploids arose through hybridization of females of *T. cryptosis* with males of *T. delalandii*. This is the first evidence for allopolyploidy in bisexually reproducing tetraploid frogs, except for the various species of *Xenopus*, which are all thought to have arisen through ancient hybridization events (Tymowska 1991). In polyploid plants, allopolyploidy has been assumed to be the major mode of speciation but there is some evidence to suggest that autopolyploidy may have been a more significant evolutionary factor than previously presumed (Soltis *et al.* 1989).

Another interesting observation suggested by my data has to do with the relative age of tetraploids in each complex. While little is known about evolution in the genus *Neobatrachus*, divergence of this genus from other burrowing frogs (*Heleioporus*, *Notaden*) has been described. Based on this evidence, it has been assumed that tetraploidy within *Neobatrachus* was a relatively recent event. While there is little evidence to predict actual times of divergence, my mtDNA data suggest that polyploidy in *Neobatrachus* probably occurred much earlier than in the North American complex. In the North American group, the highest divergence levels in COI were those between tetraploids from the northwestern region and diploids from central (=western diploid) Texas (4.2%). In the Australian group, the most closely

related populations showed 4% sequence divergence and occurred within the "eastern" tetraploid clade, within nominal species groups (*N. sudelli, N. centralis*). Differences between the eastern and western tetraploid clades averaged about 11%, while the smallest difference between a tetraploid and any of the diploids was 13%, and the highest difference between any pair of taxa was 15% (*N. pelobatoides* and *N. fulvus*). If it were assumed that both sets of taxa showed similar rates of divergence in this gene, the most closely related taxa in the *Neobatrachus* complex are as divergent as the most distantly related taxa in the North American complex. The split between eastern and western tetraploids in the Australian group in an analysis of cytb (unpublished data) is as deep as that between *Hyla versicolor* and the outgroup (*Hyla arenicolor*) used in the study of Ptacek *et al.* (1994). It is not surprising therefore that tetraploids within the Australian group have morphological and behavioural differences that are recognizable while the North American group have remained completely cryptic.

While it would be tempting to predict when tetraploids in each group arose, there is really no basis on which to make this prediction. Rates of change of nucleotide substitution in mitochondrial genes have been established for mammals for up to about 20 million years of divergence (e.g., Brown *et al.* 1979). However, this rate cannot be applied across all groups. For example, sharks have been shown to exhibit a much lower (7-8 fold) rate of nucleotide substitution than mammals in both cytb and cytochrome oxidase, which was attributed possibly to a lower rate of metabolism in ectotherms (Kocher *et al.* 1989; Martin *et al.* 1992). The rationale is that oxygen radicals are potent intracellular mutagens and a lower metabolic rate could thus result in slower mtDNA evolution (Martin *et al.* 1992). Bargelloni *et al.* (1994)

also found a slow rate of mtDNA evolution in antarctic fish that they attibuted to decreased metabolism at cold temperatures. Burrowing frogs experience a significant decrease in metabolic rate during times aestivation and cocoon formation (Withers 1993), and they are ectothermic. They therefore might have relatively slow rates of mtDNA evolution and may be even older than the sequences suggest. Tetraploids are found across the central desert regions while diploids are restricted to the more arid and coastal regions. It is not known how this distribution might correlate with metabolic rate effects. In the same vein, tetraploid grey treefrogs exploit colder environments than diploids and could have reduced metabolic rates. This, however, has not been found to be true, at least as measured by relative oxygen consumption in diploids and tetraploids (Kamel *et al.* 1985).

Origins of Tetraploids: Isolating Mechanisms

One of the consequences that has been proposed for polyploidy in anurans is that the change in cell size may lead to automatic isolation from the diploid progenitor species through changes in call parameters directly tied to the polyploidization event (Bogart and Wasserman 1972; Bogart 1980). While one of the original goals of this thesis was to address this hypothesis in both groups examined, the limited phylogenetic resolution achieved does not allow definite conclusions to be drawn. However, some trends can be noted.

Variation among calls of diploids and tetraploids in the grey treefrog complex has been described from various parts of the ranges (Zweifel 1970; Ralin 1977; Gerhardt 1974; Jaslow and Vogt 1977; Bogart and Jaslow 1979; Gerhardt 1978) but an extensive survey of variation across the northwestern part of the range has not

been conducted. Other studies have found significant differences in call duration that varied according to the geographic boundaries predicted by mtDNA studies (Ptacek *et al.* 1994). However, since pulse rate is the feature which females have been demonstrated to use in mate discrimination (Littlejohn 1960; Gerhardt 1978; 1982; 1991; Doherty and Gerhardt 1984; Klump and Gerhardt 1987; Gerhardt and Doherty 1988), it is not known how significant differences in duration would be in male-female interactions. Multiple origins of tetraploids, all having reduced pulse rate calls compared to all diploids examined, suggest that this could have been the result of an automatic change in the calls with the increase in cell size. However, as yet a physiological explanation does not exist for how pulse rate might be altered by cell size. Since the tetraploid has fewer, larger cells than the diploid, it may be that pulse rate is determined by the change in cell size of the vocal apparatus resulting from the doubling of the genome. However, this could only lead to an automatic prezygotic isolation mechanism in the newly developed species if a concomitant change occurred in the female auditory system (Bogart 1980). Female grey treefrogs have shown a clear preference for calls of their own ploidy level at the same temperature (Gerhardt 1982; Klump and Gerhardt 1987; Gerhardt and Doherty 1988). Furthermore, it has been found that the temporal tuning in cells in the midbrain (torus semicircularis) in diploid and tetraploid grey treefrogs are closely related to the respective pulse rates in the two species (Rose *et al.* 1985) and show a trend in sensitivity corresponding to changes in pulse rate that occur with temperature in each species (Brenowitz *et al.* 1985). However, McLister *et al.* (1995) were unable to identify changes in morphology, firing rate, or recovery period of laryngeal muscles of diploid and

tetraploid grey treefrogs. Therefore, it remains unclear exactly why the differences in pulse rate exist.

In the Australian group, a definite diploid ancestor was not found for most of the tetraploids. However, *N. fulvus* appeared to be the most closely related diploid to the tetraploid *N. aquilonius* (and possibly to *N. kunapalari)* and has a pulse rate that is within the range of all of the tetraploids. This would not support the hypothesis that a direct decrease in pulse rate occurs with polyploidization. In fact, if it were the ancestor of *N. kunapalari*, there would actually be an increase in pulse rate (J.D. Roberts, personal communication). The slope of the regression of pulse rate on temperature for *N. aquilonius* is similar to that of *N. fulvus* and pulse rates are generally similar over the temperature range examined (see Figure 3.6). *Neobatrachus aquilonius* is morphologically almost indistinguishable from *N. fulvus*, corresponding with morphological similarity in diploids and tetraploids in some of the other complexes known (*H. versicolor*--Ralin and Rogers 1979; *Ondontophrynus americanus*--Beçak and Beçak 1971). The populations of *N. fulvus* and *N. aquilonius* used in this study were both from an extreme northwestern region of Western Australia, near the Exmouth peninsula. Two of the other diploids, *N. sutor* and *N. wilsmorei* were also from this general region. These taxa have single pulsed calls and therefore it is unlikely that either gave rise to the tetraploid, *N. aquilonius*, unless as a result of a hybridization event with one of the diploids with a faster pulse rate but none of the other diploids currently have a range in this region.

N. kunapalari has a different karyotype from both *N. fulvus* and *N. aquilonius*, is morphologically distinctive and has a mating call which has a higher pulse rate than *N. fulvus*. The populations of *N. kunapalari* used in this study were

from central Western Australia. It is therefore difficult to explain its origin if *N. fulvus* is the progenitor species. In many of the analyses, *N. albipes* appeared to be most closely related to one of the *N. kunapalari* populations, but discrepancies between the placement of the two *N. kunapalari* populations does not permit much confidence in this result. However, a separate origin for *N. kunapalari* would fit with its unique characteristics and the *N. kunapalari* and *N. albipes* populations used in this study were from the same region of mid-Western Australia so this relationship would make biogeographic sense. If *N. albipes* was the ancestor of *N. kunapalari*, this would support the hypothesis of decreased pulse rate with polyploidy, as it has a pulse rate that is nearly double that of the tetraploids. However, the distinctive karyotype of *N. kunapalari* makes it difficult to support or refute any of the extant diploids as its ancestor and it may be that it arose from an extinct or unidentified diploid lineage. The "eastern" clade of tetraploids did not appear to be closely related to any of the diploid taxa and may have diverged from a common ancestor with them.

Therefore it would appear that call characteristics of tetraploids in the North American group (and in South American *Odontophrynus*) support a hypothesis of automatic call change with polyploidy while the Australian group provides ambiguous conclusions. It is possible that this difference is tied to the relative importance of species advertisement calls to the reproductive biology of these species. In all North American species of *Hyla* examined, females have the capacity to selectively choose mates based on specific mating call parameters (e.g., Gerhardt and Doherty 1988). These species also tend to call over a period of a few nights to several weeks or months and therefore there is a relatively long period of advertisement over which mistakes could potentially occur. Many species also display male-male interactions

that result in separation of calling males from one another, even in a crowded breeding area. It is not known whether female *Neobatrachus* are able to discriminate calls of conspecifics or whether they show selective phonotaxis based on specific call parameters. However, they also have a different breeding strategy. These species are all explosive breeders which rely on heavy rainfall to trigger emergence from their burrows for reproduction. Some species breed on a single night a year or may not breed every year. Breeding is more of an opportunistic orgy than an organized event and it is possible that while mating calls are used to attract females to the breeding site and generally cue which species they should orient to, specific call parameters may not be as important and mate choice may be more random than carefully planned (e.g., Roberts 1978). This would be an interesting hypothesis to test. A decrease in pulse rate with increased ploidy level also has been observed in tetraploids of the South American *Odontophrynus americanus* complex (Beçak and Beçak 1970; Bogart 1980) but diploids and tetraploid *Phyllomedusa* produce calls that are not distinguishable from each other (Haddad *et al.* 1994). A more careful analysis of call variation in the calls of other diploid-tetraploid species pairs is required before definite conclusions are made on the factors controlling mating call changes with polyploid speciation.

Species Concepts

One implication for the results presented here relates to concepts of speciation and species identification. While some of the problems discussed below might be interpreted as special cases that apply only to polyploid taxa, the general discussion

could be seen to apply to the overall problem of taxonomic naming if a single concept of what defines a "species" is applied to all types of independently evolving units.

Most discrepancies in species definitions have resulted from concepts of species that are currently geographically or functionally isolated and are thought to have arisen allopatrically. Specifically, the relative importance of species isolating mechanisms to species identification or speciation has been a point of contention. Polyploid frogs may not be postzygotically isolated from closely related diploid species (although they may be functionally isolated because they would not normally encounter these species) but have been shown to be prezygotically and/or postzygotically isolated from the species with which they are sympatric (e.g., Mable 1989; Mable and Bogart 1995). Because sympatric speciation by any definition requires a separation of gene pools sharing the same physical and reproductive space, some form of isolation is necessary between them. Whether or not identification of isolating mechanisms is critical to the delineation of species boundaries, their existence is essential to the underlying evolutionary process. Therefore, it is not surprising that sympatrically occurring diploids and tetraploids within the groups of frogs we examined produce mating calls which are recognizable from each other in most cases. In addition, it is not known how much introgression might occur between the various species of *Neobatrachus* but in *Hyla* it is not thought to be common, emphasizing the importance of reproductive isolation in sympatry. This introduces another conceptual problem: if species are entities that can be identified by specific characteristics, which characteristics are the most important, especially under sympatric conditions? Should species status only be conferred when it can be documented that pre- or postzygotic isolating mechanisms exist or is it sufficient that

genetic changes are apparent between the populations and isolating mechanisms only invoked as a means of explaining how gene pools have been isolated to allow independent evolution? This discrepancy has been central to the "species problem" debate, although from a functional standpoint, it is only an issue when comparing populations that are in potential contact, as geographic isolation alone is probably sufficient to allow genetic divergence. However, an issue that is central to all types of species-forming events is how do you know when what you are looking at should be classified as a "species", rather than some other taxonomic subdivision?

The two groups of polyploids examined in this thesis provide a good demonstration of the difficulty in establishing semantic classification of identified lineages. The original "concept" of species was part of a hierarchical classification of types that could be identified (mainly by morphological differences) from other such types. There was little philosophy involved in this process, and in accordance with the Linnaean code (still in use in the present) each type was assigned a binomial name to identify it as a particular recognizable type. The first part of the name identified a larger class of similar organisms, the genus, and the second identified the finer scale class, species. The original debate over species concepts arose out of recognition that there are other important biological differences that separate organisms at the lowest level of classification (i.e., species) beyond simply morphological differentiation. The type concept was replaced by the biological species concept (Mayr 1942), which emphasized the importance of reproductive isolating mechanisms to the separation of types that merited species status. With the improvement and more widespread use of systematic methods that emphasized the evolutionary histories and relatedness of organisms, the biological species concept was also questioned because it was realized

that differentiation by geographic separation did not require actual development of identifiable isolating mechanisms and not all independently evolving units could be characterized by reproductive isolating mechanisms (e.g., Erhlich 1961). The evolutionary species concept (Simpson 1961; Ghiselin 1969; Wiley 1978; 1980) ascribed less importance to isolating factors and emphasized that isolation of gene pools was the critical factor. Other lineage-based concepts have also been proposed (e.g., phylogenetic species concept: Cracraft 1983; Nixon and Wheeler 1990), some of which do not invoke any more biology than did the original type concept. There has also been discussion of the fact that it is necessary to identify both factors which result in the separation of gene pools and those that maintain cohesiveness within gene pools (e.g., recognition concept: Paterson 1985). There have also been philosophical debates about whether species should be viewed as individuals or classes (e.g., Ghiselin 1966, 1975; Hull 1976; Frost and Hillis 1990) and whether present attributes or predictions of future interactions are the essential factors to identify (e.g., Frost and Hillis 1990). Several authors have more recently suggested that the debate is not a realistic one because "species" is too broad a category, and it might be more useful to have several systems of classification to identify particular types of organisms or particular types of interactions among organisms (e.g., O'Hara 1993; Graybeal 1995). Unfortunately, debates on these presentations have tended to center on the particular names chosen rather than the real take-home message, which is that not all "species" have arisen through the same evolutionary mechanisms so it is unrealistic to think that one definition could adequately describe and classify them all. In addition, "species" delineation may vary depending on what scale of measurement is used to draw the boundaries, no matter which definitions are applied.

For example, the grey treefrogs of North America were originally classified as a single type under the name *Hyla versicolor* Le Conte because populations from across the continent are morphologically indistinguishable and they form a clear morphological type that is different from any other treefrogs. Anyone who picked up a frog that fit the description (e.g., from a field guide or morphological key) could confidently say "This is a grey treefrog". However, in the late 1950s and early 1960s when it first became recognized that frogs were capable of choosing mates based on differences in the calls that they produced (Blair 1958; Littlejohn 1960), two "call races" were identified (Blair 1958). It was further found that these call races produced sterile offspring when they were artificially hybridized (Johnson 1959; 1963). Following the biological species concept, these two forms were classified as potential species because they showed prezygotic and postzygotic isolation from one another: the slower trilling species was given the original name and the faster trilling species was given the name *Hyla chrysoscelis* (Johnson 1966). Recognition of the two species then depended on the ability to distinguish between their calls. However, picking up a female (who does not call) or a museum specimen would not allow distinction of the two species and would only allow application of the name that might be considered as a class distinction, "grey treefrog". It was subsequently discovered that the slower trilling species was a tetraploid and could be reliably distinguished from the faster trilling one by the fact that it had twice as many chromosomes (Wasserman 1970; Bogart and Wasserman 1972). Museum specimens could now be reliably classified based on corresponding differences in cell size and nucleolar number but live females could still only be grouped to the class, "grey treefrog".

Further genetic and behavioural analysis complicated this picture because it was suggested that the diploid (the faster trilling species) appeared to be divided into an eastern and western form that could be distinguished by allozymes (e.g., Ralin 1978; Ralin and Selander 1979), immunological characteristics (Maxson *et al.* 1977) and differences in duration of mating calls (Gerhardt 1974). It was further suggested that the tetraploids may have arisen through hybridization of these two diploid forms. An extensive debate ensued in the late 1970s because researchers did not agree on whether the two forms of the diploid actually deserved classification as a species. This is where the confusion begins. If the two diploid "forms" warrant species status, and the tetraploids arose through hybridization between them, how many species should be recognized in this group? In addition, what if the tetraploids actually arose separately from each species, or had separate origins from different hybrid combinations? Although researchers did not agree with one another on the relative distinctiveness of the diploids and how many tetraploid origins had probably occurred (mainly because of differences in interpretation of allozyme patterns), the very safe classification "complex" was usually used to describe the class, "grey treefrog". This issue was left largely unresolved until the mitochondrial DNA sequence study of Ptacek *et al.* (1994). In this study, mitochondrial haplotypes supported the distinction of the diploids into two clades: eastern and western (that generally agreed with previous boundary predictions based on earlier studies). Tetraploids from the east appeared to be most closely related to the western diploids and tetraploids from the southwest appeared to be most closely related to the eastern diploids, while tetraploids from the north did not appear to be closely related to either group. The mitochondrial study in this dissertation was unable to identify a diploid

relative for the northern tetraploids but supported the eastern and western division of diploids and the suggestion that tetraploids were most distantly related to the diploids with which they are currently sympatric (Chapter 2). Proponents of a strict lineage-based species concept might suggest that sufficient information has thus been provided to propose attaching additional species names to the organisms within this complex: two to the diploids and three to the tetraploids. However, within the northern tetraploids, there also appeared to be some degree of divergence and it is possible that more than three mitochondrial tetraploid lineages could be considered. However, because a putative diploid ancestor is lacking in this region and tetraploids are more closely related to each other than to diploids from other regions it is difficult to know where to "draw the line".

This brings up a very important problem with lineage-based species concepts: how do you determine when sufficient variation exists among organisms to warrant division into independently evolving lineages? In the Australian genus *Neobatrachus*, the mitochondrial phylogeny generated in this thesis suggests an opposite problem to the grey treefrogs: rather than predicting that not enough species have been named, it suggests that there are currently too many names. However, this could be interpreted as a question of scale. The relative degree of divergence in COI among the most distant lineages of grey treefrogs identified (and in cytb) was on the same scale as the within-population variation determined in the Australian group. Just for sake of argument, if we were to apply the same scale of comparison to both groups, different conclusions would be reached about the number of species in each. For example, let's say we were to conclude that populations in the "eastern" clade of tetraploids in the Australian group should be viewed as a single species with clinal variation across

its range (based on the relative degree of divergence within this clade compared to the other taxa in the genus). If we applied the same scale to the North American group we might conclude that only a single species was identified. However, we know that other factors (chromosome number, mating calls, hybrid sterility) clearly identify two species. By the same reasoning, if we defined five species in the North American group and applied that scale of divergence to the Australian group, we could not identify how many species actually existed in that group until we had done much more extensive population level sampling within nominal species. Although it is not realistic to assume one rate of evolution for organisms with very different evolutionary histories (such as the two groups studied here), the point is that lineage-based names are inherently subjective and may not be the most realistic way to identify species if they are viewed as individuals. Templeton (1994) suggested that rather than viewing species from the usual top down approach of hierarchy that the problem instead be viewed from the population level up. If species names are viewed as hypotheses, he suggests that their status could be tested by a rigorous system of criteria (Templeton 1994).

Another problem with species definitions for the polyploid species studied here has to do with what conclusions to draw when different types of data sets predict different things about the evolutionary history of the organisms under study. Mitochondrial DNA is maternally inherited and therefore provides only a partial picture of evolutionary interactions, especially in organisms that may have arisen through hybridization. It also may evolve independently of the nuclear genome and therefore a mitochondrial phylogeny may not always be congruent with the "host" species phylogeny. For example, in a study of unisexual salamanders in the genus

Ambystoma, Hedges *et al.* (1992) found that new combinations of nuclear genomes were continually being cycled through an ancient cytoplasm. In addition, the more rapid rate of evolution in mitochondrial DNA may reveal differences in taxonomic groups that are not identifiable by nuclear markers. For example, in the grey treefrogs, populations of diploid and tetraploids share most of the same alleles and are separated mainly by allelic frequency differences. The analysis presented in this thesis suggests that eastern and western forms can be identified but divergence of allozymic types does not appear to correspond with ploidy level. If we didn't know about the behavioural and chromosomal characters that separate diploids and tetraploids into independently evolving units, we might conclude that there are only two species in this complex (i.e., eastern and western), demonstrating the danger of relying on one type of character to distinguish evolutionary interactions. The allozyme study also suggested that tetraploids were most closely related to their diploid neighbours—directly contradicting the conclusions based on mtDNA. A potential dilemma thus arises: which data set (if any) reflects the true relationships among the taxa? In this case, mitochondrial and nuclear evidence would not be congruent but both provide meaningful information on the phylogenetic history of the taxa in question. The discrepancy found could be evidence of a hybrid origin for the tetraploids, in which case both data sets provide meaningful information. However, it could also reflect differences in relative rates of divergence of protein changes versus synonymous mtDNA changes or could be evidence that selection on allele frequencies has obscured actual relationships. If we were to add in the limited results of the nuclear gene study, another level of divergence would be added. From the nuclear intron regions from which sequences were obtained, all populations of grey

treefrogs had identical sequences and, for the Australian frogs, sequence divergence was extremely low within genera but there were vast differences between genera. If we were to use these results to define species, conclusions would be very different than using either of the other data sets. Admittedly, the discrepancies found in these studies may be attributed to the use of inappropriate molecular markers for the phylogenetic questions asked and/or differences among methods of analysis. However, problems of this type are increasingly being recognized as more DNA level phylogenetic comparisons are being made.

In conclusion, while it would be nice to think of "species" as real biological entities that satisfy the criteria of "individuals" we may not have the necessary tools to identify the limits to these entities. Naming is in itself an arbitrary process and unfortunately, the only tools we have available to identify "species" are also rather arbitrary criteria (e.g., choice of characters that are used to identify lineages). Therefore, perhaps the term "species" should be reserved for the definition to which it was originally applied; i.e., to distinguish classes of distinctive types that can be recognized by a set of fixed criteria. The study of evolutionary relationships and interactions among these types could be seen as a separate discipline in which various types of names are given to individuals showing particular patterns of evolution (e.g., Graybeal 1995). The names themselves do not matter, but the implications for elucidating the underlying evolutionary interactions do.

CONCLUSIONS

The original purpose of this dissertation was to determine phylogenetic relationships between tetraploid frogs and their diploid relatives as a framework in which to examine factors leading to speciation and subsequent consequences of polyploidy. However, discrepancies between allozyme and mitochondrial data sets in the North American group and limited phylogenetic resolution of tetraploid relationships to diploids for the Australian group did not allow conclusions beyond hypotheses of historical relationships within each group.

Conclusions that were supported were that tetraploids have arisen multiple times within each group and that patterns of relationships may be even more complex than previously suggested. For the Australian group, evidence from mitochondrial DNA, morphological differentiation, and mating call variation suggests that there are at least two (and probably three) lineages of tetraploids and that further speciation within the "eastern" lineage may have occurred following the polyploidization event. For the North American group, inclusion of more northernly distributed populations to the previous body of knowledge suggested that there be more lineages of tetraploids than previously predicted. The inability to definitively identify a diploid progenitor for the tetraploids in either group could be due to incomplete taxonomic sampling or to limited phylogenetic resolution, but could also be evidence that tetraploids have replaced their diploid progenitors. Based on degree of divergence in mitochondrial DNA, speciation of tetraploids within the Australian group appeared to be a much more older event than that within the North American group.

Conclusions based on nuclear evidence were less clear. Intron regions of single-copy nuclear genes did not appear to have sufficient variation to be phylogenetically informative. Allozymes may have limited utility because diploids and tetraploids share most alleles. In the North American group, allozymes suggested that tetraploids tended to be most similar to the diploids with which they were sympatric whereas mitochondrial data suggested that they were more closely related to tetraploids from other regions than they were to their sympatric diploids. While this could provide evidence that tetraploids have arisen through hybridization, other explanations are possible given that allozymes represent gene products that may be subject to selection or differential expression.

Table 5.1. Comparison of synonymous per synonymous site versus nonsynonymous changes per synonymous site in cytochrome oxidase 1 (COI) and cytochrome b (cyt b). Note the large increase in the ratio within the ingroup when the outgroup is removed for COI but not for cytb. Also note that the ratio appears to change roughly in proportion to the degree of divergence in cytochrome oxidase: the shortest branch lengths are found among the eastern tetraploids while the longest branches are among the diploids and the western tetraploids.

Taxa included	**C01**			**Cyt b**		
	Nonsynonymous	Synonymous	Ratio	Nonsynonymous	Synonymous	Ratio
Neobatrachus-all taxa	0.028	0.522	18.6	0.034	0.522	15.4
No outgroups	**0.008**	**0.463**	**57.9**	**0.029**	**0.459**	**15.8**
Tetraploids only	0.007	0.360	51.4			
Eastern tetraploids only	0.002	0.145	72.5			
Diploids only	0.012	0.494	41.2			
Diploids + Eastern 4n	0.008	0.480	60.0			
Diploids + Western 4n	0.013	0.478	36.8			
Hyla	**0.001**	**0.082**	**82.0**			

Appendix A : Parsimony Analysis of Allozyme Data

```
#NEXUS
BEGIN DATA;
DIMENSIONS  NTAX=18 NCHAR=16;
FORMAT MISSING=? GAP=-  SYMBOLS= " A B C D E F G H I J K L M N O P Q R";

CHARLABELS
      [1]   6PGD  [2]   'Acon-1'    [3]   'Acon-2'    [4]   'Got-1'     [5]   'Got-2'
      [6]   'Idh-1'     [7]   'Idh-2'     [8]   'Mdh-1'     [9]   'Mdh-2'     [10]  'Mpi-2'
      [11]  'Pgi-1'     [12]  'Pgi-2'     [13]  'Pgm-1'     [14]  'Pgm-2'     [15]  'Sod-1'     [16]  aGpd;

MATRIX
[                       10    ]
wisconsin1.2n    AAAAAAAAAAAAAAAA
wisconsin2.4n    BABBBAABABAABABA
wisconsin3.2n    CACACAACACAAACCC
minnesota.4n     DDDDDDDDDDDDDDDA
minnesota.2n     EEEACAEEAEAEEAEA
manitoba.4n      FFFAFFFFAFFFFFFF
manitoba.2n      GGGACAGGAGAGGAEG
virginia.4n      HHHAHHHHAHHHHHHA
virginia.2n      IIIAIIIHAIAAIIII
maryland.2n      JAJJCJJHAJJJJJJJ
LaAqua.4n        KK?KCKAKAK??KKK?
aberfoyle.4n     LL?LCLLLALALLLL?
wvirginia.2n     GM?MCAAHAM??AMM?
wvirginia.4n     NN?NCAAHAN??NNN?
nmc.4n           OA?OCOAOAO??IOO?
guelph.4n        PP?PCAAPAP??PPP?
rosseau.4n       QQ?QCAAQAQ??QQQ?
michigan.4n      RR?RCRARAR??RRR?
;
END;

BEGIN ASSUMPTIONS;

USERTYPE I=2 [Mdh-2]
      A     D
[A]   .     005
[D]   005   .
```

;

USERTYPE A=17 [6Pgd]

	A	B	C	D	E	F	G	H	I	J	K	L	N	O	P	Q	R
[A]	.	044	007	022	017	036	025	001	004	007	023	008	008	014	008	011	009
[B]	044	.	051	022	061	008	019	043	040	037	067	036	052	044	036	033	035
[C]	007	051	.	029	010	043	.32	008	011	014	016	015	001	021	008	018	016
[D]	022	022	029	.	038	014	003	021	018	015	049	014	030	007	014	022	013
[E]	017	061	010	038	.	053	042	018	021	024	006	025	008	031	025	028	017
[F]	036	008	043	014	053	.	011	035	032	029	059	028	044	022	028	025	027
[G]	025	019	.32	003	042	011	.	024	021	018	048	017	033	011	017	014	016
[H]	001	043	008	021	018	035	024	.	003	006	024	007	009	013	007	010	008
[I]	004	040	011	018	021	032	021	003	.	003	027	004	013	010	004	007	005
[I]	007	037	014	015	024	029	018	006	003	.	001	001	015	008	001	004	002
[K]	023	067	016	049	006	059	048	024	027	001	.	015	037	031	034	032	050
[L]	008	036	015	014	025	028	017	007	004	001	015	.	016	006	0	003	0
[N]	008	052	001	030	008	044	033	009	013	015	037	016	.	023	016	019	017
[O]	014	044	021	007	031	022	011	013	010	008	031	006	023	.	006	003	005
[P]	008	036	008	014	025	028	017	007	004	001	034	0	016	006	.	003	001
[Q]	011	033	018	022	028	025	014	010	007	004	032	003	019	003	003	.	002
[R]	009	035	016	013	017	027	016	008	005	002	050	0	017	005	001	002	.

;

USERTYPE B=14 [Acon-1]

	A	D	E	F	G	H	I	K	L	M	N	P	Q	R
[A]	.	010	017	011	007	010	002	004	004	050	013	006	003	002
[D]	010	.	007	001	006	008	008	010	007	040	003	004	010	010
[E]	017	007	.	006	010	015	015	017	014	033	004	010	017	017
[F]	011	001	006	.	006	009	009	011	008	039	002	005	011	011
[G]	007	006	010	006	.	006	006	007	006	043	006	006	007	007
[H]	010	008	015	009	006	.	008	006	006	048	011	008	007	008
[I]	002	008	015	009	006	008	.	004	002	048	010	004	003	002
[K]	004	010	017	011	007	006	004	.	002	050	013	006	001	002
[L]	004	007	014	008	006	006	002	002	.	048	010	004	002	002
[M]	050	040	033	039	043	048	048	050	048	.	038	044	050	050
[N]	013	003	004	002	006	011	010	013	010	038	.	006	013	013
[P]	006	004	010	005	006	008	004	006	004	044	006	.	006	006
[Q]	003	010	017	011	007	007	003	001	002	050	013	013	.	001
[R]	002	010	017	011	007	008	002	002	002	050	013	006	001	.

;

TYPESET *ALLWTS = A:1, B:2, C:3, D:4, E:5, F:6, G:7, H:8, I:9, J:10, K:11, L:12, M:13, N:14, O:15, P:16;

Appendix B : PAUP* File for COI Sequences of Grey Treefrogs

```
#NEXUS
BEGIN DATA;
DIMENSIONS  NTAX=12 NCHAR=590;
FORMAT MISSING=? GAP=-  MATCHCHAR=.  INTERLEAVE  DATATYPE=DNA;

MATRIX

minnesota.4n    CACCCAGAAG TTTACATTCT TATTCTCCCT GGGTTCGGCA TTATTTCTCA   [50]
nontario.4n     .......... .......... .......... .......... ..........   [50]
sontario.4n     .......... ....T..... .......... .......... ..........   [50]
minnesota.2n    .......... ....T..... .........C .......... ..........   [50]
etexasL.2n      .......... ....T..... .........C .......... ..........   [50]
ctexas.4n       .......... ....T..... .......... .......... ..........   [50]
etexC.2n        .......... ....T..... .........C .......... ..........   [50]
wisconsin.2n    .......... .......... .........C .......... ..........   [50]
wisconsin.4n    .......... .......... .........C .......... ..........   [50]
michigan.2n     .......... .......... .........C .......... ..........   [50]
michigan.4n     .......... .......... .......... .......... ..........   [50]
ctexas.2n       .......... .......... .........C .......... ..........   [50]

minnesota.4n    TGTAGTAACA TTTTACTCAA GTAAAAAGGA GCCCTTTGGC TATATAGGAA   [100]
nontario.4n     .......... .......... .......... .......... ..........   [100]
sontario.4n     .......... .......... .......... .......... ..........   [100]
minnesota.2n    .......... .......... .......... .......... ..........   [100]
etexasL.2n      .......... .......... .......A.. .......... .....G..G.   [100]
ctexas.4n       .......... .......... .......... .......... ..........   [100]
etexC.2n        .......... .......... .......... .......... ..........   [100]
wisconsin.2n    .......... .......... .......... .......... ..........   [100]
wisconsin.4n    .......... .......... .......... .......... ..........   [100]
michigan.2n     .......... .......... .......... .......... ..........   [100]
michigan.4n     .......... .......... .......... .......... ..........   [100]
ctexas.2n       .......... .......... .......... .......... ..........   [100]

minnesota.4n    TAGTATGAGC TATAATATCT ATTGGACTTC TAGGCTTCAT CGTCTGAGCA   [150]
nontario.4n     .......... .......... .......... .......... ..........   [150]
sontario.4n     .......... ...G...... .......... .......... ..........   [150]
minnesota.2n    .......... C......... .......... ....G..... ..........   [150]
etexasL.2n      .......... C..G...... .......... ....G..... ..........   [150]
ctexas.4n       .......... .......... .......... ....G..... ..........   [150]
etexC.2n        .......... C......... .......... ....G..... ..........   [150]
wisconsin.2n    .......... C......... .......... ....G..... ..........   [150]
wisconsin.4n    .......... C......... .......... ....G..... ..........   [150]
michigan.2n     .......... C......... .......... ....G..... ..........   [150]
michigan.4n     .......... ...G...... ..G....... .......... ..........   [150]
ctexas.2n       .......... C......... .......... ....G..... ..........   [150]

minnesota.4n    CATCATATAT TTACAACTGA CTTAAACGTA GACACCCGAG CCTACTTCAC   [200]
nontario.4n     .......... .......... .......... .......... ..........   [200]
sontario.4n     .......... .......... .......... .......... ..........   [200]
minnesota.2n    ........G. .......... .......... .......... ..........   [200]
etexasL.2n      ........G. .......... .......... .......... .......T..   [200]
ctexas.4n       .......... .......... .......... .......... ..........   [200]
etexC.2n        ........G. .......... .......... .......... .......T..   [200]
wisconsin.2n    ........G. .......... .......... .......... ..........   [200]
wisconsin.4n    ........G. .......... .......... .......... ..........   [200]
michigan.2n     ........G. .......... .......... .......... ..........   [200]
```

```
michigan.4n     ........?? ?????????? ?......... .......... ..........   [200]
ctexas.2n       .......... .......... .......... .......... ..........   [200]

minnesota.4n    CTCTGCCACT ATAATTATTG CCATCCCAAC AGGGGTAAAA GTATTTAGCT   [250]
nontario.4n     .......... .......... .......... .......... ..........   [250]
sontario.4n     .......... .......... .......... .......... ..........   [250]
minnesota.2n    .......... .......... .T........ .......... ..........   [250]
etexasL.2n      .......... .......... .T........ .......... ..........   [250]
ctexas.4n       .......... .......... .......... ....T..... ..........   [250]
etexC.2n        .......... .......... .T........ .......... ..........   [250]
wisconsin.2n    .......... .......... .T........ .......... ..........   [250]
wisconsin.4n    .......... .......... .T........ .......... ..........   [250]
michigan.2n     .......... .......... .T........ .......... ..........   [250]
michigan.4n     .......... .......... .......... .......... ..........   [250]
ctexas.2n       .......... .......... .T........ .......... ..........   [250]

minnesota.4n    GACTGGCTAC ATGGAGGAAT TATTAAATGA GACGCTGCTA TGCTATGAGC   [300]
nontario.4n     .......... .......... .......... .......... ..........   [300]
sontario.4n     .......... .......... .......... .......... .......G..   [300]
minnesota.2n    .......... .......... .......... .......... ..........   [300]
etexasL.2n      ....A..... .......... .......... ......A... ....C.....   [300]
ctexas.4n       .......... .......... .......... .......... ..........   [300]
etexC.2n        .......... .......... .......... ......A... ..........   [300]
wisconsin.2n    ....A..... .......... .......... ......A... ..........   [300]
wisconsin.4n    ....A..... .......... .......... ......A... ..........   [300]
michigan.2n     ....A..... .......... .......... ......A... ....T.....   [300]
michigan.4n     .......... .......... .......... .......... ..........   [300]
ctexas.2n       ....A..... .......... .......... ......A... ..........   [300]

minnesota.4n    TCTGGGCTTT ATTTTCTTAT TTACAGTAGG CGGCTTGACA GGAATTGTTT   [350]
nontario.4n     .......... .......... .......... .......... ..........   [350]
sontario.4n     .......... .......... .......... ......A... ........C.   [350]
minnesota.2n    ...A...... ........G. .......... ......A... ..........   [350]
etexasL.2n      ...A...... ........G. .......... ......A... ..........   [350]
ctexas.4n       .......... .......... .......... ......A... ..........   [350]
etexC.2n        ...A...... .......... .......... ......A... ..........   [350]
wisconsin.2n    ...A...... .......... .......... ......A... ..........   [350]
wisconsin.4n    ...A...... .......... .......... ......A... ..........   [350]
michigan.2n     ...A...... ........G. .......... ......A... ..........   [350]
michigan.4n     .......... .......... .......??. ......A... ..........   [350]
ctexas.2n       ...A...... .......... .......... ......A... ..........   [350]

minnesota.4n    TAGCTAATTC ATCCCTGGAC ATTGTTTTAC ATGACACCTA CTATGTTGTA   [400]
nontario.4n     .......... .......... .......... .......... ..........   [400]
sontario.4n     .......... ...T...... .......... .......... ..........   [400]
minnesota.2n    .......... .......... .......... .......... ..........   [400]
etexasL.2n      .......... .......... .......... .......... ..........   [400]
ctexas.4n       .......... .......... .......... .......... ..........   [400]
etexC.2n        .......... .......... .......... .......... ..........   [400]
wisconsin.2n    .......... .......... .......... .......... ..........   [400]
wisconsin.4n    .......... .......... .......... .......... ..........   [400]
michigan.2n     .......... .......... .......... .......... ..........   [400]
michigan.4n     .......... ...T...... .......... .......... ..........   [400]
ctexas.2n       .......... .......... .......... .......... ..........   [400]
```

```
minnesota.4n    GCCCACTTTC ATTATGTCCT TTCTATAGGC GCGGTATTCG CAATTATAGC   [450]
nontario.4n     .......... .......... .......... .......... ..........   [450]
sontario.4n     .......... .......... .......... .......... ..........   [450]
minnesota.2n    .......... .......... .......... .......... ..........   [450]
etexasL.2n      .......... .......T.. .......... .......... ..........   [450]
ctexas.4n       .......... .......... .......... .......... ..........   [450]
etexC.2n        .......... .......... .......... .......... ..........   [450]
wisconsin.2n    .......... .......T.. .......... .......... ..........   [450]
wisconsin.4n    .......... .......T.. .......... .......... ..........   [450]
michigan.2n     .......... .......T.. .......... .......... ..........   [450]
michigan.4n     .......... .......... .......... .......... ..........   [450]
ctexas.2n       .......... .......T.. .......... .......... ..........   [450]

minnesota.4n    AGGATTTGTT CACTGATTTC CGCTATTTAC AGGATTCACC CTTCATAAAA   [500]
nontario.4n     .......... .......... .......... .......... ..........   [500]
sontario.4n     .......... .......... .......... .......... ..........   [500]
minnesota.2n    .......... .......... .A........ .......... ..........   [500]
etexasL.2n      .......... .......... .A........ .......... ..........   [500]
ctexas.4n       .......... .......... .A........ .......... ..........   [500]
etexC.2n        .......... .......... .A........ .......... ..........   [500]
wisconsin.2n    .......... .......... .A........ .......... ..........   [500]
wisconsin.4n    .......... .......... .A........ .......... ..........   [500]
michigan.2n     .......... .......... .A........ .......... ..........   [500]
michigan.4n     .......... .......... .......... .......... ..........   [500]
ctexas.2n       .......... .......... .A........ .......... ..........   [500]

minnesota.4n    CTTGAGCAAA AATCCAATTT GGAGTTATGT TCACTGGAGT AAATCTAACA   [550]
nontario.4n     .......... .......... .......... .......... ..........   [550]
sontario.4n     .......... .......... .......... .......... ..........   [550]
minnesota.2n    .......... ...T...... .......... ....C..... ..........   [550]
etexasL.2n      .......... ...T...... ..G....... ....C..... ..........   [550]
ctexas.4n       .......... ...T...... .......... .......... ..........   [550]
etexC.2n        .......... ...T...... .......... ....C..... ..........   [550]
wisconsin.2n    .......... ...T...... ..G....... ....C..... ..........   [550]
wisconsin.4n    .......... ...T...... ..G....... ....C..... ..........   [550]
michigan.2n     .......... ...T...... ..G....... ....C..... ..........   [550]
michigan.4n     .......... .......... .......... .......... ..........   [550]
ctexas.2n       .......... ...T...... ..G....... ....C..... ..........   [550]

minnesota.4n    TTCTTCCCAC AACATTTCCT GGGCTTAGCC GGGATACCAC   [590]
nontario.4n     .......... .......... .......... ..........   [590]
sontario.4n     .......... .......... .......... ..........   [590]
minnesota.2n    .......... .......... ....C..... ..A.......   [590]
etexasL.2n      .......... .......... A...C..... ..A.......   [590]
ctexas.4n       .......... .......... ....C..... ..........   [590]
etexC.2n        .......... .......... ....C..... ..A.......   [590]
wisconsin.2n    .......... .......... A...C..... ..A.......   [590]
wisconsin.4n    .......... .......... A...C..... ..A.......   [590]
michigan.2n     .......... .......... A...C..... ..A.......   [590]
michigan.4n     .......... .......... ....C..... ........T.   [590]
ctexas.2n       .......... .......... A...C..... ..A.......   [590]
;
END;
```

```
BEGIN ASSUMPTIONS;

      USERTYPE a STEPMATRIX = 4
           A C G T
       [A] . 1 0 1
       [C] 1 . 1 0
       [G] 0 1 . 1
       [T] 1 0 1 .
      ;

      USERTYPE b STEPMATRIX = 4
           A C G T
       [A] . 7 1 7
       [C] 7 . 7 1
       [G] 1 7 . 7
       [T] 7 1 7 .
      ;

      OPTIONS  DEFTYPE=unord PolyTcount=MINSTEPS ;

      CHARSET  pos1  =  1-589\3;
      CHARSET  pos3  =  3-588\3;

      TYPESET  notrans  = a:1-590;
      TYPESET  ts7  = b:1-590;

END;
```

Appendix C : PAUP* File for COI Sequences of *Neobatrachus*

```
#NEXUS

BEGIN DATA;
DIMENSIONS  NCHAR=568 NTAX=31;
FORMAT MISSING=? GAP=-  MATCHCHAR=.  INTERLEAVE  DATATYPE=DNA ;
OPTIONS  MSTAXA=UNCERTAIN ;

MATRIX

[                          10         20         30         40         50]
[                          .          .          .          .          .]
'sudelli-SA1'         CCTGAGGTCTACATTCTCATCCTTCCCGGATTTGGCATTATCTCTCATGT   [50]
'sudelli-SA2'         ..................................................   [50]
'sudelli-QLD1'        ..................T........G.......................   [50]
'aquilonius-WA1'      ..G..A..T........T..T..C..A..C.............C..C..   [50]
'aquilonius-WA2'      ..G..A..T..T.....T..T..C..A..C.............C..C..   [50]
'centralis-SA2'       ................T..T........G.......................   [50]
'kunapalari-WA1'      ..........T.............A..C.....T..C.....C..C..   [50]
'kunapalari-WA2'      ..........T.............A..C.....T..C.....C..C..   [50]
'albipes-WA1'         ..C........T.............A..C.....T..C.....C..C..   [50]
'albipes-WA2'         ..C........T.............A..C.....T..C.....C..C..   [50]
'wilsmorei-WA1'       ..........T.....T.....C..G..C.....T........C.....   [50]
'wilsmorei-WA2'       ..........T.....T.....C..G..C.....T........C.....   [50]
'pictus-Vic1'         ..G.....A..T.....T.....C.....C..C.....C..T..C..C..   [50]
'pictus-Vic2'         ..G.....A..T.....T.....C.....C..C.....C..T..C..C..   [50]
'pelobatoides-WA1'    ..C..A..........G..T..C..A..C.....T........C..C..   [50]
'pelobatoides-WA2'    ..G..A..............T..C..A..C.....T........C..C..   [50]
'centralis-WA1'       .....A............................................   [50]
'centralis-WA2'       .....A............................................   [50]
'centralis-SA1'       ..................T........G.......................   [50]
'fulvus-WA1'          ..G..A..T..T.....T.....C..A..C..C.....C.....C..C..   [50]
'fulvus-WA2'          ..G..A..T..T.....T..T..C..A..C..C..........C..C..   [50]
'fulvus-WA3'          ..G........T.....?........A..C........C.....C..C..   [50]
'fulvus-WA4'          ..G..A..T..T.....T..T..C..A..C..C.....C.....C..C..   [50]
'sutor-WA'            ..........T.....T..T..C.....C.....T........A..C..   [50]
'Heleioporus-1'       ......................A.................T........   [50]
'Heleioporus-2'       ..G..A..T..T.....T..T..C..A..C.............C..C..   [50]
'Heleioporus-3'       ..G........T.....T........A..C........C.....C..C..   [50]
Notaden               ..A..A.....T.....T..TT.A..A........A.....T..A.....   [50]
Xenopus               .......A..G.............A..A..G........G.....C...   [50]
Litoria               .......A..T.............A..A........A........C...   [50]
'aquilonius-NT'       ..............................G...................   [50]


[                          60         70         80         90        100]
[                          .          .          .          .          .]

'sudelli-SA1'         AGTAGCCTACTACTCAAGCAAAAAAGAACCGTTTGGCTACATGGGCATAG   [100]
'sudelli-SA2'         ..................................................   [100]
'sudelli-QLD1'        .................................................G.   [100]
'aquilonius-WA1'      ......T................G..G.......................   [100]
'aquilonius-WA2'      ......T..................G.......................   [100]
'centralis-SA2'       .................................................G.   [100]
'kunapalari-WA1'      ..................T.....G.............T..........   [100]
'kunapalari-WA2'      ..................T.....G..G..........T..A.....G.   [100]
'albipes-WA1'         ..................T.....G..G..........T........G.   [100]
'albipes-WA2'         ..................T.....G.............T........G.   [100]
'wilsmorei-WA1'       ............T.....T........G.....C.....T..A.....G.   [100]
'wilsmorei-WA2'       ............T.....T........G.....C.....T..A.....G.   [100]
'pictus-Vic1'         ......................G..G..A........T.....T....   [100]
'pictus-Vic2'         ......................G..G..A........T.....T....   [100]
```

```
'pelobatoides-WA1'    G..G.....................G.....C...........A.....G.    [100]
'pelobatoides-WA2'    G.......................G.....C...........A.....G.    [100]
'centralis-WA1'       ..................................................    [100]
'centralis-WA2'       ..................................................    [100]
'centralis-SA1'       ..............................................G...    [100]
'fulvus-WA1'          ........................G..G.....................    [100]
'fulvus-WA2'          ......T.................G..G....................G?    [100]
'fulvus-WA3'          ........................G..G.....................    [100]
'fulvus-WA4'          ......T.................G..G.....................    [100]
'sutor-WA'            G.................T...........A.................G.    [100]
'Heleioporus-1'       ........................G.....................G.    [100]
'Heleioporus-2'       ......T.................G..G.....................    [100]
'Heleioporus-3'       ........................G..G.....................    [100]
Notaden               ....A.......T...G.A...........A..C..A..T..A..A....    [100]
Xenopus               ..C...A.T..T......G...............C........A..A..    [100]
Litoria               ......A.T..T.....T..T.....G..G.....................    [100]
'aquilonius-NT'       ..................................................    [100]

[                              110         120         130         140         150]
[                              .           .           .           .           .]

'sudelli-SA1'         TATGAGCCATGATATCTATCGGCTTCCTGGGCTTTATTGTGTGGGCTCAC    [150]
'sudelli-SA2'         ..??????????????????????????????????????????????    [150]
'sudelli-QLD1'        ..........................A.........C.....A......    [150]
'aquilonius-WA1'      .G...........G..A..T..................C........C...    [150]
'aquilonius-WA2'      .G...........G..C..T..................C........C...    [150]
'centralis-SA2'       ............................A........C.....A......    [150]
'kunapalari-WA1'      .G...........G..A..T..................C........C...    [150]
'kunapalari-WA2'      .G.............C...........C........C..A.....C...    [150]
'albipes-WA1'         .G.............C...........C........C..A.....C...    [150]
'albipes-WA2'         .G.............C...........C........C..A.....C...    [150]
'wilsmorei-WA1'       .G........A..G..A...........A.....C..C.....A.....T    [150]
'wilsmorei-WA2'       .G........A..G..A...........A.....C..C.....A.....T    [150]
'pictus-Vic1'         .G..G...........C........T...........C.....A.....T    [150]
'pictus-Vic2'         .G..G...........C........T..A........C.....A.....T    [150]
'pelobatoides-WA1'    .G...........G..C........T...........C.....A......    [150]
'pelobatoides-WA2'    .G...........G..C........T...........C.....A......    [150]
'centralis-WA1'       ............................A........C.....A......    [150]
'centralis-WA2'       ..............??????????????????????????????????    [150]
'centralis-SA1'       ............................A........C.....A......    [150]
'fulvus-WA1'          .............G..A..................A...C..A..A..C..T    [150]
'fulvus-WA2'          ??????????????????????????????????????????????????    [150]
'fulvus-WA3'          .............G..A.....................C..A..A..C..T    [150]
'fulvus-WA4'          .............G..A..T..................C..A..A..C..T    [150]
'sutor-WA'            .G..G.....A.....A.....A..............C..A..A..C..T    [150]
'Heleioporus-1'       ............................A..............A......    [150]
'Heleioporus-2'       .G...........G..A..T..................C..A.....C...    [150]
'Heleioporus-3'       ............??????????????????????..C..A.....C...    [150]
Notaden               .T.....T..A.....C..T..T.....A..T..C.....A..A..C..T    [150]
Xenopus               ...C..G..A..A.....A..T..A....CA...........C..A..C.    [150]
Litoria               G..T........A.....C..T.....A..A..A...........A..C.    [150]
'aquilonius-NT'       G.............................A........C.....A....    [150]

[                              160         170         180         190         200]
[                              .           .           .           .           .]

'sudelli-SA1'         CATATGTTCACTACAGACCTAAACGTCGACACCCGAGCCTACTTTACATC    [200]
'sudelli-SA2'         ??????????????????????????????????????????????????    [200]
'sudelli-QLD1'        ........T..C.....T..............T..................    [200]
'aquilonius-WA1'      ...........C..............T.....T..G........C..G..    [200]
'aquilonius-WA2'      ...........C..............T.....T..G........C..G..    [200]
'centralis-SA2'       ........T..C.....................T..................    [200]
```

```
'kunapalari-WA1'      ..........C.............T.....T..G........C..G..  [200]
'kunapalari-WA2'      ..C........C.............T.......................  [200]
'albipes-WA1'         ..C........C.............T.......................  [200]
'albipes-WA2'         ..C........C.............T.......................  [200]
'wilsmorei-WA1'       ..C..A.....C..G..........T.......................  [200]
'wilsmorei-WA2'       ..C..A.....C..G..........T.......................  [200]
'pictus-Vic1'         ..C.....T.....G..........T........C...........G..  [200]
'pictus-Vic2'         ..C.....T.....G......??????????????????????????  [200]
'pelobatoides-WA1'    ..C..A.....................T.....................  [200]
'pelobatoides-WA2'    ..C..A.....................T.....................  [200]
'centralis-WA1'       ..C.....T..C.....................T..G.............  [200]
'centralis-WA2'       ??????????????????????????????????????????????  [200]
'centralis-SA1'       ..C.....T..C.....................T...............  [200]
'fulvus-WA1'          ..C........C.................G..T..G.............  [200]
'fulvus-WA2'          ??????????????????????????????????????????????  [200]
'fulvus-WA3'          ..C........C.................??.T...............  [200]
'fulvus-WA4'          ..C........C.................??.T...............  [200]
'sutor-WA'            ..C........C.........................T.....C.....  [200]
'Heleioporus-1'       ........T..C.....T.............T.................  [200]
'Heleioporus-2'       ..........C.............T.....T..G........C..G..  [200]
'Heleioporus-3'       ..C........C.....................................  [200]
Notaden               ..C.....T..AGT..GAA..G....A.....A.....A..T.......  [200]
Xenopus               .T..C..A..T..GGTT.............T..T.....T.........  [200]
Litoria               ....C..A..T..G..T...T.............T.....T.........  [200]
'aquilonius-NT'       ..........T.......................T...........Y..R  [200]

[                              210        220        230        240        250]
[                              .          .          .          .          .]

'sudelli-SA1'         AGCCACTATAATCATTGCTATCCCAACCGGCGTAAAAGTCTTTAGCTGAC  [250]
'sudelli-SA2'         ??????????????????????????????????????????????  [250]
'sudelli-QLD1'        ..................................G........C..T...T  [250]
'aquilonius-WA1'      ......AC.......C..C...........T..............T....  [250]
'aquilonius-WA2'      ......AC.......C..C.......................C..T...T  [250]
'centralis-SA2'       ...................................................??  [250]
'kunapalari-WA1'      ......A........C..C.............................T  [250]
'kunapalari-WA2'      ......A........C.................G..............T  [250]
'albipes-WA1'         ......A........C..C.............................T  [250]
'albipes-WA2'         ......A........C..C.....C.......................T  [250]
'wilsmorei-WA1'       ......A........C..A.....T..T.....................  [250]
'wilsmorei-WA2'       ......A........C..A.....T..T.....................  [250]
'pictus-Vic1'         ............T..C...........T.....................T  [250]
'pictus-Vic2'         ??????????????????????????????????????????????  [250]
'pelobatoides-WA1'    ......A..G.....C..C........T.....................  [250]
'pelobatoides-WA2'    ......Y.....??????????????????????????????????  [250]
'centralis-WA1'       ...................................................  [250]
'centralis-WA2'       ??????????????????????????????????????????????  [250]
'centralis-SA1'       ..........................T..............C......T  [250]
'fulvus-WA1'          ..................C...........T..........C...T....  [250]
'fulvus-WA2'          ??????????????????????????????????????????????  [250]
'fulvus-WA3'          ..................C...........T..............T....  [250]
'fulvus-WA4'          ..................Y...........??????????????????  [250]
'sutor-WA'            ......C........C..............T...................  [250]
'Heleioporus-1'       ...........................................C.......  [250]
'Heleioporus-2'       ......C........C..C...........T..............T....  [250]
'Heleioporus-3'       ......Y.....??????????????????????????????????  [250]
Notaden               ............G..C.....T..T..........C..............  [250]
Xenopus               .....A..A...........A..TY.T..A.....T.....A........  [250]
Litoria               .....A..A...........C........T...................  [250]
'aquilonius-NT'       ...................................................  [250]
```

```
[                          260         270         280         290         300]
[                          .           .           .           .           .]

'sudelli-SA1'              TGGCCACAATACACGGCGGAATCATTAAATGAGACGCCGCAATGCTATGA   [300]
'sudelli-SA2'              ??????????????????????????????????????????????????   [300]
'sudelli-QLD1'             ..................................................   [300]
'aquilonius-WA1'           .A........G......C.G..T.C..............G.........   [300]
'aquilonius-WA2'           .A........G........G..T..C..............G.........   [300]
'centralis-SA2'            ??????????????????????????????????????????????????   [300]
'kunapalari-WA1'           .A.....G?...................????????????????????   [300]
'kunapalari-WA2'           .A.....G?...................????????????????????   [300]
'albipes-WA1'              .A.....G.....................G.....................   [300]
'albipes-WA2'              .A.....G.....................G.....................   [300]
'wilsmorei-WA1'            .A........G.............................G..A..G..G   [300]
'wilsmorei-WA2'            .A........G.............................G..A..G..G   [300]
'pictus-Vic1'              .A..T.....G..............C..G..............A......   [300]
'pictus-Vic2'              ??????????????????????????????????????????????????   [300]
'pelobatoides-WA1'         .A........G..................G..G...........A.....G   [300]
'pelobatoides-WA2'         ??????????????????????????????????????????????????   [300]
'centralis-WA1'            ...............................G...........A......   [300]
'centralis-WA2'            ??????????????????????????????????????????????????   [300]
'centralis-SA1'            ..................................................   [300]
'fulvus-WA1'               .A........G........G..T..C..G.......G...G.....G...   [300]
'fulvus-WA2'               ??????????????????????????????????????????????????   [300]
'fulvus-WA3'               .A........G..............C...........?????????????   [300]
'fulvus-WA4'               ??????????????????????????????????????????????????   [300]
'sutor-WA'                 ........G.G............T..C..................A..G...   [300]
'Heleioporus-1'            ..................................................   [300]
'Heleioporus-2'            .A.....C..G..............C..............G.........   [300]
'Heleioporus-3'            ??????????????????????????????????????????????????   [300]
Notaden                    .A..A...C.G..T..A....AT..C.......T.A....CC.AT.....   [300]
Xenopus                    .T.A.GT...........T..G.CA..............C....A..T.   [300]
Litoria                    ......T..G...........C..................A......T...   [300]
'aquilonius-NT'            .T..........R.....................??????????????   [300]


[                          310         320         330         340         350]
[                          .           .           .           .           .]

'sudelli-SA1'              GCCCTGGGCTTCATCTTCCTCTTTACTGTCGGTGGTCTTACTGGAATCGT   [350]
'sudelli-SA2'              ??????????????????????????????????????????????????   [350]
'sudelli-QLD1'             .....A.............????????????Y...........C........   [350]
'aquilonius-WA1'           ...............A..T.A..C.....T.....C.....C.....T..   [350]
'aquilonius-WA2'           ..................TAT........Y..G..C.....C.....T..   [350]
'centralis-SA2'            ??????????????????????????????????????????????????   [350]
'kunapalari-WA1'           ??????????????????????????????????????????????????   [350]
'kunapalari-WA2'           ??????????????????????????????????????????????????   [350]
'albipes-WA1'              ...T..............TT.A.....C.....C..C........G.....   [350]
'albipes-WA2'              ...T..............TT.A.....C.....C..C........G.....   [350]
'wilsmorei-WA1'            ...T.A............TT.A.....C..T..C..C...........T..   [350]
'wilsmorei-WA2'            ...T.A............TT.A.....C..T..C..C...........T..   [350]
'pictus-Vic1'              .....A............TT.A.....C.....G..A..C.....G..T..   [350]
'pictus-Vic2'              ??????????????????????????????????????????????????   [350]
'pelobatoides-WA1'         ..................TT.A..C..C.....C..C..C..?..G.....   [350]
'pelobatoides-WA2'         ??????????????????????????????????????????????????   [350]
'centralis-WA1'            .....A..............T.....C....AC........?........   [350]
'centralis-WA2'            ??????????????????????????????????????????????????   [350]
'centralis-SA1'            .....A..............T..C..C.......................   [350]
'fulvus-WA1'               ....................TG..........G..C...C...........   [350]
'fulvus-WA2'               ??????????????????????????????????????????????????   [350]
'fulvus-WA3'               ??????????????????????????????????????????????????   [350]
'fulvus-WA4'               ??????????????????????????????????????????????????   [350]
'sutor-WA'                 ........G........TT.A..C..C..T..C..G..A..C.....T..   [350]
'Heleioporus-1'            .....A..............T............................   [350]
```

```
'Heleioporus-2'      .................TTAT.C?.....Y..G..C.....C.....T..   [350]
'Heleioporus-3'      ??????????????????????????????????????????????????   [350]
Notaden              .....A.....T......T.......A..A..G.A.T.A..C........   [350]
Xenopus              .....T.A...........T.G........A..A.....A..A......   [350]
Litoria              .......C...........T.G........T.....A..A..A......   [350]
'aquilonius-NT'      ??????????????????????????????????????????????????   [350]

[                              360        370        380        390        400]
[                              .          .          .          .          .]

'sudelli-SA1'        CCTAGCCAATTCCTCCCTAGACATTGTCCTCCACGACACCTACTATGTAG   [400]
'sudelli-SA2'        ??????????????????????????????????????????????????   [400]
'sudelli-QLD1'       ...........T...................T..T.............   [400]
'aquilonius-WA1'     .........C........G........T.....T.............   [400]
'aquilonius-WA2'     .........Y........G........T.....T.............   [400]
'centralis-SA2'      ??????????????????????..T........T.............   [400]
'kunapalari-WA1'     ??????????????????????????????????????????????????   [400]
'kunapalari-WA2'     ??????????????????????????????????????????????????   [400]
'albipes-WA1'        ...G.....C......T....T.....T..A.....T.....T.......   [400]
'albipes-WA2'        ...G.....C......T....T.....T..A.....T.....T.......   [400]
'wilsmorei-WA1'      T........C...........T..C.....A.................T.   [400]
'wilsmorei-WA2'      T........C...........T..C.....A.................T.   [400]
'pictus-Vic1'        T........C.........................T...........G.   [400]
'pictus-Vic2'        ??????????????????????????????????????????????????   [400]
'pelobatoides-WA1'   .........C..A.....G.......................T.......   [400]
'pelobatoides-WA2'   ??????????????????????????????????????????????????   [400]
'centralis-WA1'      .........C.............................T.............   [400]
'centralis-WA2'      ??????????????????????????.....................     [400]
'centralis-SA1'      ............T.......................T.............   [400]
'fulvus-WA1'         T..G.....C.................T........T.............   [400]
'fulvus-WA2'         ??????????????????????????????????????????????????   [400]
'fulvus-WA3'         ??????????????????????????????????????????????????   [400]
'fulvus-WA4'         ??????????????????????????????????????????????????   [400]
'sutor-WA'           .T....T..C.....T.............................C....   [400]
'Heleioporus-1'      ............T.....G.................T.............   [400]
'Heleioporus-2'      .........C........G........T.....T..G.............   [400]
'Heleioporus-3'      ??????????????????????????????????????????????????   [400]
Notaden              A..T.G...C..A..A..C..T.....A...........A.....C....   [400]
Xenopus              .....T........A..AGCTT.....A.A......ATAG..........   [400]
Litoria              ..A...........T.................T...........T.....   [400]
'aquilonius-NT'      ??????????????????????????????????????????????????   [400]

[                              410        420        430        440        450]
[                              .          .          .          .          .]

'sudelli-SA1'        TAGCTCACTTCCACTATGTTCTATCGATGGGCGCAGTTTTCGCCATTATA   [450]
'sudelli-SA2'        ??????????????????????????????????????????????????   [450]
'sudelli-QLD1'       ..........................A..R.....................   [450]
'aquilonius-WA1'     .G..............C..C.....A.................T......   [450]
'aquilonius-WA2'     .G..............C..C..G..A.................T......   [450]
'centralis-SA2'      ..........................A.......................G   [450]
'kunapalari-WA1'     ????????........C.....G.....................T......   [450]
'kunapalari-WA2'     ??????????????????????????????????????????????????   [450]
'albipes-WA1'        ....C...........C........A..A.................C..G   [450]
'albipes-WA2'        ....C...........C........A..A.................C..G   [450]
'wilsmorei-WA1'      ....C........T........G..A..A.................C..G   [450]
'wilsmorei-WA2'      ....C........T........G..A..A.................C..G   [450]
'pictus-Vic1'        ................C..C.....A.................T.....G   [450]
'pictus-Vic2'        ???????????????????C........A..............?...T.....G   [450]
'pelobatoides-WA1'   ....C..................G.....A..T..G..............G   [450]
'pelobatoides-WA2'   ??????????????????????????????????????????????????   [450]
'centralis-WA1'      ..................................................     [450]
```

```
'centralis-WA2'      ..................C.............................  [450]
'centralis-SA1'      .......................A......................G  [450]
'fulvus-WA1'         ..................C.....A.....?.?.........T......  [450]
'fulvus-WA2'         ??????????????????????????????????????..T......  [450]
'fulvus-WA3'         ????????????????????????????????????????????????  [450]
'fulvus-WA4'         ????????????????????????????????????????????????  [450]
'sutor-WA'           ....C...........C..C.....A..A.....G...........C..G  [450]
'Heleioporus-1'      ............A...........A......................R  [450]
'Heleioporus-2'      .G.............C..C..G..A.................T......  [450]
'Heleioporus-3'      ????????????????????????????????????????????????  [450]
Notaden              ..........................T..A..A..C..A.....T.....  [450]
Xenopus              ......C........T.....A..T.....A..A.....A..T..G..C.  [450]
Litoria              ...T........T..T.....G.....A.....T..C..A..T.....C.  [450]
'aquilonius-NT'      ????????????????????????????????????????????????  [450]

[                              460        470        480        490        500]
[                              .          .          .          .          .]

'sudelli-SA1'        GCCGGCTTTGTACACTGATTCCCACTCTTCACTGGTTACACTCTTCATAA  [500]
'sudelli-SA2'        ??????????????????????.....................................  [500]
'sudelli-QLD1'       ..................................C.............  [500]
'aquilonius-WA1'     ...........G.............T........C.....C........  [500]
'aquilonius-WA2'     ...........C.............T.....C..C.....C..C..C..  [500]
'centralis-SA2'      ..................................C.............  [500]
'kunapalari-WA1'     ..............................C..C...........C..  [500]
'kunapalari-WA2'     ????????????????????????????????????????????????  [500]
'albipes-WA1'        ..G........G...........G..T.....C..C.....C..C..CG.  [500]
'albipes-WA2'        ..T........G...........G..T.....C..C.....C..C..CG.  [500]
'wilsmorei-WA1'      ...........G.............T.....C..C.....CT.A.....  [500]
'wilsmorei-WA2'      ...........G.............T.....C..C.....CT.A.....  [500]
'pictus-Vic1'        ...........G.............T.....C..C.....C..C..C.G  [500]
'pictus-Vic2'        .........................T.....C..C.....C..C..C.G  [500]
'pelobatoides-WA1'   ...................T.....T.....C..A.....C..C..C..  [500]
'pelobatoides-WA2'   ??????????????????????????????..C..C.....C..C..C..  [500]
'centralis-WA1'      ..................................................  [500]
'centralis-WA2'      ..................................................  [500]
'centralis-SA1'      ..................................C.............G  [500]
'fulvus-WA1'         ...........C.....G........T.....C..C.....C..C..C..  [500]
'fulvus-WA2'         ...........C.....G........T.....C..C.....C..C..C..  [500]
'fulvus-WA3'         ????????????????????????????????????????????????  [500]
'fulvus-WA4'         ????????????????????????????????????????????????  [500]
'sutor-WA'           ..A.....C..G.............T.....C..C.....C..C..C..  [500]
'Heleioporus-1'      ............................T.....................  [500]
'Heleioporus-2'      ...........C.............T.....Y..C.....C..C..C..  [500]
'Heleioporus-3'      ????????????.............T.....C..C.....C.....CG.  [500]
Notaden              ..........................A..TT.A..C..T.......CG.  [500]
Xenopus              .G..A..G..C..T..........GT.A..T........T..A.....  [500]
Litoria              ....T..A.....................A..T..C..............  [500]
'aquilonius-NT'      ????????????????????????????????????????????????  [500]

[                              510        520        530        540        550]
[                              .          .          .          .          .]

'sudelli-SA1'        CACCTGAACTAAAGCTCACTTCGGGGTGATGTTTACAGGGGTCAACCTTA  [550]
'sudelli-SA2'        ................................C.............A.  [550]
'sudelli-QLD1'       ............................A...................A.  [550]
'aquilonius-WA1'     ............G..C.....................T..A..T.......  [550]
'aquilonius-WA2'     ............G..C...........T........C..A..T.....A.  [550]
'centralis-SA2'      ............G.............A...................A.  [550]
'kunapalari-WA1'     ............................T........C..A..T.......  [550]
'kunapalari-WA2'     ????????????????????????????????????????????????  [550]
'albipes-WA1'        T...........G..C........A..C...........A........A.  [550]
```

```
'albipes-WA2'          T..........G..C........A..C..........A........A.   [550]
'wilsmorei-WA1'        ...........G........T..A..T..A..C..G..A..G.....A.   [550]
'wilsmorei-WA2'        ...........G........T..A..T..A..C..G..A..G.....A.   [550]
'pictus-Vic1'          ..............C........A..C..A..C.....A..T.....G.   [550]
'pictus-Vic2'          ..............C........A..C..A..C.....A..T.....G.   [550]
'pelobatoides-WA1'     T.............C.....T..A..C..A..C.....A..T.......   [550]
'pelobatoides-WA2'     ...........G..C.....T.....T........C..A..T.....A.   [550]
'centralis-WA1'        .........C..G..........A..A..........A..........   [550]
'centralis-WA2'        .........C..G..........A..A..........A........G.   [550]
'centralis-SA1'        ......................A..A...................G.   [550]
'fulvus-WA1'           ...........G..C........A..T.....C..T..A.G...AT.A.   [550]
'fulvus-WA2'           ...........G..C........A..T.....C..T..A..Y..AT.A.   [550]
'fulvus-WA3'           ??????????????????????????????????????????????   [550]
'fulvus-WA4'           ??????????????????????????????????????????????   [550]
'sutor-WA'             .........C..G..C.....T.....C..............A...T.A.   [550]
'Heleioporus-1'        ........................A...................A.   [550]
'Heleioporus-2'        ...........G.............T........C..A..T.....A.   [550]
'Heleioporus-3'        T..........G..C........A..T..........A..A.....A.   [550]
Notaden                ...A...G.A............CTA.T..A...GT...A..T.....A.   [550]
Xenopus                G.A..A...G.....ATC.....T..A..A..A...G.T.....T.....   [550]
Litoria                ..A......G......T......T..A.....A..................   [550]
'aquilonius-NT'        ??????????????????????????????????????????????   [550]

[                              560       ]
[                              .         ]

'sudelli-SA1'          CATTCTTCCCCCAACACT   [568]
'sudelli-SA2'          ..................   [568]
'sudelli-QLD1'         ..................   [568]
'aquilonius-WA1'       ..................   [568]
'aquilonius-WA2'       ..........A..G....   [568]
'centralis-SA2'        ..................   [568]
'kunapalari-WA1'       ..........A.......   [568]
'kunapalari-WA2'       ???????????????...   [568]
'albipes-WA1'          ....T..T..A..G....   [568]
'albipes-WA2'          ....T..T..A..G....   [568]
'wilsmorei-WA1'        ..........A.....T.   [568]
'wilsmorei-WA2'        ..........A.....T.   [568]
'pictus-Vic1'          ....T...........T.   [568]
'pictus-Vic2'          ....T...........T.   [568]
'pelobatoides-WA1'     ....T........G....   [568]
'pelobatoides-WA2'     ..........A..G....   [568]
'centralis-WA1'        ............G....   [568]
'centralis-WA2'        ............G....   [568]
'centralis-SA1'        ..................   [568]
'fulvus-WA1'           ..........A..G..T.   [568]
'fulvus-WA2'           ..........A..G..T.   [568]
'fulvus-WA3'           ???????????????....   [568]
'fulvus-WA4'           ???????????????....   [568]
'sutor-WA'             .T........T..G....   [568]
'Heleioporus-1'        ..................   [568]
'Heleioporus-2'        ..........A..G....   [568]
'Heleioporus-3'        ....T..T..A..G....   [568]
Notaden                .T........A.....T.   [568]
Xenopus                A..C........T.....   [568]
Litoria                G.................   [568]
'aquilonius-NT'        ???????????????????   [568]
;
END;
```

```
BEGIN ASSUMPTIONS;
      USERTYPE a STEPMATRIX = 4
           A C G T
       [A] . 10 1 10
       [C] 10 . 10 1
       [G] 1 10 . 10
       [T] 10 1 10 .
      ;


      USERTYPE b STEPMATRIX = 4
           A C G T
       [A] . 1 0 1
       [C] 1 . 1 0
       [G] 0 1 . 1
       [T] 1 0 1 .
      ;

      USERTYPE c STEPMATRIX = 4
           A C G T
       [A] . 4 1 4
       [C] 4 . 4 1
       [G] 1 4 . 4
       [T] 4 1 4 .
      ;


      USERTYPE d STEPMATRIX = 4
           A C G T
       [A] . 5 1 8
       [C] 8 . 16 1
       [G] 3 16 . 16
       [T] 8 1 16 .
      ;

      OPTIONS  DEFTYPE=unord PolyTcount=MINSTEPS ;

      CHARSET  pos1  =  1-568\3;
      CHARSET  pos3  =  1-567\3;

      TYPESET  ts10  = a: 1-568;
      TYPESET  notv  = b: 1-568;
      TYPESET * ts4  = c: 1-568;
      TYPESET  asym  = d: 1-568;

      TAXSET delete = 'Heleioporus-1' 'Heleioporus-2' 'Heleioporus-3' Xenopus Litoria
         'sudelli-SA2' 'centralis-SA2' 'centralis-WA2'  'pictus-Vic2' 'pelobatoides-WA2'
         'wilsmorei-WA2' 'fulvus-WA2' 'fulvus-WA4' 'albipes-WA2';
      TAXSET outgroups = 'Heleioporus-1' 'Heleioporus-2' 'Heleioporus-3' Xenopus
         Litoria;

END;

BEGIN PAUP;
      SET outroot = monophyl;
      DELETE outgroups;
      OUTGROUP Notaden;
      CONSTRAINT  4nmono = (('sudelli-SA1','centralis-WA1','sudelli-QLD1','centralis-
         SA1','aquilonius-NT','aquilonius-WA1','aquilonius-WA2','kunapalari-
         WA1','kunapalari-WA2'),('fulvus-WA1','fulvus-WA3','albipes-WA1','pictus-
         Vic1','pelobatoides-WA1','wilsmorei-WA1','sutor-WA'));

END;
```

REFERENCES

Abelson, J. 1992. Recognition of tRNA precursors: A role for the intron. Science 255:1390.

Avise, J. C. and R. A. Lansman. 1983. Polymorphism of mitochondrial DNA in populations of higher animals, pp. 165–190. *In* M. Nei and R. K. Koehn (eds.), *Evolution of Genes and Proteins*. Sinauer, Sunderland, Massachusetts.

Allendorf, F.W. 1978. Protein polymorphism and the rate of loss of duplicate gene expression. Nature 272:76-78.

Bachmann, K. and J.P. Bogart. 1975. Comparative cytochemical measurements in the diploid-tetraploid species pair of hylid frogs *Hyla chrysoscelis* and *H. versicolor*. Cytogenet. Cell Genet. 15:186-194.

Bailey, G.S., R.T.M. Poulter and P.A. Stockwell. 1978. Gene duplication in tetraploid fish: Model for gene silencing at unlinked duplicate loci. P.N.A.S. 11:5575-5579.

Baldi, I., E. Mattoccia, E. Bufardeci, S. Fabbri and G.P. Tocchini-Valentini. 1992. Participation of the intron in the reaction catalyzed by the *Xenopus* tRNA splicing endonuclease. Science 255:1404-1408

Barber, N. H. 1970. Hybridization and the evolution of plants. Taxon 19:154-160.

Bargelloni, L., R.A. Ritchie, T. Patarnello, B. Battaglia, D.M. Lambert and A. Meyer. 1994. Molecular evolution at subzero temperatures: Mitochondrial and nuclear phylogenies of fishes from Antarctica (Suborder Notothenioidei) and the evolution of antifreeze glycopeptides. Mol. Biol. Evol. 11:854-863.

Barnes, W. M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from l bacteriophage templates. Proc. Natl. Acad. Sci. USA 91:2216–2220.

Barrio, A. 1980. Una nueva especie de *Ceratophrys* (Anura, Ceratophryidae) del Dominio Chaqueno. Physis, Secc. C 39:21-30.

Batistic, R.F., M. Soma, M.L. Beçak and W. Beçak. 1975. Further studies on polyploid amphibians: A diploid population of *Phyllomedusa burmeisteri*. J. Hered. 66:160-162.

Beçak, M.L. and W. Beçak. 1970. Further studies on polyloid amphibians (Ceratophrydidae). III Meiotic aspects of the interspecific triploid hybrid: *Odontophyrnus cultripes* (2n=22) x *O. americanus* (4n-44). Chromosoma 31:377-385.

Beçak, M.L. and W. Beçak. 1971. Studies on polyloid amphibians: Karyotype evolution and the phylogeny of the genus *Odontophyrnus*. J.Herpetol. 8:337-341.

Beçak, M.L., W. Beçak and M.N. Rabello. 1966. Cytological evidence for constant tetraploidy in the bisexual South American frog *Odontophrynus americanus*. Chromosoma 19:188-193.

Beçak, M.L., W. Beçak and M.N. Rabello. 1967. Further studies on polyploid amphibians. I Mitotic and meiotic aspects. Chromosoma 22:192-201.

Beçak, M.L., W. Beçak and L.D. Vizotto. 1970a. A diploid population of the polyploid amphibian *Odontophrynus americanus* and an artificial intraspecific triploid hybrid. Experientia 26:545-546.

Beçak, M.L., L. Danaro, and W. Beçak. 1970b. Polyploidy and mechanisms of karyotypic diversification in Amphibia. Cytogenetics 9:225-238.

Beçak, W. 1969. Genic action and polymorphism in polyploid species of amphibians. Genetics 61:183-190.

Beçak, W. and G. Goissis. 1971. DNA and RNA content in diploid and tetraploid amphibians. Experientia 27:345-346.

Beçak, W., M.L. Beçak, D.Lavalle, and G. Schreiber. 1967. Further studies on polyploid amphibians (Ceratophrydidae). II DNA content and nuclear volume. Chromosoma 23:14-23.

Beçak, W., M.L. Beçak and F.G. de Langlada. 1968. Artificial triploid hybrids by interspecific mating of *Odontophrynus* (Amphibia, Anura). Experientia 24:1162-1164.

Bennett, A.F. and P. Licht. 1974. Enzymatic correlates of activity metabolism in anuran amphibians. Am. J. Phys. 226:1149-1151.

Berger, S. L. and A. R. Kimmel (eds.) 1987. *Guide to Molecular Cloning Techniques*. Meth. Enzymol. 152:1–812.

Berlocher, S. H. and D. L. Swofford. 1996. Searching for phylogenetic trees under the frequency parsimony criterion: An approximation using generalized parsimony. (unpublished manuscript)

Blair, W.F. 1958. Mating call in the speciation of anuran amphibians. Amer. Natur. 62:27-51.

Blair, W.F. 1964. Isolating mechanisms and interspecies interactions in anuran amphibians. Quart. Rev. Biol. 39:334-344.

Bogart, J.P. 1967. Chromosomes of the South American amphibian family Ceratophidae with a reconsideration of the taxonomic status of *Odontophrynus americanus*. Can. J. Cytogenet. 9:531-542.

Bogart, J.P. 1980. Evolutionary implications of polyploidy in amphibians and reptiles. In: W.H. Lewis (ed) Polyploidy: Biological Relevance. Plenum Press, New York and London. pp. 341-378.

Bogart, J.P. and A.P. Jaslow. 1979. Distribution and call parameters of *Hyla chrysoscelis* and *Hyla versicolor* in Michigan. Royal Ont. Mus. Pub. Life Sci. Cont. No. 117:3-13.

Bogart, J.P. and A.O. Wasserman. 1972. Diploid-polyploid cryptic species pairs: A possible clue to evolution by polyploidization in anuran amphibians. Cytogenetics 11:7-24.

Bogart, J.P. L.E. Licht, M.J. Oldham and S.J. Darbyshire. 1985. Electrophoretic identification of *Ambystoma laterale* and *Ambystoma texanum* as well as their diploid and triploid interspecific hybrids (Amphibia: Caudata) on Pelee Island, Ontario. Can. J. Zool. **63**:340-347.

Borkin, L.J., A.E. Vinogradov, Z.K. Brushko & S.L. Kuzmin. 1995. New records of tetraploid toads (*Bufo viridis* group) from Alma-ata and Taldy-kurgan region, Kazakhstan. Russian J. Herpetology 2(2):174-175.

Bradley, R. D., J. J. Bull, A. D. Johnson and D. M. Hillis. 1993. Origin of a novel allele in a mammalian hybrid zone. Proc. Natl. Acad. Sci. USA 90:8939–8941.

Brenowitz, E.A., G. Rose and R.R. Capranica. 1985. Neural correlates of temperature coupling in the vocal communication system of the gray treefrog (*Hyla versicolor*). Brain Res. 359:364-367.

Briggs, R. 1947. The experimental production and development of triploid frog embryos.

Britten, R.J. and E. Davidson. 1969. Gene regulation for higher cells: A theory. Science 165: 349-357.

Brown, W.M., M. George, Jr. and A.C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. Annals NY Acad. Sci. 361:119-134.

Brown, W.M., E. Prager, A. Wang, and A.C. Wilson. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. 18: 225-239.

Campbell, N.J.H., F.C. Hariss, M.S. Elphinstone and P.R. Baverstock. 1995. Outgroup heteroduplex analysis using temperature gradient gel electrophoresis: High resolution, large scale, screening of DNA variation in the mitochondrial control region. Mol. Ecol. 4:407-418.

Cash, M.N. and J.P. Bogart. 1978. Cytological differentiation of the diploid-tetraploid species pair of North American Treefrogs (Amphibia, Anura, Hylidae). J. Herpetol. 12:555-558.

Cavener, D.R. 1989. Evolution of developmental regulation. Amer. Natur. 3:459-473.

Clayton, J.W. and D.N. Tretiak. 1972. Amine-citrate buffers for pH control in starch gel electrophoresis. J. Fish Res. Bd. Can. **29**:1169-1172.

Cogger, H.G. 1992. *Reptiles and amphibians of Australia.* Reed International Books, Chatswood, NSW.

Coyne, J.A. 1992. Genetics and speciation. Nature 355:511-515.

Coyne, J.A. and H.A. Orr. 1989. Patterns of speciation in *Drosophila.* Evolution 43:362-381.

Cracraft, J. 1983. Species concepts and speciation analysis. Curr. Ornithol. 1:159-187.

Craxton, M. 1991. Linear amplification sequencing: a powerful method for sequencing DNA. *Methods: Companion to Methods in Enzymology.* 3:20-

Czelusniak, J., M. Goodman, D. Hewett-Emmett, M.L. Weiss, P.J. Venta and R.E. Tashian. 1982. Phylogenetic origins and adaptive evolution of avian and mammalian haemoglobin genes. Nature 268:297-300.

Danzmann, R.G. and J.P. Bogart. 1982a. Evidence for a polymorphism in gametic segregation using a malate dehydrogenase locus in the tetraploid treefrog *Hyla versicolor.* Genetics 100:287-306.

Danzmann, R.G. and J.P. Bogart. 1982b. Gene dosage effects on MDH isozyme expression in diploid, triploid, and tetraploid treefrogs of the genus *Hyla.* J. Hered. 73:277-280.

Danzmann, R.G. and J.P. Bogart. 1983. Further evidence for a polymorphism in gametic segregation in the tetraploid treefrog *Hyla versicolor* using a glutamate oxaloacetic transaminase locus. Genetics 103:753-769.

Delseny, M., J.M. McGrath, P.This, A.M. Chevre and C.F. Quiros. 1990. Ribosomal RNA genes in diploid and amphidiploid *Brassica* and related species: organization, polymorphism and evolution. Genome 33:733-744.

Doherty, J.A. and H.C. Gerhardt. 1984. Acoustic communication in hybrid treefrogs: sound production by males and selective phonotaxis by females. J. Comp. Physiol. A 319:319-330.

Donelson *et al.* 1985. Microbiol. Rev. 49:107-125.

Dorit, R. L., H. Akashi and W. Gilbert. 1995. Absence of polymorphism at the *ZFY* locus on the human Y chromosome. Science 268:1183–1185.

Dowling, T.E., G.R. Smith and W.M. Brown. 1989. Reproductive isolation and introgression between *Notropis cornutus* and *Notropis chrysocephalus* (family Cyprinidae): comparison of morphology, allozymes, and mitochondrial DNA. Evolution 43:620-634.

Dowling, T.E., C. Moritz, J.D. Palmer and L.H. Rieseberg. 1996. Nucleic acids III: Analysis of fragments and restriction sites, pp. 249-320. *In*: D.M. Hillis, C. Moritz and B.K. Mable (eds.) *Molecular Systematics*, Second Edition. Sinauer, Sunderland, Massachussetts.

Ehrlich, P. 1961. Has the biological species concept outlived its usefulness? Syst. Zool. 10:167-176.

Engel, W., J. Schmidtke, W. Vogel and U. Wolf. 1973. Genetic polymorphism of lactate dehydrogenase isoenzymes in the carp (*Cyprinus carpio*) apparently due to a "null allele". Biochem. Genet. 8:281-289.

Farris, J.S., A.G. Kluge & M.F. Mickevich. 1982. Phylogenetic analysis, the monothetic group method, and myobatrachid frogs. Sytematic Zoology 31:317-327.

Ferris, S.D. and G.S. Whitt. 1977a. Duplicate gene expression in diploid and tetraploid loaches (Cypriniformes, Cobitidae). Biochem. Genet. 15:1097-1112.

Ferris, S.D. and G.S. Whitt. 1977b. Loss of duplicate gene expression after polyploidisation. Nature 265:258-260.

Fitch, W.M. 1986. The estimate of total nucleotide substitutions from pairwise differences is biased. Philos. Trans. R. Soc. London B 312:317-324.

Frost, D.R. and D.M. Hillis. 1990. Species in concept and practice: Herpetological applications. Herpetologica 46:87-104.

Gerhardt, H.C. 1974. Mating call differences between eastern and western populations of the treefrog *Hyla chrysoscelis*. Copeia 1974:534-536.

Gerhardt, H.C. 1978. Temperature coupling in the vocal communication system of the gray tree frog, *Hyla versicolor*. Science 199:992-994.

Gerhardt, H.C. 1982. Sound pattern recognition in some North American treefrogs (Anura: Hylidae): implications for mate choice. Amer. Zool. 22:581-595.

Gerhardt. H.C. 1991. Female mate choice in treefrogs: static and dynamic acoustic criteria. Anim. Behav. 42:615-635.

Gerhardt, H.C. and J.A. Doherty. 1988. Acoustic communication in the gray treefrog, *Hyla versicolor*: evolutionary and neurobiological implications. J. Comp. Physiol. A 162:261-278.

Gerhardt, H.C., M.B. Ptacek, L. Barnett and K.G. Torke. 1994. Hybridization in the diploid-tetraploid treefrogs *Hyla chrysoscelis* and *Hyla versicolor*. Copeia 1994:51-59.

Ghiselin, M.T. 1966. On psychologism in the logic of taxonomic controversies. Syst. Zool. 15:207-215.

Ghiselin, M.T. 1969. *The Triumph of the Darwinian Method*. University of California Press, Berkeley.

Ghiselin, M.T. 1975. A radical solution to the species problem. Systm. Zool. 23:536-544.

Goldman, N. 1990. Maximum likelihood of phylogenetic trees, with special reference to Poisson process models of DNA substitution and to parsimony analysis. Syst. Zool. 39:345–361.

Goldman, N. and Z. Yang. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11:724-736.

Graybeal, A. 1993. The phylogenetic utility of cytochrome b: lessons from Bufonid frogs. Mol. Phylogenet. Evol. 2:256-269.

Graybeal, A. 1994. Evaluating the phylogenetic utility of genes: a search for genes informative about deep divergences among vertebrates. Syst. Biol. 43:174-193.

Graybeal, A. 1995. Naming species. Syst. Biol. 44:237-250.

Haddad, C.F.B., J.P. Pombal, Jr. and R.F. Batistic. 1994. Natural hybridization between diploid and tetraploid species of leaf-frogs, genus *Phyllomedusa* (Amphibia). J. Herpetology 28:425-430.

Hare, M.P., S.A. Karl and J.C. Avise. 1996. Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. Mol. Biol. Evol. 113:334-345.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160-174.

Hayashi, K. 1991a. PCR–SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. PCR Meth. Applica. 1:34–38.

Hayashi, K. 1991b. PCR–SSCP: A method for detction of mutations. GATA 9:73–79.

Hedges, S.B., J. P. Bogart and L.R. Maxson. 1992. Ancestry of unisexual salamanders. Nature 356:708-710.

Heyer, W.R. and D.S. Liem. 1976. Analysis of the intergeneric relationships of the Australian frog family Myobatrachidae. Smithson. Contrib. Zool. 233:1-29.

Higgins, D. G. and P. M. Sharp. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244.

Higgins, D. G. and P. M. Sharp. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS 5:151–153.

Higgins, D. G., A. J. Bleasby and R. Fuchs. 1992. CLUSTAL V: improved software for multiple sequence alignment. Comput. Appl. Biosci. 8:189–191.

Hillis, D.M. 1988. Systematics of the *Rana pipiens* complex: Puzzle and paradigm. Ann. Rev. Ecol. Syst. 19:39-63.

Hillis, D. M. 1994. Homology in molecular biology, pp. 339–367. *In* B. K. Hall (ed.), *Homology: The Hierarchical Basis of Comparative Biology*. Academic Press, New York.

Hillis, D.M. and J.J. Bull. 1992. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42:182-192.

Hillis, D.M. and S.K. Davis. 1987. Evolution of the 28S ribosomal RNA gene in anurans: regions of variability and their phylogenetic implications. Mol. Biol. Evol. 4:117-125.

Hillis, D.M. and S.K. Davis. 1988. Ribosomal DNA: intraspecific polymorphism, concerted evolution, and phylogeny reconstruction. Systematic Zoology 37:63-66.

Hillis, D.M. and M.T. Dixon. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. Quart. Rev. Biol. 66:411-453.

Hillis, D.M., J.T. Collins and J.P. Bogart. 1986. Distribution of diploid and tetraploid species of gray tree frogs (*Hyla chrysoscelis* and *Hyla versicolor*) in Kansas. Amer. Midl. Natur. 117:214-217.

Hillis, D.M., B.K. Mable, A. Larson, S.K. Davis and E.A. Zimmer. 1996a. Nucleic Acids IV: Sequencing and Cloning, pp. 321-381. *In*: D.M. Hillis, C. Moritz and B.K. Mable (eds.) *Molecular Systematics*, Second Edition. Sinauer, Sunderland, Massachussetts.

Hillis, D.M., B.K. Mable and C. Moritz. 1996b. Applications of molecular systematics, pp. 515-543. *In*: D.M. Hillis, C. Moritz and B.K. Mable (eds.) *Molecular Systematics*, Second Edition. Sinauer, Sunderland, Massachussetts.

Hillis, D.M., C. Moritz, C.A. Porter and R.J. Baker. 1991. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. Science 251:308-310.

Hillis, D.M., C. Moritz and B.K. Mable. 1996c. *Molecular Systematics*, Second Edition. Sinauer, Sunderland, Massachussetts.

Huelsenbeck, J. P. 1995. Performance of phylogenetic methods in simulation. Syst. Biol. 44:17–48.

Hughes, A.L. 1993. Contrasting evolutionary rates in the duplicate chaperonin genes of Mycobacterium tuberculosis and M. leprae. Mol. Biol. Evol. 10:1353-1359.

Hughes, M.K. and A.L. Hughes. 1993. Evolution of duplicate genes in a tetraploid animal, *Xenopus laevis*. Mol. Biol. Evol. 10:1360-1369.

Hull, D.L. 1976. Are species really individuals? Syst. Zool. 25:174-191.

Jaslow, A.P. and J.C. Vogt. 1977. Identification and distribution of *Hyla versicolor* and *Hyla chrysoscelis* in Wisconsin. Herpetologica 33:201-205.

Johnson, C. 1959. Genetic incompatibility in the call races of *Hyla versicolor* Le Conte in Texas. Copeia 1959:327-335.

Johnson, C. 1963. Additional evidence of sterility between call types in the *Hyla versicolor* complex. Copeia 1963:139-143.

Johnson, C. 1966. Species recognition in the *Hyla versicolor* complex. Tex. J. Sci. 18:361-364.

Johnson, G.B. 1974. Enzyme polymorphism and metabolism. Science 184:28-37.

Johnson, G.B. 1976. Genetic polymorphisms and enzyme function, pp. 46-59. *In* F.J. Ayala (ed.), *Molecular Evolution*. Sinauer, Sunderland, Massachusetts.

Jukes, T.H. and C.R. Cantor. 1969. Evolution of protein molecules. pp. 21-132 *in* H.M. Munro, ed. *Mammalian protein metabolism*. Academic Press, New York.

Kambhampati, S. and K. S. Rai. 1991. Temporal variation in the ribosomal DNA nontranscribed spacer of *Aedes albopictus* (Diptera: Culicidae). Genome 34:293–297.

Kamel, S., J.E. Marsden and J.H. Pough. 1985. Diploid and tetraploid gray treefrogs (*Hyla chrysoscelis* and *Hyla versicolor)* have similar metabolic rates. Comp. Biochem. Physiol. 82A:217-220.

Kamper, S.M. and A.F. Barbet. 1992. Surface epitope variation via mosaic gene formation of potential key to long-term survival of *Trypanosoma brucei*. Mol. Biochem. Parasitol. 53:33-44.

Karl, S. A. and J. C. Avise. 1993. PCR-based assays of Mendelian polymorphisms

from anonymous single-copy nuclear DNA: Techniques and applications for population genetics. Mol. Biol. Evol. 10:342–361.

Kawamura, T. 1984. Polyploidy in amphibians. Zool. Sci. 1:1-15.

Kimura, M. 1983. *The neutral theory of molecular evolution.* Cambridge University Press, Cambridge.

Kimura, M. and T. Ota. 1974. On some principles governing molecular evolution. Proc. Natl. Acad. Sci. U.S.A. 71:2848-2852.

Kluge, A.G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. 38:315-328.

Klump, G.M. and H.C. Gerhardt. 1987. Use of non-arbitrary acoustic criteria in mate choice by female gray treefrogs. Nature 326:286-288.

Kobel, H.R. and L. du Pasquier. 1986. Genetics of polyploid *Xenopus*. Trends in Genet. 2:310-315.

Kocher, T. D., W. K. Thomas, A. Meyer, S. V. Edwards, S. Pääbo, F. X. Villablanca and A. C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA 86:6196–6200.

Krishnan, A. 1975. Rapid flow cytofluorometric analysis of mammalian cell cycle by propidium iodide staining. J. Cell Biol. 66:188-193.

Kuramoto, M. 1990. A list of chromosome numbers of anuran amphibians. Bull. Fukuoa University of Education Vol. 39, Part III, 83-127.

Kuramoto, M. and A. Allison 1989. Karyotypes of microhylid frogs of Papua New Guinea and their systematic implications. Herpetologica 45: 250-259.

Kwiatkowski, J., D. Skarecky and F.J. Ayala. 1992. Structure and sequence comparison of the Cu, Zn Sod gene in the Mediterranean fruit-fly Ceratitis capitata: intron insertion/deletion and evolution of the gene. Mol. Phylogenet. Evol. 1:72-82.

Lamb, J. 1911. Description of three new batrachians from southern Queensland. Ann. Qld. Mus. 10:26-28.

Lanave, C., G. Preparata, C. Saccone and G. Serio. 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20:86–93.

Larhammar, D. and C. Risinger. 1994. Molecular genetic aspects of tetraploidy in the common carp, *Cyprinus carpio*. Mol. Phylogenet. Evol. 3:59-68.

Lawton, J.H. 1992. There are not 10 million kinds of population dynamics. OIKOS 63:337-338.

Lecointre, G., H. Philippe, H. L. V. Lê and H. L. Guyader. 1993. Species sampling has a major impact on phylogenetic inference. Mol. Phylogenet. Evol. 2:205–224.

Lee, S. B. and J. W. Taylor. 1992. Phylogeny of five fungus-like protoctistan Phytophthora species, inferred from the internal transcribed spacers of ribosomal DNA. Mol. Biol. Evol. 9:636–653.

Lessa, E.P. 1992. Rapid surveying of DNA sequence variation in natural populations. Mol. Biol. Evol. 9:323-330.

Lessa, E.P. and G. Applebaum. 1993. Screening techniques for detecting allelic variation in DNA sequences. Mol. Ecol. 2:121-129.

Levanon, D., J. Lieman-Hurwitz, N. Dafui, M. Wigderson, L. Sherman, Y. Bernstein, Z. Laver-Rudich, E. Danciger, O.Stein and Y. Groner. 1985.

Architecture and anatomy of the chromosomal locus in human chromosome 21 encoding the Cu/Zn superoxide dismutase. EMBO 4:77-84.

Levin, D.A. 1983. Polyploidy and novelty in flowering plants. Amer. Natur. 122:1-25.

Li, W.-H. 1980. Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. Genetics 95:237-258.

Li, W.-H. 1990. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J. Mol. Evol. 36:96-99.

Li, W-H. , C,-I. Wu, and C-C. Luo. 1984. Nonrandomness of point mutation reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. J. Mol.Evol. 21:58-71.

Li, W.-H., C.-I. Wu, and C.-C. Luo. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol. Biol. Evol. 2:150-174.

Little, M.L., B.L. Monroe and J.E. Wiley. 1989. The distribution of the *Hyla versicolor* complex in the northern appalachian highlands. J. Herpetol. 23:303-305.

Littlejohn, M. J. 1960. Call discrimination by female frogs of the *Hyla versicolor* complex. Copeia 1960: 47-49.

Littlejohn, M. J. 1965. Vocal communication in frogs. Australian natural history 15(2): 52-55.

Littlejohn, M.J. 1981. Reproductive isolation: a critical review. In: W.R. Atchley and D.S. Woodruff (eds.), Evolution and Speciation: Essays in Honour of M.J.D. White. Cambridge University Press, Melbourne.

Lockhart, P. J., M. A. Steel, M. D. Hendy and D. Penny. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11:605–612.

Longino, J.T. 1993. Scientific Naming. National Geographic Research and Exploration 9(1): 80-85.

Mable, B.K. 1989. Genetic investigation of interspecific hybrids between tetraploid and diploid species of treefrogs, genus *Hyla*. M.Sc. Masters Thesis, University of Guelph.

Mable, B.K. and J.P. Bogart. 1991. Call analysis of triploid hybrids resulting from diploid-tetraploid species crosses of hylid tree frogs. Bioacoustics 3:111-119.

Mable, B.K. and J.P. Bogart. 1995. Genetic investigation of interspecific hybrids between tetraploid and diploid species of treefrogs, genus *Hyla*. J. Heredity 86:432-440.

Maddison, W.P. and D.R. Maddison. 1992. MacClade: analysis of phylogeny and character evolution, version 3.1. Sinauer, Sunderland, Mass.

Mahony, M. 1986. Cytogenetic studies on Australian frogs of the family Myobatrachidae. Unpublished PhD. thesis. Macquarie University, Australia. 270 pp.

Mahony, M. and E.S. Robinson. 1980. Polyploidy in the Australian leptodactylid frog genus *Neobatrachus*. Chromosoma 81:199-212.

Mahony, M. and J.D. Roberts. 1986. Two new species of desert burrowing frogs of the genus *Neobatrachus* (Anura: Myobatrachidae) from Western Australia. Records of the Western Australian Museum. 13:155-170.

Main, A.R., M.J. Littlejohn and A.K. Lee. 1959. Ecology of Australian frogs, pp. 396-411. *In*: A. Keast, R.L. Crocker and C.S. Christian (eds.), *Biogeography and Ecology in Australia.* Monographiae Biologicae 8, W. Junk, Hague.

Marsden, J.E., S.J. Schwager and B. May. 1987. Single-locus inheritance in the tetraploid treefrog *Hyla versicolor* with an analysis of expected progeny ratios in tetraploid organisms. Genetics 116:299-311.

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

Matson, T.O. 1990. Erythrocyte size as a taxonomic character in the identification of Ohio *Hyla chrysoscelis* and *H. versicolor*. Herpetologica 46:457-462.

Maxson, L., E. Pepper and R.D. Maxson. 1977. Immunological resolution of a diploid-tetraploid species complex of tree frogs. Science 199:1012-1013.

Maxson, L. R. 1992. Tempo and pattern in anuran speciation and phylogeny: an albumin perspective. p. 41-57. *In*, K. Adler (ed.), *Herpetology: Current research on the Biology of Amphibians and Reptiles*. Proceedings of the First World Congress of Herpetology. Society for the Study of Amphibians and Reptiles, Oxford, Ohio.

May, B., J.E. Wright and M. Stoneking. 1979. Joint segregation of biochemical loci in salmonidae: results from experiments with *Salvelinus* and review of the literature on other species. J. Fish. Res. Bd. Can. 36:1114-1128.

Mayr, E. 1942. Systematics and the origin of species. Columbia University Press, N.Y.

McLister, J.D., Stevens, E.D. & Bogart, J P. 1995. Comparative contractile dynamics of calling and locomotor muscles in three hylid frogs. J Exp. Biol. 198: 1527-1538.

Mecham, J.S. 1965. Genetic relationships and reproductive isolation in southeastern frogs of the genera *Pseudacris* and *Hyla*. Amer. Midl. Natur. 74:269-308.

Moritz, C., T.E. Dowling and W.M. Brown. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu. Rev. Ecol. Syst. 18:269-292.

Moritz, C. and D.M. Hillis. 1996. Molecular systematics: Context and controversies, pp. 1-13. *In*: D.M. Hillis, C. Moritz and B.K. Mable (eds.) *Molecular Systematics*, Second Edition. Sinauer, Sunderland, Massachussetts.

Moritz, C., C.J. Schneider and D.B. Wake. 1992. Evolutionary relationships within the *Ensatina escholtzi* complex confirm the ring species interpretation. Sys. Biol. 41:273-291.

Muller, H.J. 1925. Why polyploidy is rarer in animals than in plants. Amer. Natur. 59:346-353.

Murphy, R.W., J.W. Sites, Jr., D.G. Buth and C.H.Haufler. 1996. Proteins: Isozyme Electrophoresis, pp. 51-120. *In*: D.M. Hillis, C. Moritz and B.K. Mable (eds.) *Molecular Systematics*, Second Edition. Sinauer, Sunderland, Massachussetts.

Murray, V. 1989. Improved double-stranded DNA sequencing using the linear polymerase chain reaction. Nucl. Acids Res. 17:8889.

Muse, S.V. and B.S. Gaut. 1994. A likelihood approach to comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol. Biol. Evol. 11:715-724.

Myers, R. M., T. Maniatus and L. S. Lerman. 1986. Detection and localization of single base changes by denaturing gradient gel electrophoresis. Meth. Enzymol. 155:501–527.

Myers, R. M., V. C. Sheffield and D. R. Cox. 1989. Mutation detection, GC-clamps, and denaturing gradient gel electrophoresis, pp. 71–88. *In*: H. A. Erlich (ed), *PCR Technology: Principles and Applications for DNA Amplification* . Stockton Press, New York.

Nei, M. 1969. Gene duplication and nucleotide substitutions in evolution. Nature 221:40-42

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283-292.

Nei, M. 1987. *Molecular evolutionary genetics*. Columbia Univ. Press, New York.

Nei, M. and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3:418-426.

Nishioka, M. And H. Ueda. 1983. Studies on polyploidy in Japanese Treefrogs. Sci. Rep. Lab. Amphibian Biol., Hiroshima Univ. 6:207-252.

Nixon, K.C. and Q.D. Wheeler. 1990. An amplification of the phylogenetic species concept. Cladistics 6:211-223.

O'Hara, R.J. 1993. Systematic generalization, historical fate, and the species problem. Syst. Biol. 42:231-246.

Ohno, S. 1985. Dispensable genes. Trends in Genetics 1:160-164.

Ohno, S. 1993. Patterns in genome evolution. Current Opinions in Genetics and Development 3: 911-914.

Ohno, S., U. Wolf, and N.B. Atkin. 1968. Evolution from fish to mammals by gene duplication. Hereditas 59:169-187.

Ohta, T. 1992. The nearly neutral theory of molecular evolution. Ann. Rev. Ecol. Syst. 23:263-286.

Orita, M., H. Iwahana, H. Kanazawa, K. Hayashi and T. Sekiya. 1989. Detection of polymorphisms of human DNA by gel electrophoresis and single-strand conformation polymorphisms. Proc. Natl. Acad. Sci. USA 86:2766–2770.

Orr H.A. 1990. "Why polyploidy is rarer in animals than in plants" revisited. Amer. Natur. 6:759-770.

Ota, T. and M. Nei. 1995. Evolution of immunoglobulin V_H pseudogenes in chickens. Mol. Biol. Evol. 12:94-102.

Paigen, K. 1989. Experimental approaches to the study of regulatory evolution. Amer. Natur. 134:440-458.

Palumbi, S. 1996. Nucleic acids I: PCR, pp. 205-247. *In*: D.M. Hillis, C. Moritz and B.K. Mable, eds. *Molecular Systematics*, Second edition. Sinauer Assoc., Sunderland, Massachusetts.

Palumbi, S.R. and C.S. Baker. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. Mol. Biol. Evol. 11:426-435.

Palumbi, S., A. Martin. S. Romano, W.O. McMillan, L. Stice, G. Grabowski. 1991. *The simple fool's guide to PCR*. Dept. of Zool., special publication, Univ. of Hawaii, Honolulu.

Parker, H.R., D.P. Philipp, and G.S. Whitt. 1984. Relative developmental success of interspecific *Lepomis* hybrids as an estimate of gene regulatory divergence between species. J.Exp.Zool. 201:354-360.

Parker, H.W. 1940. Australasian frogs of the family Leptodactylidae. Novit. Zool. 42:1-106.

Paterson, H.E.H. 1985. The recognition concept of species. *In*: E.S. Vrba (ed.). *Species and speciation.* Transvaal Museum Monograph No 4, Transvaal Museum, Pretoria, South Afirca, pp. 21-29.

Phillips, R.B. and K.A. Pleyte. 1991. Nuclear DNA and salmonid phylogenetics. J. Fish Biol. 39(suppl.A):259-275.

Pleyte, K.A., S.D. Duncan and R.B. Phillips. 1992. Evolutionary relationships of the salmonid fish genus *Salvelinus* inferred from DNA sequences of the first internal transcribed spacer (ITS1) of the ribosomal DNA. Mol. Phylogenet. Evol. 1:223-230.

Principaud, E. and G. Spohr. 1991. *Xenopus laevis* c-myc I and II genes: Molecular structure and developmental expression. Nucleic Acids Res. 19:3081-3088.

Ptacek, M., H.C. Gerhardt, and R.D. Sage. 1994. Speciation by polyploidy in treefrogs: multiple origins of the tetraploid, *Hyla versicolor*. Evolution 48:.893-903.

Ralin, D.B. 1968. Ecological and reproductive differentiation in the cryptic species of the *Hyla versicolor* complex. (Hylidae). Southwest. Natur. 13:283-300.

Ralin, D.B. 1977. Evolutionary aspects of mating call variation in a diploid-tetraploid species complex of treefrogs (Anura). Evolution 31:721-736.

Ralin, D.B. 1978. "Resolution" of diploid-tetraploid treefrogs. Science 202:335-336.

Ralin, D.B. 1981. Ecophysiological adaptation in a diploid-tetraploid complex of treefrogs (Hylidae). Biochem. Physiol. 68A:175-179.

Ralin, D.B. and J.S. Rogers. 1979. A morphological analysis of a North American diploid -tetraploid complex of treefrogs. J. Herp. 13:261-269.

Ralin, D.B. and R.K. Selander. 1979. Evolutionary genetics of diploid-tetraploid species of treefrogs of the genus *Hyla*. Evolution 33:595-608.

Ralin, D.B., M.A. Romano and C.W. Kilpatrick. 1983. The tetraploid treefrog *Hyla versicolor:* evidence for a single origin form the diploid *H. chrysoscelis*. Herpetologica 39:212-225.

Richards, C.M. and G.W. Nace. 1977. The occurrence of diploid ova in *Rana pipiens*. J. Herpetol. 68: 307-312.

Ritland, C. E., K. Ritland and N. A. Straus. 1993. Variation in the ribosomal internal transcribed spacers (ITS1 and ITS2) among eight taxa of the *Mimulus guttatus* species complex. Mol. Biol. Evol. 10:1273–1288.

Roberts, J.D. 1978. Redefinition of the Australian leptodactylid frog *Neobatrachus pictus* Peters. Transactions of the Royal Society of South Australia 102:97-105.

Roberts, J.D. and C.M. Majors. 1993. Range extensions, range definitions and call structures for frogs from Western Australia. Rec. West. Aust. Mus. 16:315-322.

Roberts, J.D., M. Mahoney, P. Kendrick and C.M. Majors. 1991. A new species of burrowing frog, *Neobatrachus* (Anura: Myobatrachidae), from the eastern wheatbelt of Western Australia. Records of the Western Australian Museum. 15:23-32.

Roberts, J.D., Maxson, L.R. & N. Plummer. 1996. (in press) Recent speciation in diploid-tetraploid burrowing frogs of the genus Neobatrachus (Anura: Myobatrachidae): immunological evidence. Chapter in, Gondwanan Heritage: Proceedings of the Australian Systematic Botany Society meeting, Perth, 1993.

Rodríguez, F., J. L. Oliver, A. Marín and J. R. Medina. 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142:485–501.

Rogers, J.S. 1972. Measures of genetic similarity and genetic distances. Studies in Genet. VII. University of Texas Pub. 7213:145-153.

Romano, M.A., D.B. Ralin, S.I. Guttman and J.H. Skillings. 1987. Parallel electromoph variation in the diploid-tetraploid gray treefrog complex. Amer. Natur. 130:864-878.

Romano, P.R. and J.C. Vaughn. 1985. Diploidization of rRNA gene function may proceed by ordered nucleotide sequence change at the genomic level in a tetraploid frog species. J. Cell Biol. 101:452a.

Romano, P.R. and J.C. Vaughn. 1986. Restriction endonuclease mapping of ribosomal RNA genes: sequence divergence and the origin of the tetraploid treefrog Hyla versicolor. Biochem. Genet. 24:329-347.

Rose, G.J., E.A. Brenowitz and R.R. Capranica. 1985. Species specificity and temperature dependency of temporal processing by the auditory midbrain of two species of treefrogs. J.Comp.Physiol. A 157:763-769.

Roth, C., F. Bringaud, R.E. Layden, T. Baltz and H. Eisen. 1989. Active late-appearing variable surface antigen genes in *Trypanosoma equiperdum* are constructed entirely from pseudogenes. Proc. Natl. Acad. Sci. U.S.A. 86:9375-9377.

Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horu, H. A. Erlich and N. Arnheim. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350–1354.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491.

Salemma, H. 1984. Polyploidy in the evolution of the glacial relict *Pontoporeia* spp. (Amphipoda, Crustacea). Heriditas 100:53-60.

Sambrook, J., E.F. Fritxch and T. Maniatus. 1989. Molecular Cloning: A Laboratory Manual. 2nd Edition. Cold Spring Harbor Laboratory Press, New York.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Nat. Acad. Sci. U.S.A. 74:5463-5467.

Schlötterer, C., M. T. Hauser, A. von Haeseler and D. Tautz. 1994. Comparative evolutionary analysis of rDNA ITS regions in *Drosophila*. Mol. Biol. Evol. 11:513–522.

Schuldiner, A.R., F. de Pablo, C.A. Moore and J. Roth. 1991. Two nonallelic insulin genes in *Xenopus laevis* are expressed differentially during neurulation in prepancreatic embryos. Proc. Natl. Acad. Sci. 88:7679-7683.

Schultz, R.J. 1980. Role of polyploidy in the evolution of fishes, pp. 313-340. *In*: W.H. Lewis (ed.), *Polyploidy, Biological Relevance*. Plenum Press, New York

Selander, R.K., M.H. Smith, S.Y. Young, W.E. Johnson, and J.B. Gentry. 1971. IV Biochemical polymorphism and systematics in the genus *Peromyscus*. I.

Variation in the old-field mouse (*Peromyscus polionotus*). In: Studies in Genetics VI. Univ. Texas Publ. **7103**: 49-90.

Shaw, C.R. and Prasad, R. 1970. Starch gel electrophoresis of enzymes - a compilation of recipes. Biochem. Genet. **4**:297-320.

Shyue, S.-K., L.Li, B.H.-J. Chang and W.-H. Li. 1994. Intronic gene conversion in the evolution of X-linked color vision genes. Mol. Biol. Evol. 11:548-551.

Simpson, G.G. 1961. *Principles of Animal Taxonomy.* Columbia University Press, New York.

Slade, R. W., C.M. Moritz, A. Heideman and P. T. Hale. 1993. Rapid assessment of single copy nuclear DNA variation in diverse species. Mol. Ecol. 2:359-373.

Slade, R.W., C.M. Moritz and A. Heideman. 1994. Multiple nuclear-gene phylogenies: application to pinnipeds and comparison with a mitochondrial DNA gene phylogeny. Mol. Biol. Evol. 11:341-356.

Sollner-Webb, B. 1993. Novel intron-encoded small nucleolar RNAs. Cell 75:403-405.

Soltis, D.E. and P.S. Soltis. 1989. Genetic consequences of autopolyploidy in *Tomiea* (Saxifragaceae). Evolution 43:586-594.

Soltis, D.E., P.S. Soltis and B.D. Ness. 1989. Chloroplast-DNA variation and multiple origins of autopolyploidy in *Heuchera micrantha* (Saxifragaceae). Evolution 43:650-656.

Stoneking, M., B. May and J.E. Wright. 1981. Loss of duplicate gene expression in salmonids: evidence for a null allele polymorphism at the duplicate aspartate aminotransferase loci in brook trout (*Salvelinus fontinalis*). Biochem. Genet. 19:1063-1077.

Sullivan, D.T., W.T. Starmer, S.W. Cutiss, M. Menotti-Raymond and J. Yum. 1994. Unusual moleular evolution of an *Adh* pseudogene in *Drosophila*. Mol. Biol. Evol. 11:443-457.

Swofford, D.L. 1996. PAUP*, version 4.0. Sinauer Assoc., Sunderland, Mass.

Swofford, D.L., and S.H. Berlocher. 1987. Inferring evolutionary trees under the principle of maximum parsimony. Syst. Zool. 36:293-325.

Swofford, D. L. and R. B. Selander. 1981. BIOSYS-1: A FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Hered. 72:281–283.

Swofford , D.L., G.J. Olsen, P.J. Waddell and D.M. Hillis. 1996. Phylogenetic Inference, pp. 407-514. *In:* D.M. Hillis, C. Moritz and B.K. Mable, eds. *Molecular Systematics*, second edition. Sinauer Assoc., Sunderland, Mass.

Tanaka, T. And M. Nei. 1989. Positive Darwinian selection observed at the variable-region genes of immunoglobulins. Mol. Biol. Evol. 6:447-459.

Theriot, E. 1992. Clusters, species concepts, and morphological evolution of diatoms. Syst. Biol. 41:141-157.

Thomas, W.K. and A.T. Beckenbach. 1989. Variation in salmonid mitochondrial DNA: Evolutionary constraints and mechanisms of substitution. J. Mol. Evol. 29:233-245.

Toivonen, L.A., D.T. Crowe, R.J. Detrick, S.W. Klemann, and J.C. Vaughn. 1983. Ribosomal RNA gene number and sequence divergence in the diploid-tetraploid species pair of North American hylid tree frogs. Biochem. Genet. 21:299-307.

Tyler, M.J. and M. Davies. 1986. *Frogs of the Northern Territory*. Conservation Commission of the Northern Territory, N.T.

Tyler, M.J., M. Davies and A.A. Martin. 1981 New and rediscovered species of frogs from the Derby-Broome area of Western Australia. Records West. Aust. Mus. 9:147-172.

Tyler, M.J., L.A. Smith and R.E. Johnstone. 1994. *Frogs of Western Australia*, Revised Edition. Western Australian Museum, Perth, WA.

Tymowska, J. 1991. Polyploidy and cytogenetic variation in frogs of the genus *Xenopus*., pp. 259-297. *In* D.M. Green and L.K. Sessions (eds.), *Amphibian Cytogenetics and Evolution*. Academic Press, San Diego.

Utter, F.M., H.O. Hodgins and F.W. Allendorf. 1975. Biochemical genetic studies of fishes: potentialities and limitations, pp. 213-238. *In*: P.C. Malins and J.R. Sargent (eds.), *Biochemical and biophysical perspectives in marine biology*, Volume 1. Academic Press, New York.

Valentine, J.E., M.J. Boyle and W.A. Sewell. 1992. Presence of single-stranded DNA in PCR products in slow electrophoretic mobility. BioTechniques 13:222-224.

Vilgalys, R. and B. L. Sun. 1994. Ancient and recent patterns of geographic speciation in the oyster mushroom *Pleurotus* revealed by phylogenetic analysis of ribosomal DNA sequences. Proc. Natl. Acad. Sci. USA 91:4599–4603.

Vogler, A. P. and R. DeSalle. 1994. Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle *Cicindela dorsalis*. Mol. Biol. Evol. 11:393–405.

Vriz, S., M. Taylor and M. Mechali. 1989. Differentiai expression of two *Xenopus* c-myc proto-oncogenes during development. EMBO 8:4091-4097.

Wakely, J. 1994. Substitution-rate variation among sites and the estimation of transition bias. Mol. Biol. Evol. 11:436-442.

Wasserman, A.O. 1970. Polyploidy in the common tree toad *Hyla versicolor* LeConte. Science 167:385-386.

Watt, W.B., K. Donohue and P.A. Carter. 1996. Adaptation at specific loci. VI. Divergence vs. parallelism of polypmorhic allozymes in molecular function and fitness-component effects among *Colias* species (Lepidoptera, Pieridae). Mol. Biol. Evol. 13:699-709.

Wendel, J.F., A. Schnabel and T. Seelanan. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc. Nat. Acad. Sci. USA 92:280-284.

Werth, C.R. and M.D. Windham. 1991. A model for divergent, allopatric speciation of polyloid pteridophytes resulting from silencing of duplicate-gene expression.

Wiley, E.O. 1978. The evolutionary species concept reconsidered. Syst. Zool. 27:17-26.

Wiley, E.O. 1980. Is the evolutionary species concept fiction?—A consideration of classes, individuals and historical entities. Syst. Zool. 29:76-80.

Wiley, J.E. 1982. Chromosome banding patterns of treefrogs (Hylidae) of the eastern United States. Herpetologica 38:507-520.

Wiley, J.E. 1983. Chromosome polymorphism in *Hyla chrysoscelis*. Copeia 1983:275-279.

Wiley, J.E., M.L. Little, M.A. Romano, D.A. Blount and G.R. Cline. 1989. Polymorphism in the location of the 18s and 28s rRNA genes on the chromosomes of the diploid tetraploid treefrogs *Hyla chrysoscelis* and *Hyla versicolor*. Chromosoma 97:481-487.

Wilson, A. C., R. L. Cann, S. M. Carr, M. George, Jr., U. B. Gyllensten, K. Helm-Bychowski, R. C. Higuchi, S. R. Palumbi, E. M. Prager, R. D. Sage and M. Stoneking. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linnean Soc. 26:375–400.

Withers, P.C. 1993.. Metabolic depression during aestivation in the Australian frogs, *Neobatrachus* and *Cyclorana*. Aust J Zoology 41:467-473.

Withers, P.C. 1995. Cocoon formation and structure in the aestivating Australian desert frogs, *Neobatrachus* and*Cyclorana*. Aust J Zoology 43:429-41.

Woodruff, D.S. 1989. Genetic anomalies associated with *Cerion* hybrid zones: the origin and maintenance of new electromorphic variants called hybrizymes. Biol. J. Linn. Soc. 36:281-294.

Yang, Z. 1994. Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39:105-111.

Yang, Z., N. Goldman, and A. Friday. 1994. Comparison of models of nucleotide substitution used in maximum likelihood phylogenetic estimation. Mol. Biol. Evol. 11:316-324.

Zevering, C. E., C. Moritz, A. Heideman and R. Sturm. 1991. Parallel origin of duplications and the formation of pseudogenes in mitochondrial DNA from parthenogenetic lizards (*Heteronotia binoei*: Gekkonidae). J. Mol. Evol. 33:431–441.

Zhang, H.-B. and J. Dvorak. 1991. The genome origin of tetraploid species of *Leymus* (Poaceae:Triticeae) inferred from variation in repeated nucleotide sequences. Amer. J. Bot. 78:871-884.

Zorn, A.M. and P.A. Krieg. 1991. PCR analysis of alternative splicing pathways: Identification of artifacts generated by heteroduplex formation. BioTech. 11:181-183.

Zhu, D. B.G.M. Jamieson, A. Hugall and C. Moritz. 1994. Sequence evolution and phylogenetic signal in control-region and cytochrome b sequences of rainbow fishes (Melanotaeniidae). Mol. Biol. Evol. 11:672-683.

Zweifel, R.G. 1970. Distribution and mating call of the treefrog, *Hyla chrysoscelis*, at the Northeastern edge of its range. Chesapeake Sci. 11:94-97.

Vita

Barbara Kim Mable was born in St. Catharines, Ontario, Canada on 31 December, 1963, the daughter of James B. and Laura Mable. After graduating from Beamsville District Secondary School, Beamsville, Ontario, in May 1982, she entered the University of Guelph in Guelph, Ontario. She received her Bachelor of Science degree from the University of Guelph in May 1986. She received her Masters of Science degree from the University of Guelph under the supervision of James P. Bogart in May 1989. She was then employed as a research associate at the Agriculture Canada research station in Vineland, Ontario until July 1991. In August 1991 she entered the Graduate School of the University of Texas at Austin. Her publications include papers on unisexual salamanders, genetic and behavioral characteristics of polyploid hybrid frogs, and inheritance of resistance to pesticides in orchard pests. She is also the coeditor (with David M. Hillis and Craig Moritz) of a book, Molecular Systematics, Second Edition. Her awards include an Ontario Graduate Scholarship, an NSERC Postgraduate Scholarship, and an Izaak Walton Killam Memorial Postdoctoral Fellowship tenurable at the University of British Columbia in Vancouver, B.C.

Current Address: Department of Zoology, University of British Columbia, Vancouver, B.C., CANADA V6T 1Z4

This dissertation was typed by the author.