From Molecules to Organisms: Research Applications of Modern Genetic Tools for Turtle Biology and Conservation

Suzanne E. McGaugh[1], Erika A. Alacs[2], Scott V. Edwards[3], Chris R. Feldman[4], Arthur Georges[2], Jack W. Sites, Jr.[5], and Nicole Valenzuela[1]

[1]*Department of Ecology, Evolution, and Organismal Biology, Iowa State University,*
253 Bessey Hall, Ames, Iowa 50011 USA [smcgaugh@iastate.edu, nvalenzu@iastate.edu];
[2]*Institute for Applied Ecology, University of Canberra, ACT 2601 Australia*
[alacs@aerg.canberra.edu.au, georges@aerg.canberra.edu.au];
[3]*Department of Organismic and Evolutionary Biology, Harvard University,*
Cambridge, Massachusetts 02138 USA [sedwards@fas.harvard.edu];
[4]*Department of Biology, Utah State University, Logan, Utah 84322-5305 USA [elgaria@biology.usu.edu]*
[5]*Integrative Biology, Brigham Young University, Provo, Utah 84602 USA [jack_sites@byu.edu]*

Abstract. – **Molecular methods are a powerful complement to traditional field practices in illuminating the evolution and ecology of turtles. We illustrate how standard approaches such as DNA sequencing and microsatellites have, and will continue, to shed light on numerous aspects of turtle biology. We also forecast the impact of selected technologies such as amplified fragment length polymorphisms (AFLPs), small interspersed nuclear elements (SINEs), single nucleotide polymorphisms (SNPs), bacterial artificial chromosome libraries (BACs), and gene expression techniques. These tools continue to help clarify the demography, population genetics, phylogeography, and phylogenetics of turtles, and hold great potential to elucidate developmental and life history questions in this group. This additional insight, allowed by molecular methods, may ultimately aid in the preservation of turtles by honing conservation and management efforts.**

Key Words. – turtles, small interspersed nuclear elements (SINEs), single nucleotide polymorphisms (SNPs), bacterial artificial chromosome libraries (BACs), gene expression, microsatellites, mitochondrial DNA, amplified length polymorphisms (AFLPs)

Molecular genetic techniques have allowed invaluable insight while complementing traditional field and morphological studies fundamental to ecological and evolutionary questions. Especially welcomed by the turtle community are non-invasive methods that have been a great tool in elucidating demographics (Pearse et al., 2001), mating systems (Pearse et al., 2002), and **phylogenetic** and **phylogeographic** (see glossary for highlighted words) relationships (Spinks et al., 2004; Spinks and Shaffer, 2005; Krenz et al., 2005; Parham et al., 2006b) in these long-lived, wide-ranging, and often highly endangered taxa. In addition to the well-established and widely used genetic methods, emerging techniques will allow studies of genome-wide variation and gene expression, thereby accessing some important questions in turtle biology. Implementing such technologies has the potential to revolutionize our ability to address ecological and evolutionary questions in turtles, including adaptation, longevity, and sex determining mechanisms, and this information will ultimately be useful in conservation efforts.

This review is intended to highlight the capabilities and limitations of traditional and emerging molecular techniques while emphasizing their utility in studies of conservation, evolution, and ecology of turtles. We show how standard approaches such as DNA sequencing and **microsatellite** analysis have, and will continue, to shed light

on numerous aspects of turtle biology (see Fig. 1), and we also forecast the impact of a selected few new techniques such as **bacterial artificial chromosome** libraries and **microarrays**.

MOLECULAR MARKERS

Mitochondrial Genes and Genomes

Mitochondria are small organelles found in the cytoplasm of eukaryotic cells that possess their own **genomes** that encode products crucial to cellular adenosine triphosphate (ATP) production. The typical vertebrate mitochondrial (mt) genome is a circular, haploid genome (ca. 16,500 base pairs) that contains 37 **genes** (Boore, 1999). Because the mt genome is usually transmitted maternally, and generally lacks recombination, it is inherited as a single **locus** (Avise, 2004). These features, along with a relatively high mutation rate, make sequences from the **mtDNA** locus ideal for many kinds of evolutionary studies (Fig. 1).

Bowen et al. (1989) and Lamb et al. (1989) were the first workers to apply mtDNA data to chelonian questions, using variation in mtDNA to assess phylogeographic structure in *Chelonia mydas* and *Gopherus agassizii*, respectively. The first complete mt
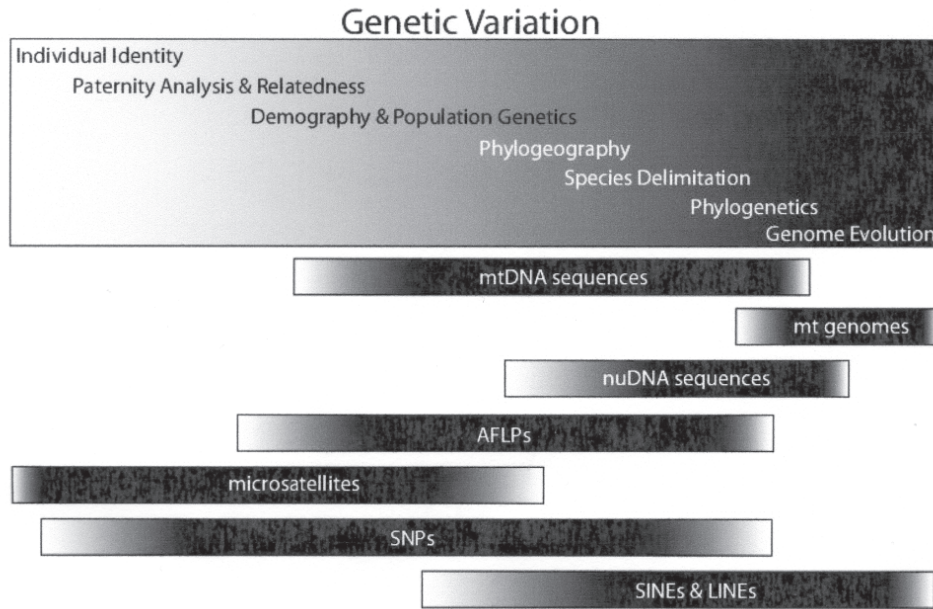
Genetic Variation

Figure 1. Diagram representing the continuum of genetic variation exhibited in biological systems, from the smallest amount of molecular differences (light) distinguishing conspecific individuals to the greatest amount of genetic divergence (dark) seen between phylogenetically distant taxa. Particular fields of inquiry within ecology and evolutionary biology typically deal with only a subset of this gradient of genetic variation and therefore only certain markers will be appropriate for such fields. The text elaborates on technical and logistical aspects of each tool's utility within this context. Technological advances may enable each class of molecular markers to span beyond the boundaries illustrated here, which show where markers are well-suited (dark) or of limited utility (light).

genome sequenced from a turtle (*Pelomedusa subrufa*) was used to assess the phylogenetic position of turtles relative to other amniotes (Zardoya and Meyer, 1998), while the first study to use mt genome data exclusively in turtles examined the phylogenetic relationships of a small group of Old World tortoises including *Testudo*, *Indotestudo*, and *Malacochersus* (Parham et al., 2006b).

Applications of mtDNA

Demography and Population Genetics. — Mitochondrial DNA has been widely used to study processes that determine the geographic distribution of genetic diversity within and among populations. Early comparisons of population genetic structure in mtDNA and nuclear markers performed in *Chelonia mydas* were landmark studies in demonstrating how sex-biased gene flow in turtles could be inferred from such data (Karl et al., 1992, FitzSimmons et al., 1997a,b). Beyond elucidating the current spatial distribution of genetic variation, mtDNA has been used in other vertebrate systems to examine change in genetic diversity and population structure through time. Because of its high copy number, mtDNA remains the most probable source of genetic population signature from ancient specimens.

Phylogeography. — Since mtDNA is haploid, maternally inherited, and possesses a rapid mutation rate, it should track recent population splitting events with higher fidelity than a single nuclear marker under many biologically plausible scenarios (Moore, 1995; Hickerson and Cunningham, 2005). Thus, mtDNA gene regions have

been the most widely used molecular markers to reconstruct population histories and assess phylogeographic structure in turtle species (Fig. 2; e.g., Starkey et al., 2003; Spinks and Shaffer, 2005).

Species Identification and Forensics. — Because multiple copies of the mitochondrion exist in each cell, mtDNA analysis can be particularly useful in identifying the taxonomic or geographic origin of otherwise unidentifiable or poor quality samples (e.g., cooked meat, egg shells, carapace, blood smears, feces). For example, Hsieh et al. (2006) sequenced sections of cytochrome *b* to identify *Kachuga tecta* from poorly stored shells, helping the Council of Agriculture in Taiwan positively document violations of CITES regulations. In another case, Roman and Bowen (2000) used mtDNA to assess whether turtle meat in southeastern U.S. markets was harvested from legitimate sources (e.g., unprotected species). This study showed that even alligator meat was being sold as turtle and led the authors to coin the name "mock turtle syndrome" (Roman and Bowen, 2000).

Mitochondrial DNA can be used in conjunction with other datasets, including either morphological or nuclear molecular markers, to identify hybrid individuals. This approach has recently shown that numerous specimens purported to represent rare and endangered turtle species were actually hybrid individuals from the pet trade (Parham et al., 2001).

Phylogeny. — As mentioned above, mtDNA is particularly amenable to genealogical reconstruction and several features of mt genomes suggest that entire mt genomes are especially well suited for chelonian phylogenetics. First, be-
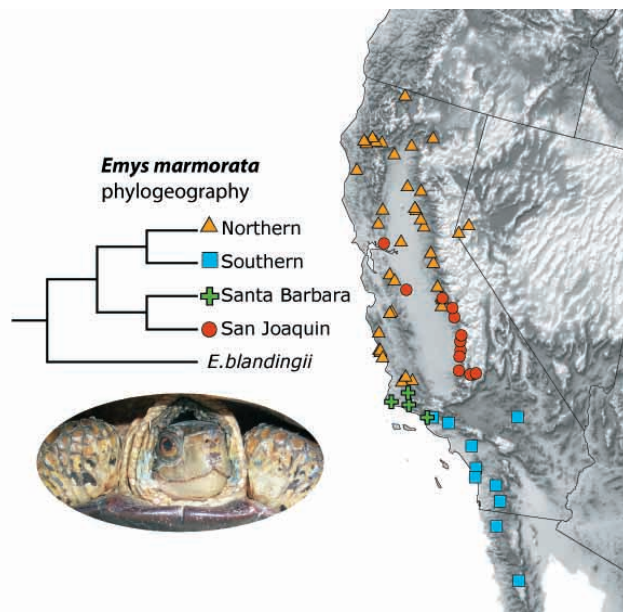
Figure 2. Phylogeography of *Emys* (or *Actinemys*) *marmorata* in western North America (Spinks and Shaffer, 2005). Employing both mtDNA and nuDNA markers and integrating phylogenetic and population genetic analyses in *marmorata* highlights the value of phylogeographies in assessing the evolutionary and biogeographic history of turtle taxa. The phylogeographic approach is also useful in revealing spatial patterns of genetic diversity and setting management priorities. Relationships between the four major mtDNA lineages their geographic distribution follow Spinks and Shaffer (2005). Photo by James Parham.

cause the order and content of genes is highly conserved within mitochondria (Boore, 1999), with no **introns** and spacer DNA, these genomes are easy to align and provide thousands of homologous characters for phylogenetic analyses (Fig. 3). Such large datasets are often necessary to resolve uncertain or incorrect relationships recovered from smaller DNA fragment data (Cummings et al., 1995; Zardoya and Meyer, 1996). For example, small mtDNA and nuclear (nu) DNA fragments initially suggested that turtles are nested within the Diapsida, rather than sister to all other reptiles (Hedges, 1994; Platz and Conlon, 1997). However, the placement of turtles within the Diapsida varied and support for any particular hypothesis was weak. Subsequently, complete mt genome data firmly placed turtles sister to archosaurs (Zardoya and Meyer, 1998; Kumazawa and Nishida, 1999), an arrangement later corroborated by multiple nuclear loci (Hedges and Poling, 1999). Second, the mt genome is composed of loci with vastly different rates of evolution, from hypervariable third positions in cytochrome *b*, to nearly immutable stems in 16S r**RNA**. Thus, mitogenomic datasets should provide resolution at various phylogenetic levels. Third, mt genome features and gene rearrangements have been shown to be valuable phylogenetic characters. Because gene rearrangements are typically rare, and generally considered irreversible, such characters are assumed to be virtually **homoplasy** free (Boore et al., 1995; Boore, 1999).

Genome Evolution. — Complete mt genome sequences may provide a better understanding of how genomes evolve. Because mt genomes are so small, aspects of evolution that

are difficult to study in the nuclear genome may be tractable in the mt genome (Boore, 1999). A number of obvious questions include whether certain gene rearrangements or **duplications** occur more commonly than others, whether rates of rearrangements or duplications correlate with rates of sequence evolution, and whether novel mitochondrial features correspond to particular physiological or life history attributes (Boore, 1999). For instance, through sequencing the mt genome of *Platysternon* a partial mt genome duplication was inferred (Parham et al., 2006a). Further, the hypothesized loss of supernumerary genes excluded the duplicated control region, a characteristic that is relatively unique in metazoans and sets *Platysternon* apart from most other extant species of turtles (Parham et al., 2006a).

Marker Development. — Comparing levels of variation across entire mt genomes for a few focal taxa allows researchers to identify the most appropriate mtDNA markers for their research. Additionally, obtaining mt genome data for the major turtle clades would facilitate the quick development of other mtDNA markers in related taxa.

Data Collection and Analyses. — Total DNA is generally isolated by one of a number of conventional DNA extraction methods (Maniatis et al., 1982), often from tissue samples taken nondestructively in the field. Targeted mtDNA gene regions can then be amplified via **PCR** (Saiki et al., 1988) using a wide array of **primers** known to work in turtles (Engstrom et al., 2007). Long PCR can be employed to amplify large portions of the mt
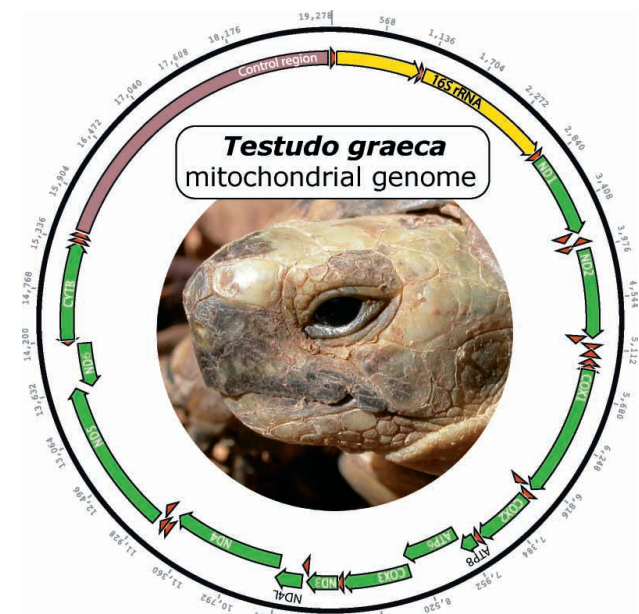
Figure 3. Diagram of the complete mitochondrial genome of *Testudo graeca* (GenBank NC 007692; Parham et al., 2006a). The mt genome of *T. graeca* is representative of turtles and most other vertebrates in overall size, gene content, and gene order. This circular, haploid genome is less than 20 kb with a single control region (brown), two rRNA genes (yellow), thirteen protein coding genes (green) and 22 tRNA genes (red). The mt genome is drawn to scale; arrows indicate the direction of transcription of loci (5' to 3'). The utility of specific loci and the primers used to capture those markers are reported by Engstrom et al. (2007). Photo by James Parham.

genome, or rolling circle amplification (RCA; Dean et al., 2001; Hawkins et al., 2002) can be used to generate entire mt genomes. These amplified products are then used in sequencing reactions that label the four DNA nucleotides (Sanger et al., 1977) and run on an automated machine that reads the labeled nucleotides.

Obtaining sequences from mtDNA gene regions is relatively inexpensive and efficient compared to the cost and time involved in collecting equivalent data from other classes of markers with similar properties and applications. Furthermore, primers that readily amplify many mtDNA regions in turtles are common (Engstrom et al., 2007), and rapid screening of variation in small mtDNA regions for large numbers of individuals is now possible (Avise, 2004; DeSalle and Amato, 2004).

However, collecting entire mt genome data is nontrivial, and the most efficient way to gather these data may be in collaboration with genome centers that have perfected the rapid and efficient acquisition of whole mt genomes (e.g., Joint Genome Institute, Lawrence Berkeley National Laboratory).

Limitations. — Because the haploid mt genome does not recombine, and is uniparentally inherited, all the genes in the mt genome effectively represent a single, linked locus. Thus, analyses based on multiple mt genes or entire mt genomes only represent single-locus estimates of **demography**, population history, or phylogeny. Likewise, inferences made from mtDNA to delimit species or reconstruct population or species histories should be made judiciously. Mitochondrial DNA phylogenies represent the branching history of mitochondria (**gene tree**) and may not track organismal history (**species tree**) flawlessly (reviewed in Avise, 2004), and thus should be corroborated by other evidence (Morando et al., 2004; Avila et al., 2006).

Phylogeographic studies of single species or closely related taxa focus on how evolutionary processes operate in natural populations (Avise, 2000), but the abundance of these studies in the literature belie the difficulties inherent in reconstructing complex demographic histories. The possible influences of past migration, divergence in isolation or with gene flow, and population bottlenecks or expansions, are difficult to disentangle (Knowles, 2004). Furthermore, introgression, incomplete lineage sorting, and natural selection may confound phylogeographic studies (Funk and Omland, 2003). As a consequence, mtDNA phylogeographic analyses have become increasingly sophisticated to accommodate these limitations (Ballard and Whitlock, 2004; Templeton, 2004; Hickerson and Cunningham, 2005).

For deeper phylogenetic questions the rapid rate of mtDNA evolution may lead to homoplasy between deep clades, possibly misleading even mitogenomic estimates of phylogeny (Curole and Kocher, 1999). However, some mtDNA and mt genome data collected in turtles (e.g., Feldman and Parham, 2002; Parham et al., 2006b) do not appear to have suffered from **saturated** data, and newer mixed model methods of analysis (Yang, 1996) may accommodate and correct for at least some mutational history that

can mislead phylogenetic inference (Engstrom et al., 2004; Brandley et al., 2005). Conversely, gene duplications and rearrangements that should be useful for deep level questions in mt genome data (Boore, 1999) may be rare or **autapomorphic**. Parham et al. (2006b) examined both sequence variation and mitogenomic features among major chelonian clades and found that gene rearrangements and duplications were restricted to a single taxon, and thus were phylogenetically uninformative.

Finally, nuclear sequences of mitochondrial origin (**numts**) are relatively common among metazoan taxa (Zhang and Hewitt, 1996; Bensasson et al., 2001) and can seriously mislead any genetic analysis if these nuclear copies of mtDNA are mistaken for authentic mtDNA (Zhang and Hewitt, 1996). Nuclear **pseudogenes** of mtDNA have been reported in turtles (Stuart and Parham, 2004; Spinks and Shaffer, 2007), and may be relatively common.

Future Applications. — We suggest several directions for the future use of mtDNA gene regions and mt genomes in chelonian biology. Most likely, these directions will include a combination of both mitochondrial and nuclear data to address a range of conservation and evolutionary questions. For example, maternally inherited mtDNA and paternally inherited nuDNA markers (Y or W linked loci in taxa with genotypic sex determination) could be used in combination to estimate sex-specific gene flow or other demographic parameters and assess population genetic structure. Already mtDNA and single copy nuDNA sequences have been used in concert to tackle phylogenetic and phylogeographic questions (Krenz et al., 2005; Spinks and Shaffer, 2005; Parham et al., 2006a). Mitogenomic data, in particular, might be combined with nuclear sequences to build a robust chelonian phylogeny that could provide the backbone for any comparative turtle study. Rapidly evolving mtDNA sequences can also be used in combination with **Mendelian** markers, such as microsatellites and single nucleotide polymorphisms (**SNPs**; Morin et al., 2004) for studies of metapopulation structure and conservation genetics (Pearse et al., 2006a). Because mitochondria play an essential role in cellular metabolism, investigations of the molecular evolution of the mt genome may convey metabolic and respiratory adaptations in turtles (e.g., Doiron et al., 2002). Lastly, we anticipate the expanded use of mtDNA, and molecular markers in general, to address broader ecological and evolutionary questions in turtles (Stephens and Wiens, 2003, 2004), and the extension of these findings into conservation biology.

Nuclear Markers: Sequences, Microsatellites, and AFLPs

In contrast to the mitochondrial genome, the nuclear genome contains a huge number of coding and non-coding regions (introns and intergenic spacers) that are subject to different mutation mechanisms and rates (Li, 1997). Thus the nuclear genome offers a virtually unlimited set of potential markers that are informative across the entire range of phylogenetic divergence and can be applied to a wider array

of questions relative to mtDNA data, including studies of adaptive radiation, life histories, hybridization, species delimitation, and phylogenetic inference (including estimates of divergence times [Near et al., 2005]; Fig. 1 summary; Avise 2004; but see Zhang and Hewitt, 2003, for a description of all technologies as well as an in-depth implementation guide).

Nuclear Gene Regions
Applications of Nuclear Gene Data

Phylogeography. — Karl et al. (1992) first used nuclear markers (**restriction digests** of anonymous loci) to estimate global population structure of the marine turtle *Chelonia mydas*, but few subsequent nuclear-sequence based phylogeographic studies have been published on freshwater turtles (FitzSimmons et al., this volume). Phylogeographic studies of Galapagos tortoises *Geochelone* (Caccone et al., 2004) and the western pond turtle *Emys [= Actinemys] marmorata* (Spinks and Shaffer, 2005) have met with limited success because of extremely low variation of nuclear relative to mitochondrial gene regions. This may be a general limitation of most nuclear gene regions accessible by conventional technologies (Zhang and Hewitt, 2003), but newer methods of screening for large numbers of anonymous nuclear loci will likely offer multiple unlinked high resolution markers for future phylogeographic studies (see Jennings and Edwards, 2005, for a recent example in birds).

Species Delimitation. — Allozymes have been used for species delimitation in turtles (e.g., Georges et al., 2002), but the use of nuclear DNA sequence for this purpose is not as prevalent in vertebrates as is the use of mitochondrial markers. Nuclear ribosomal DNA (e.g., internal transcribed spacer [ITS] DNA) has been used for studies of species classifications in algae and nematodes (LaJeunesse, 2001; Chilton, 2004), and primers are available for ITS in turtles although it is not known if this marker would provide an appropriate amount of variability for species delimitation in Testudines (Engstrom et al., 2007).

Phylogeny. — Combining nuclear gene regions can resolve the Testudines phylogenetic history, which has long terminal branches that may result in ambiguous placement of some taxa (Bergsten, 2005). In fact, multiple nuclear genes have been informative about the placement of turtles within Amniota (Hedges and Poling, 1999; Iwabe et al., 2005), single loci have been useful for resolving relationships within Testudines (Fujita et al., 2004), and a combination of nuclear and mtDNA indicated the separation between Platysternidae and Chelydridae (Krenz et al., 2005)

Data Collection and Analyses. — While data are collected using the same protocols as those used for mtDNA gene regions (conventional extraction from field preserved tissue samples; conventional PCR followed by automated sequencing of product, albeit **cloning** of the product is sometimes needed before sequencing), the efficiency of collecting sequence data is usually more difficult because

primers are often borrowed from published sequences developed for other vertebrate groups, and must then be optimized. For example, the nuclear gene *glyceraldehydes-3 phosphate dehydrogenase* (*GADPH*) used by Spinks and Shaffer (2005) was amplified with primers originally developed for birds (Friesen et al., 1997).

Furthermore, duplicated regions can cause problems for phylogenetic and other analyses if one is unknowingly comparing **paralogs** and not **orthologs** (Li, 1997). Therefore, for every nuclear marker developed, a Southern hybridization should be performed to confirm single-copy status as Fujita et al. (2004) did when introducing the nuclear intron R35 as a phylogenetic tool in turtles. Lastly, heterozygosity is more prevalent in nuclear regions and generally requires cloning to resolve.

Limitations. — While nuclear sequences offer many advantages, there are multiple processes operating with greater frequency than in mtDNA and these may confound both data collection and various types of analyses. Additional efforts may be needed to evaluate possible influences of recombination, **codon** bias, duplicated genes, rate variation across characters or taxa, compositional bias, and heterozygosity (Maddison, 1997; Posada and Crandall, 2002; Harris, 2003), and to resolve gene tree – species tree discordance (Edwards and Beerli, 2000; Hudson and Turelli, 2003).

Future Directions. — Data analyses are improving as increasingly refined methods become available for mixed-model analyses (Yang, 1996) of multi-gene data sets for phylogenetic inference (Pagel and Meade, 2004), delimiting species (Sites and Marshall, 2003), and phylogeographic analyses (Templeton, 2004). In addition, steps to improve the alignment process of multigene data sets over a large number of taxa have been taken. At shallower levels of divergence, network methods will become more sophisticated (Cassens et al., 2003, 2005), as will demographic modeling under more biologically plausible scenarios (Hickerson and Cunningham, 2005; Jennings and Edwards, 2005).

Lastly, many conservative vertebrate nuclear gene primers will become applicable for turtle studies, as a result of the National Science Foundation's "Assembling the Tree of Life" (ATOL) initiative (Crandall and Buhay, 2004). Of the 22 projects supported by the ATOL project, five focus exclusively on vertebrates (including birds, archosaurs, amphibians, squamate reptiles, and cypriniform fishes; see: http://ucjeps.berkeley.edu/bryolab/ATOL/?page=projects), and other eukaryote projects are also likely to discover at least some highly conserved regions that can be employed in turtle studies.

Microsatellites

Microsatellite markers, or simple-sequence repeat (SSR) loci, are hyper-variable, iterated 1-6 bp motifs that have been detected in virtually all organismal genomes (Ellegren, 2000; Li et al., 2002). SSR markers constitute a subset of codominant Mendelian loci that are usually assumed to be selectively **neutral** and randomly distributed across **eu-**

chromatic genomes, although these assumptions are not always met (Li et al., 2002). **Alleles** originate by a number of non-conventional mutation mechanisms, which alter the number of repeat units in the alleles segregating at a given locus, and are easily distinguishable based on the length of a PCR product amplified with primers flanking the SSR region. The ease of screening polymorphisms, along with the typically high variability (up to 50 alleles per locus in a population; DeWoody and Avise, 2000), has made SSRs the markers of choice for a wide array of analyses (Avise, 2004; see Bennett, 2000, for in-depth technical review).

Applications of Microsatellites

Paternity Analysis and Relatedness. — Microsatellites are frequently used to estimate individual fitness and some components of breeding structure, in the context of single vs. multiple paternity, and the related phenomenon of sperm storage (both relevant issues in freshwater turtles; see Pearse and Avise, 2001; Pearse et al., 2002, 2006b; for examples).

Demography, Population Genetics, and Phylogeography. — Microsatellites have been utilized to estimate population genetic and phylogeographic structure, especially with regard to the identification of genetic 'breaks' – **Evolutionarily Significant Unit (ESU) or Management Unit (MU)** boundaries – an issue of crucial importance in the design of conservation strategies for endangered species (reviews in Fraser and Bernatchez, 2001; Frankham et al., 2002; Moritz, 2002; DeSalle and Amato, 2004; see Pearse et al., 2006a, for a turtle example). Similarly, microsatellites have been recently used to: (1) evaluate the genetic consequences of recent population bottlenecks (Waldick et al., 2002; Kuo and Janzen, 2004), (2) estimate population sizes and between-deme migration rates (Nichols and Freeman, 2004), (3) estimate natal dispersal (Berry et al., 2004), (4) detect hybridization (see Burns et al., 2003, for an example in turtles), and (5) provide identification in wildlife forensics (Avise, 2004).

In a recent study, Fritz et al. (2005) used microsatellite repeat motifs as primers to amplify ISSRs (inter-simple sequence repeats). By using the repeat motif as a primer, these authors were able to amplify a suite of bands particular to different *Testudo* species. This DNA "fingerprinting" method, in conjunction with mitochondrial DNA, was then used to reject the uniqueness of *Testudo weissingeri* (Fritz et al. 2005).

Data Collection and Analyses. — Microsatellite loci are typically isolated via enrichment probes, which requires less time than previous methods of clone screening (see Fischer and Bachmann, 1998). Once markers are developed, DNA is typically amplified using fluorescently labeled primers, following basic PCR protocols (Sites et al., 1999; Valenzuela, 2000). Amplification reactions are analyzed by electrophoresis, and alleles are scored based on the length of fragments (**electromorphs**). High-throughput genotyping can be achieved by using different fluorescent dyes to label loci with non-overlapping allele sizes in a single automated run or in a single PCR reaction (both terms are referred to as "multiplexing").

Limitations. — Although they are widely utilized, microsatellites have well-characterized limitations as well. From a theoretical perspective, Estoup et al. (2002) reviewed the relationship between SSR mutation models and homoplasy of alleles and showed that basic assumptions about mutational mechanisms are often not met in real data sets. In addition, although SSR loci are generally assumed to be neutral, evidence implicates their influence in clearly non-neutral processes such as genetic disorders (Li et al., 2002), and Vasemägi et al. (2005) found nine microsatellites linked to **Expressed Sequence Tags (EST)** that deviated significantly from neutral expectation. There is also selection against repeat motifs that would produce **frame shifts** in coding regions (e.g., di- and tetra-nucleotide repeats; Metzgar et al., 2000). The nonrandom distribution of SSR loci in the genome further suggests that assumptions of neutral evolution are not always accurate.

Operationally, using primers from related species can affect results by leading to alleles that are shorter, and less variable due to differential amplification (i.e., **ascertainment bias**; Hutter et al., 1998; Amos et al., 2003), or that do not amplify at all (so-called "null" alleles; Zenger et al., 2003). The strength of these effects is directly proportional to the genetic distance from the species for which the loci were originally isolated (Shepherd et al., 2002; Wright et al., 2004).

Future Directions. — Recent studies showed that electromorph (fragment length) data alone tended to underestimate population divergence (Balloux et al., 2000; Fisher et al., 2000). By sequencing microsatellite alleles one can infer mutational processes directly, by checking for consistency in repeat motif for each population sampled (see Engstrom et al., 2007). Electromorph data accompanied by sequence information can paint a more accurate picture of population differentiation (Colson and Goldstein, 1999).

In addition, a variety of approaches have been developed that are appropriate for the evaluation of population genetic structure in non-equilibrium conditions, which are the most likely demographic scenarios for declining species (see reviews by Pearse and Crandall, 2004; Manel et al., 2005). Further, recent empirical studies have shown the advantages of using multiple complementary analytical methods, including equilibrium and non-equilibrium methods, to detect different signals in genetic datasets (e.g., Lemaire et al., 2005; Pearse et al., 2006a). Lastly, combining Mendelian markers and mtDNA sequences can result in powerful inferences about demographic and meta-population structure and histories (FitzSimmons et al., 1997b; Pearse et al., 2006a).

Amplified Fragment Length Polymorphism (AFLP)

The amplified fragment length polymorphism (**AFLP**) method (Vos et al., 1995) is a relatively new technique for generating genome-wide estimates of genetic variation. The AFLP method combines two older molecular techniques (RFLP and RAPD) to quickly and inexpensively produce

numerous, variably sized DNA fragments. Profiles of these anonymous DNA fragments represent multilocus **genotypes** that can be used to answer questions at a wide range of biological scales. For example, these DNA profiles can be used to create distance matrices for phylogenetic reconstruction (Koopman, 2005), estimate population structure (e.g., Mock et al., 2002) or as DNA fingerprints to assess parentage (Mueller and Wolfenbarger, 1999).

The AFLP method has seen little use in animal systems (Bensch and Akesson, 2005) and has not been applied in any chelonian studies, but shows great promise. In the absence of a well-characterized genome, the AFLP method can provide a useful assessment of genome-wide variation in turtles. While there are some limitations inherent to AFLP data, the low cost and ease of use indicate that AFLPs could become valuable markers in a wide range of turtle ecological and evolutionary studies.

Applications of AFLPs

Demography and Population Genetics. — Population genetic studies of animal populations currently emphasize the use of microsatellite or mtDNA sequence data to provide estimates of population structure, gene flow, historical bottlenecks and other population parameters. While rapid rates of evolution in both microsatellites and mtDNA provide investigators with a workable pool of genetic variation to analyze, in most systems, these markers offer a limited view of overall genetic variation in the genome. Furthermore, microsatellite development can be a time consuming and expensive endeavor that generally yields less than 20 usable loci (Zane et al., 2002). AFLPs, on the other hand, can quickly and inexpensively provide a more complete view of genome-wide variation for estimates of population level processes (Bensch and Akesson, 2005). Although AFLP data cannot be scored for more than two alleles at any locus (1/0), or used to detect **heterozygotes**, as they are dominant, rather than co-dominant markers, the shear number of polymorphic AFLP loci can be as powerful as a several variable microsatellite loci in providing highly resolved genotypes (Gerber et al., 2000). Thus, AFLP data may be a useful molecular tool for tackling demographic questions.

Phylogenetics, Phylogeography, and Species Delimitation. — AFLP data can be used to reconstruct the branching history of populations and taxa. Phylogeographic surveys using AFLPs, in particular, could quickly identify cryptic lineages that may represent important management units or cryptic species and could identify regions of hybridization and backcrossing (Miller, 2000). AFLP data can be used directly in the character-based method of maximum parsimony, or compressed into distance matrices to be analyzed with clustering methods for phylogeographic and phylogenetic analysis (Koopman, 2005). However, adequate resolution of many phylogenetic questions may require hundreds or even thousands of AFLP loci (Albertson et al., 1999).

AFLP data could also be used in conjunction with other markers to delimit species when such datasets show concordant geographic boundaries exhibited by distinct populations, similar and separate evolutionary histories, or any other number of empirical situations (reviewed in Sites and Marshall, 2003, 2004).

Adaptive Variation. — The AFLP method may be useful in helping ecologists and evolutionary biologists explore the relationship between genotype and phenotype in chelonian systems. Specifically, researchers may find sets of AFLP loci that are correlated with particular phenotypes of interest. Furthermore, researchers can identify loci that are under selection by comparing the observed distribution of genetic variation at AFLP loci with expectations based on neutral processes (Wilding et al., 2001; Campbell and Bernatchez, 2004).

Data Collection and Analyses. — Following standard DNA extraction/isolation (Maniatis et al., 1982), genomic DNA is cut with two restriction enzymes (Vos et al., 1995) creating hundreds of thousands of DNA fragments. To reduce the number of DNA fragments to a more manageable amount, two rounds of PCR (Saiki et al., 1988) are used to selectively amplify a small portion of the DNA fragments originally cut by the restriction enzymes (Vos et al., 1995). This final pool of amplified DNA fragments can be fluorescently labeled and read on any standard fragment analysis machine (e.g. ABI 3100).

Raw AFLP data consist of a number of DNA fragments of varying lengths. Each fragment is assumed to represent a unique locus in the genome. Individuals that possess a specific fragment have one allele (1), while those that lack the same fragment have the alternative allele (0). Thus, with AFLP data, heterozygotes cannot be distinguished from **homozygotes**, and each locus is assumed to be diallelic in this **dominant** marker system.

Once all the presence/absence data have been collected, any number of analyses can be conducted, though some assumptions regarding Hardy-Weinberg equilibrium may be required to calculate heterozygosity for certain population genetic measures (Bensch and Akesson, 2005).

Limitations. — The chief limitation of AFLP data is that they are not **codominant**. Furthermore, each AFLP locus contains relatively little information (presence or absence of an allele). Thus codominant markers, especially those with high allelic diversity such as microsatellites, actually contain far greater resolving power per locus than AFLPs. To compensate for this deficiency in information content per locus, an AFLP data set must contain many more loci than most other marker systems (Bensch and Akesson, 2005).

Because AFLPs are dominant markers, Hardy-Weinberg equilibrium must be assumed in order to estimate population genetic parameters. Thus AFLPs cannot be used to independently test for violations of Hardy-Weinberg equilibrium in population genetic surveys (Bensch and Akesson, 2005).

Another potential problem of AFLP data is the anonymous nature of loci. Each DNA fragment is assumed to represent a unique locus. Yet, size homoplasy has occurred among smaller DNA fragments (Vekemans et al., 2002) and could seriously confound analyses of genetic diversity

(Vekemans et al., 2002) and phylogenetic reconstruction (Koopman, 2005).

Feasibility. — The quick set-up time involved in collecting AFLP data (often less than a week) and low cost of processing samples make AFLP the most inexpensive and efficient method of assessing genome-wide variation. The AFLP technique can be used without any prior knowledge of a turtle's genome to provide genotypes for a large number samples at a sizeable number of loci. Furthermore, the genetic profiles are highly reproducible across different laboratories. The protocols and equipment required to collect AFLP data should be found in any reasonably equipped molecular genetic laboratory. Moreover, the laboratory procedures have been further streamlined and standardized by a number of commercially manufactured kits.

Regardless of the cost and ease of data collection, AFLP data are not a panacea. Depending on the question and the system, other markers that do not suffer from the same major limitations of AFLP data may be more appropriate (e.g., microsatellites, DNA sequences).

Future Directions. — Future applications in which AFLP are likely to be used include further refinement of our understanding of the genome and its expression into the phenotype. For example, applications include gene mapping such as in **QTL studies** (though crosses are required) and in the discovery of SNPs for chelonian studies (every informative AFLP potentially contains an informative SNP). Another very interesting application of AFLPs is in the study of gene expression. Instead of using whole genomic DNA as the original template for the procedure, **cDNA** generated from expressed mRNA can be used. Using AFLP on cDNA allows researchers to generate global gene expression profiles that may be associated with a particular phenotype, developmental stage, or tissue type of interest (Bachem et al., 1996, 1998).

The AFLP method has not yet been used by turtle biologists, yet the technique can easily be applied to any number of ecological and evolutionary questions. AFLP data should be used judiciously in providing complementary datasets for the estimation of demographic and population genetic parameters (better addressed with microsatellites and SNPs), and in the reconstruction of phylogeographic and phylogenetic histories (better addressed with mtDNA and nuDNA sequence data), but may be ideal in delimiting species (Fig. 1). Further, sex specific AFLPs can be used indicate the heterogametic sex in species with cryptic sex chromosomes (Griffiths and Orr, 1999). Regardless, the low cost and ease of use suggest that the AFLP method shows great potential as a powerful molecular tool for turtle biologists.

MARKERS ON THE HORIZON

Short and Long Interspersed Nuclear Elements (SINEs and LINEs)

An exciting and relatively new set of molecular markers are **SINEs and LINEs** – repetitive elements with no obvious function that are dispersed randomly throughout the ge-

nomes of most eukaryotes (reviewed by Weiner et al., 1986; Shedlock and Okada, 2000; Shedlock et al., 2004).

LINEs (long interspersed nuclear elements) are **transposons** that contain some of the basic machinery of a retrovirus, including a gene for reverse transcriptase (RTase), but do not have the ability to cross-infect cells or individuals. LINE length is variable, but most typically spans a 1-7 kb (Kidwell, 2002). LINEs maintain their integrity within the genome, functioning as self-replicating elements that proliferate randomly by a copy-and-paste process involving an RNA intermediary. Those that lose that function progressively lose their identity through mutation, but are replaced elsewhere within the genome by the continued proliferation of functional elements within the same family. Thus families of functional LINEs reside within the genome, their relationship to each other determined by sequence **homology**. Such families may be longstanding, spanning much or all of the vertebrate radiation, for example. Relatively few LINEs are functional at any one time and the frequency of their propagation is governed by the intranuclear and intragenomic environment (Weiner, 2002).

SINEs (short interspersed nuclear elements) are also transposable but are much shorter elements (70-500 bp), lack a gene for RTase, and rely on a functional corresponding LINE to provide the RTase to support their proliferation (Kajikawa and Okada, 2002). SINEs too form families that are maintained by the balanced processes of gain through replication of functional elements (requiring a functional RTase recognition site) and loss through random mutation. SINEs have attracted particular attention because of their manageable size and because they usually are represented by >10^4 copies per SINE type per vertebrate genome (Kazazian and Moran, 1998; Shedlock et al., 2004).

Applications of SINEs and LINEs to Chelonian Biology

Demography and Population Genetics. — Where a SINE family is still actively proliferating, their utility extends beyond phylogenetics into population biology (Batzer et al., 1996). For example, insertion or lack of insertion of the *Alu* element for 100 loci provided sufficient polymorphism to estimate diversity among and within human populations (Watkins et al., 2003). Sampling of many SINE loci, which are dispersed across the genome, enabled inferences regarding the genetic distance to ancestral states and population subdivision with very little sampling error. In fact, resampling methods regard 50 loci to be sufficient for future studies (Watkins et al., 2003).

Species Delimitation and Phylogenetics. — Other applications where an unambiguous marker is of value may be found in species identification for forensics where the SINEs are fixed at the level of species. In addition, SINEs have been successful at identifying the close relationship between humans and chimps and discovering previously undetected radiations in cichlid species of the east African rift lakes (Shedlock et al., 2004).

For turtle biologists, SINEs are significant because their first application to reptiles is a study of the phylogenetic relationships among geoemydid turtles (Fig. 4; Sasaki et al., 2004). SINEs and LINEs yield phylogenetic information at three levels. The first is at the sequence level, providing information on the phylogeny of the element, and thus the species (or clade) that carries it, in the same way as for any nuclear marker. The second is at the level of the presence or absence of representatives of SINE or LINE families in the entire genome, from which we can infer their origin in a common ancestor to the exclusion of other taxa of interest. The third level involves their use as positional markers, where an individual SINE element at a particular locus can be identified by developing primers for its unique flanking region and scored as present or absent.

It is as positional markers that SINEs and LINEs come into their own as phylogenetic markers. They have a suite of remarkable properties straight out of the notebook of the pioneer of phylogenetic systematics, William Hennig (1966):

(a) They are discrete and recognizable DNA elements that proliferate through the nuclear genome by a copy-and-paste mechanism, rather than the cut-and-paste mechanisms of DNA transposons, so the history of their proliferation can be uncovered using traditional approaches to phylogenetic reconstruction using sequence data. This said, it is the presence or absence of the SINE or LINE at a specific location that is the novel character, and the sequence data internal to the marker is secondary to this.

(b) They insert into the genome essentially at random (though there is a slight bias in favor of AT rich regions) so the probability of homoplasy arising through a second insertion at the same site is remote. This assertion has been supported by an intensive study of the *Alu* SINE of primates (Roy-Engel, 2002). In any case, such an insertion does not overwrite the first and so if a duplicate insertion were to occur it would most likely be easily detected when the element and flanking region are sequenced, unless substantial deterioration has occurred (e.g., Ray et al., 2005).

(c) SINE or LINE insertion at a particular locus is considered irreversible, because flanking regions are created upon insertion and provide a signature of the insertion even in the unlikely event that the element "jumps" out of the previous spot.

(d) Absence of a SINE or LINE is accompanied by a robust positive control, so that there are three possibilities – amplification product contains a SINE or LINE, amplification product does not contain the SINE or LINE, no amplification because of mutation at the primer site. An absence of a SINE or LINE is an absence, provided there is successful amplification.

(e) The marker has clear homology across taxa and the polarities of the character states are unambiguous (i.e., the absence of the SINE or LINE and flanking regions at a specific location in the genome is unambiguously the ancestral state, and presence is unambiguously the derived state).

Once found, a SINE or LINE inserted at a specific location is a nuclear marker that is essentially free of homoplasy, which can occur only through introgression of a SINE element following interspecific hybridization or through gene-tree/species-tree disparity (Hillis, 1990; Miyamoto, 1999). Phylogenetic characters with these attributes potentially offer a treasure trove for systematic biology (Shedlock and Okada, 2000).

Data Collection and Analyses. — The human genome contains nearly 1.5 million SINEs (Shedlock et al., 2004). This abundance in genomes makes isolating and characterizing new SINEs relatively easy given the large playing field. Main approaches for SINE isolation include screening a genomic library with a probe which is designed for a particular SINE family of interest or sequencing of large chunks of the genome and using this information to predict the presence of SINEs
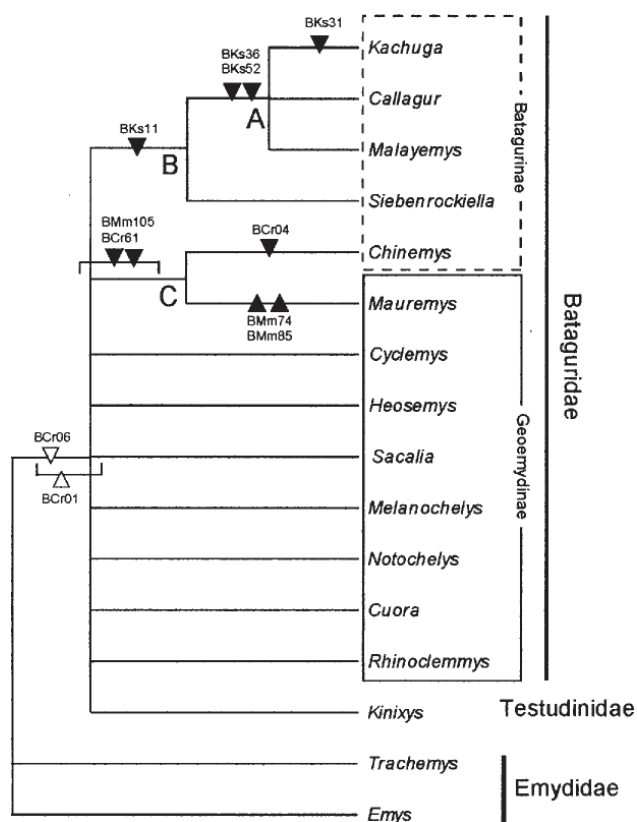
Figure 4. Phylogenetic relationships revealed from among the 16 species of Testudinoidea by the application of the SINE approach (after Sasaki et al., 2004). Arrowheads denote the insertion of tortoise polIII class SINEs. SINE insertions into loci BCr01 and BCr06 reveal a close relationships between Bataguridae (= Geoemydidae) and Testudinidae to the exclusion of the Emydidae. SINE insertions at loci BKs36 and BKs52 indicate that *Kachuga (= Pangshura) smithii, Callagur (= Batagur) borneoensis,* and *Malayemys subtrijuga* form a monophyletic group (clade A) within the *Batagur* complex. An insertion at BKs11 locus suggests monophyly of *Siebenrockiella* with the above three species (clade B). A close relationship between *Chinemys reevesii* (recently changed to *Mauremys reevesii*; Feldman and Parham, 2004; Spinks et al., 2004) and *Mauremys mutica kami* is suggested by SINE insertions at loci BCr61 and BMm105 (clade C).

computationally (Shedlock et al., 2004). Recent approaches also design a primer identical to the conserved polymerase III promoter and use PCR or genomic screening to isolate the new SINE (Shedlock et al., 2004; Borodulina and Kramerov, 2005). Additional information on the characterization of new SINEs and the use of SINEs in systematics is briefly summarized by Shedlock et al. (2004).

Limitations. — Limitations on the utility of SINES derive from the limited life of a particular retroelement as an identifiable and recoverable sequence in the genome, or the limited life of the flanking sequence that enables homology of the positional element to be established. Once inserted, the actual SINE and its flanking regions deteriorate over time through mutation to the point that they are not detectable. This aspect diminishes the utility of the technique beyond 50-150 million years (Shedlock and Okada, 2000; Shedlock et al., 2004).

A second limitation is that unlike sequence data, one cannot expect SINEs to provide information across all nodes of a phylogeny. This was evident in the turtle study (see Fig. 4; Sasaki et al., 2004) where despite considerable effort, solid information was obtained on only four nodes in the cryptodire phylogeny. This situation will improve as options for screening SINEs improve, such as when genomic information on target taxa increases, leading to greater numbers of loci. There may also be novel approaches on the horizon for targeting specific phylogenetic hypotheses at the time of screening for informative SINEs (e.g., screening after selected subtractive hybridization).

A third limitation is that these positional markers, informative as they may be for resolving tree topology, cannot be used for determining branch lengths or dating divergences. For this we must rely upon comparisons of the actual DNA sequences of the SINEs or LINEs or comparisons of sequence data from the flanking regions (Del Pozzo and Guardiola, 1990; Shedlock and Okada, 2000).

Future Directions. — Overall, the future of SINEs for resolving important questions in turtle phylogeny looks bright. Their abundance in the genome provides the opportunity to address the second limitation by identifying a very great number of SINE markers, so that resolution will ultimately be obtained across most or all of the important nodes in the turtle phylogeny, within the 50 million year window. This development will be greatly assisted by improved knowledge of the turtle genome, either through the development of selected BAC libraries (see below) and ultimately, one hopes, a turtle genome project. In the meantime, novel approaches to focusing attention on particular problematic nodes may be possible by combining subtractive hybridization with screening.

Single Nucleotide Polymorphisms (SNPs)

Single nucleotide polymorphisms (SNPs) are rapidly becoming valuable genetic markers because they are the most common source of variation among individuals – a SNP occurs on average every 300-500 bases in the human genome (Zhao et al., 2003). SNPs are generated by point mutations in the genome when one nucleotide is replaced by another (i.e., substitution). This definition is often broadened to include single-base **indels** where an extra base is inserted or deleted during the replication of DNA. In principle, SNPs can have as many as four alternative allelic states (i.e., adenine, guanidine, cytosine, or thiamine), but because of the rarity of the mutations (i.e., 10^{-8} to 10^{-9} mutations per generation per site), tri or tetra-allelic states are virtually non-existent within related taxa. As a result, SNPs are often referred to as bi-allelic markers (Vignal et al., 2002).

SNPs are informative genetic markers for population, conservation, and evolutionary genetic studies when the least abundant allele reaches a frequency of 1% or greater in the population; a threshold that eliminates sites that are variable because of infrequent sequencing errors (Kwok and Gu, 1999; Wakeley et al., 2001; De La Vega et al., 2002). These traits of ubiquitous variation and high utility have recently been harnessed and applied to studies of evolutionary genetics, population genetics, hybridization, and wildlife forensics, and show great promise in chelonian studies (Bensch et al., 2002; Stickney et al., 2002; Belfiore et al., 2003; Aitken et al., 2004; Seddon et al., 2005).

Applications of SNPs

Paternity and Relatedness. — The typically bi-allelic character of SNPs creates a requirement for many more loci to be genotyped for parentage and relatedness studies compared to multi-allelic markers. It is estimated that 60 maximally informative SNPs would be required to provide the same level of paternity exclusion and estimates of relatedness as 14 microsatellite loci with an average allelic diversity of 9.5 (Krawczak, 1999). This number jumps to 100 when SNPs are only 20-30% heterozygous, a level closer to actual diversity (Krawczak, 1999; Glaubitz et al., 2003) However, once the SNP assays are developed, they could potentially produce better quality data and be more cost effective and efficient than microsatellites.

Demography, Population Genetics, and Phylogeography. — Similar to paternity studies, a larger number of SNP versus microsatellite loci are required for estimates of genetic diversity, gene flow, effective population size, and other population parameters (Morin et al., 2004). The extra effort required in isolating loci is offset by the better resolution obtained from SNPs with fewer assumptions compared to microsatellites (Brumfield et al., 2003). Estimates of population parameters such as F_{ST} are likely to be more accurate with SNPs than with microsatellites and AFLP because (i) their mutational mechanisms are relatively well characterized, (ii) they may be less subject to homoplasy, (iii) they potentially have a reduced interlocus sampling variance as a consequence of the large number of loci available for analysis, and (iv) they have less within-population variation which guards against artificially low F_{ST} estimates (Kalinowski, 2002; Nicholson et al., 2002; Brumfield et al., 2003).

Species Delimitation. — In the identification of cryptic species and hybridization, the application of SNPs has been extremely successful using a relatively small numbers of markers. For example, Belfiore et al. (2003) developed three SNPs that were 90% effective in discriminating among four species of Eurasian vole (*Microtus*), a success rate higher than at the nuclear *p*53 locus (DeWoody, 1999). Further, a study of willow warblers used a single SNP to distinguish two subspecies that could not be differentiated using mitochondrial or microsatellite markers (Bensch et al., 2002). SNPs, owing to their codominance, are also effective in the detection of hybridization and introgression (see Saetre et al., 2001, for a more extensive description).

Wildlife Forensics. — SNPs will have immense utility in wildlife forensics, especially when poaching evidence consists of samples that may yield degraded DNA, such as a fragment of carapace or meat from a market (Sarkar and Kashyap, 2003). SNPs can be genotyped from degraded DNA more efficiently than any other nuclear marker due to the small size of the DNA fragment being amplified, and diagnostic SNPs can be found at all taxonomic levels. For example, a SNP assay has been developed in the Chinook salmon which can identify the country of origin of the fish and thereby aid in the regulation of the trade (see Smith et al., 2005).

Evolutionary Genetics. — SNPs are useful in the detection of adaptive variation and in drawing inferences on population demographic history. Signatures of natural selection in populations have been detected with comprehensive SNP maps (Akey et al., 2002; Nielsen, 2005). The abundance of SNPs in the genome and their potential for rapid genotyping makes them ideal markers to map Quantitative Trait Loci (QTL). QTL studies seek to identify the loci responsible for phenotypic traits, and can thereby shed light on how continuous traits are inherited in populations and the influence of evolutionary processes on these traits (Slate et al., 2002; Weinig and Schmitt, 2004; Slate, 2005).

Data Collection and Analyses. — SNP discovery (ascertainment) is successful through BACs and other previously sequenced information (Marth et al., 2001; Saetre et al., 2001; Bensch et al., 2002; Primmer et al., 2002; Belfiore et al., 2003; Brugmans et al., 2003; Nicod and Largiader, 2003). Alternative strategies include the reduced representation shotgun approach (RRS) in which DNA from many individuals are mixed together and subjected to restriction enzyme digestion. The resultant fragments are incorporated into plasmids. This plasmid library is then sequenced, and overlapping sequences are screened for SNP polymorphisms (Altshuler et al., 2000). More recent approaches identify the SNPs causing a polymorphism in the allelic states of an AFLP marker (i.e., present and absent states) and convert these into SNP markers (Nicod and Largiader, 2003; similar to the approach in Fitzpatrick and Shaffer, 2004). SNPs may be discovered in restriction enzyme recognition sites, the primer annealing sites or within an AFLP fragment itself (Bensch et al., 2002; Brugmans et al., 2003). These techniques appear to be promising for the discovery of large numbers of SNP loci in non-model organisms.

A diverse array of methods is available for genotyping SNPs (reviewed extensively by Kwok, 2001). Well established methods such as PCR-RFLP and PCR-SSCP can be used to cost effectively genotype SNPs using standard laboratory equipment (Doi et al., 2004). High throughput can be achieved using newer methods such as primer extension (Li et al., 1999), hybridization (Howell et al., 1999), and invasive cleavage assays (Lyamichev et al., 2000). For rapid genotyping, these techniques can be modified to be used in microarray platforms (Dalma-Weiszhausz and Murphy, 2002; Heller, 2002; Jenkins and Gibson, 2002). In addition, a rapid form of sequencing by DNA synthesis, pyrosequencing, which produces light upon the incorporation of the correct nucleotide, can be advantageous over hybridization applications (Ronaghi, 2001).

Limitations. — Discovery of SNPs without ascertainment bias has been a major limitation to their use. Biases can be reduced by the selection of a large panel of individuals for screening and inclusion of loci that display lower levels of variability (Nielsen, 2000; Schlötterer and Harr, 2002). Statistical analyses to correct for biases in SNP data are also available, such as maximum likelihood models (Kuhner et al. 2000; Nielsen, 2000).

Future Directions. — SNPs are emerging as markers with the potential for wide ranging applications in chelonian biology. For some applications, only a few SNP loci are required, such as for species diagnostics and identifying the geographic origins of individuals; applications which will be particularly useful in wildlife forensics to monitor trade of turtle populations worldwide. Furthermore, an exciting application of SNPs will be to study adaptive evolution in turtles to gain insights on how phenotypic traits are inherited and how they might respond to changes in environmental conditions. However, ascertainment bias remains a major hurdle that must be overcome before SNPs can be reliably used in population and evolutionary studies.

GENOMICS AND GENE EXPRESSION

Comparative Genomics: BACs

The comparative genomics of vertebrates is still in its infancy, with only a single avian genome sequenced thus far and no non-avian reptile genomes. Still, the time is ripe for forays into the comparative genomics of turtles. In particular, the recent availability of a Bacterial Artificial Chromosome (BAC) library from a painted turtle (*Chrysemys picta*) paves the way for a scaling-up of genomic inquiries in turtles and for amassing large-scale information on the structure and organization of turtle genomes. BAC libraries are a means by which very long pieces of DNA (100,000 – 200,000 base pairs) can be isolated (cloned), sequestered from the remainder of the genome, and studied in detail. Although the sequencing of a turtle genome may still be several years away, BAC libraries will provide a useful resource in the interim for studying turtle genomics.

Applications of BACs to Chelonian Biology

Phylogenetics, Marker Development, and Genome Evolution. — Why is cloning long pieces of DNA of interest to the evolution, comparative genomics, molecular evolution, '**evo-devo**' and systematics of turtles and other vertebrates? First, the sheer size of pieces of DNA that can be isolated, and eventually sequenced (Harris and Murphy, 2001), means that a vast number of molecular characters are immediately available for study. Unlike short pieces of DNA amplified by PCR, BAC library inserts provide contiguous stretches of DNA, thereby permitting a more seamless integration of molecular systematics and genome evolution (Pollock et al., 2000; Edwards et al., 2005). The large amounts of contiguous sequence data (contigs) that can be characterized from BAC libraries in a phylogenetic context will yield new insights into phylogenetic analysis of genomic data. For example, Thomas et al. (2003) used contigs of the region containing the cystic fibrosis gene constructed from BAC clones to sequence up to 1.8 Megabases (Mb) of DNA from several mammals and a chicken. Such sequence data yielded abundant retroelements (such as SINEs and LINEs), which in turn serve as cladistic characters in a phylogenetic analysis (Shedlock and Okada, 2000). The alignment of these sequences also revealed numerous non-coding regions that were highly conserved between species, providing a detailed view of regions that could be important for regulation and genome stability. Another recent example of large-scale discovery of phylogenetically important information comes from comparative genomic studies of the coelacanth and bichir, a primitive ray-finned fish and basal tetrapod, respectively (Chiu et al., 2004; Noonan et al., 2004). BAC libraries have proved indispensable for identifying and characterizing multigene families that are important for development. For instance, one can examine conserved and nonconserved regions in these genes in comparison to sequenced organisms to elucidate possible noncoding, conserved function regions. Also, data mining and sequence analysis from BAC libraries can identify expansions or contractions of gene families (Miyake and Amemiya, 2004). In addition, BAC libraries ultimately pave the way for whole-genome sequencing as they can effectively serve as waypoints in the landscape of the genome.

BAC libraries are an efficient means for understanding broad-scale patterns within genomes without actually sequencing entire genomes, or even targeted regions. Features such as the frequency of various families of repetitive elements and retroelements, as well as base compositional and **isochore** structure, can be mined from BAC libraries in several ways. First, one can conduct hybridizations of specific genes or repetitive elements to filters on which the entire BAC library is spotted. In this way, one can obtain an estimate of the frequency of the particular element in the genome of the interrogated species. Second, one can survey the basic structure of a vertebrate genome by conducting a BAC-end sequencing survey, which consists of amassing thousands of sequence reads from the ends of BAC clones,

primed using sequence in the BAC **vector**. Such a survey has been conducted for *Chrysemys picta*, leading to several new insights into turtle genome evolution and phylogeny (Shedlock et al., unpubl. data). An important spin-off for such BAC-end sequencing surveys (or end-sequence surveys of any type of clone) are the release of large numbers of loci for use in phylogeography and molecular systematics (Hare, 2001; Matthee et al., 2001; Jennings and Edwards, 2005). With any given clone-end read, one can immediately design primers for PCR for studying within- or between-species variation (see nuclear gene region section), although the phylogenetic resolution of any given sequence must be determined empirically. The loci typically recovered in a clone-end sequencing survey are noncoding and often 'anonymous' in so far as they do not match any known loci to a significant degree when data bases such as Genbank are interrogated by a BLAST or other similarity search. Such loci are of maximal interest to multilocus phylogeography because they will tend to be more variable than currently available markers.

Gene Function and Expression. — Another key feature of BAC clones is that they contain not only coding regions of genes but all of the noncoding, regulatory regions that affect gene expression. Such regions are frequently found immediately upstream of genes but can often be tens of kilobases away from the coding regions themselves. Thus BAC clones can often capture in a single clone all of the regulatory elements and coding regions of a particular gene or gene family. This makes possible a variety of experiments in developmental biology, such as expression of turtle gene families in developing embryos of model species to examine developmental consequences of gene misexpression (Heintz, 2000; Takahashi et al., 2000; Carvajal et al., 2001; Giraldo and Montoliu, 2001).

Chromosome Mapping. — Individual BAC clones are large enough to be visualized after fluorescent labeling and hybridization to metaphase chromosomes, as in the **FISH** technique (fluorescent in-situ hybridization). By contrast, individual PCR products and many cDNA clones are too short to use in FISH and often do not provide a reliably strong signal of hybridization to a target sequence on the chromosome. Thus BAC clones provide a critical tool for locating genes and gene families on turtle chromosomes. Such studies will provide an important window into turtle chromosome evolution. Thus far the resolution provided by hybridization of whole chicken chromosomes to turtle **karyotypes** has revealed that entire chromosomes found in turtles may have remained intact in birds, as in the example provided by hybridization of a chicken Z chromosome to the entirety of a turtle chromosome 5 and no other chromosomes (Graves and Shetty, 2001). However, even such evidence leaves room for small-scale genomic **translocations** that might not be detected using whole-chromosome hybridizations, particularly of single-copy regions that may not provide an amplified fluorescence signal. BAC clones are ideal for such purposes. Preliminary investigations of chromosome assignments

of several turtle genes are underway, particularly genes in the sex determining pathway and sex-linked genes (N. Valenzuela, unpubl. data; D. Janes, unpubl. data).

Future Directions. — Overall the prospects for robust comparative genomics of turtles are very strong provided that the appropriate resources are made available to the wider community. Ideally all such resources should be available through distribution centers; the *Chrysemys picta* BAC library and additional technical information can be accessed through the Joint Genome Institute (JGI) web site on available BAC libraries: http://evogen.jgi.doe.gov/second_levels/BACs/Our_libraries.html.

Using a 'community genomics' approach and the appropriate genomic resources, large scale projects in animal molecular systematics can be tackled by coordinated efforts of single-PI laboratories as well as genome centers, even for problems that are not of high priority to genome centers (Edwards et al., 2005). In fact, efforts are underway to identify SNPs and sequences that amplify across turtles from the *C. picta* BAC library, an endeavor that will make many more genetic markers available (Thomson, Edwards, and Shaffer, unpubl. data). Such large-scale genomics approaches are a natural complement to typical molecular systematics endeavors utilizing PCR, and will forge an even tighter link between genome evolution and systematics. With judicious use of the available BAC library, and continued attention to production of important genomic resources for turtles, the turtle community could lead the way in these important new directions.

Gene expression: cDNAs, ESTs, RT-PCR, Microarrays, Functional Assays, and RNAi

Firmly linking an organism's genotype to its phenotype is one of the most important, yet, ambitious goals of molecular genetics. Technological advances are now allowing for researchers to dissect at a molecular level fundamental questions, such as how organisms react to different environments and what contributes to morphological diversity among species. A useful starting point for such molecular studies is to profile gene expression. That is, outlining where the gene is expressed (i.e., what tissue or cell), when the gene is expressed (developmental stage, environment, season, etc.), the degree to which a gene is expressed relative to other genes or other treatments (treatment is used here to refer to a developmental stage, tissue, and environmental condition, etc.), and finally, what happens when the gene is purposefully over-expressed or physically turned off. Indeed, recent advances in assessing gene expression have allowed biologists to pinpoint the genetic basis of major evolutionary transitions (e.g., limblessness in snakes, Cohn and Tickle, 1999) and even adaptive traits contributing to species radiations (e.g., beak depth and length in Darwin's finches, Abzhanov et al., 2004, 2006).

Turtles may serve as an excellent system in which to analyze a wide array of biological phenomena, such as temperature-dependant sex determination, cold tolerance,

and shell development, in a genetic and genomic context. Thus, we review technologies that allow researchers to profile gene expression.

The Candidate Gene Approach

The candidate gene approach is one where a gene shown to perform a particular function in model systems is examined for a similar role in non-model organisms. For example, several genes known to be involved in the sex determination pathway of mammals and birds were profiled in turtles and may have important roles in temperature-dependent sex determination (Spotila et al., 1998; Kettlewell et al., 2000; Place et al., 2001; Loffler et al., 2003; Murdock and Wibbels, 2003a,b; Place and Lance, 2004; Valenzuela et al., 2006; Valenzuela and Shikano, 2007). Candidate genes have also lead to a greater understanding of shell and body plan development in turtles (Gilbert et al., 2001; Loredo et al., 2001; Vincent et al., 2003; Ohya et al., 2005). Interestingly, examination of Hox gene expression, major controllers of anterior-posterior body axis in development, in *Pelodiscus sinensis* showed definite discrepancies in the way turtles, as opposed to mammals and birds, build their body (Ohya et al., 2005). The candidate gene approach is a relatively inexpensive way to discover expression pattern and level differences among lineages and treatments and can be imagined to help unravel the several turtle queries like the ones outlined below.

Future Directions. — Convergent evolution in head shape of the bigheaded turtle, *Platysternon megacephalum*, and the alligator snapping turtle, *Macrochelys temminckii,* could be explored using the same genes that partly control beak dimensions in Darwin's finches (Abzhanov et al., 2004, 2006 [bone morphogenetic protein – 4 and calmodulin]) or molecular genetic effects of inhabiting polluted, fragmented landscapes could be assayed through examining levels of typical stress response genes (Evron et al., 2006; Grisaru et al., 2006; Song et al., 1991 [i.e. acetylcholinesterase and the glucocorticoid receptor]) in turtles living in degraded versus relatively pristine habitats. Although the candidate gene approach is extremely valuable, the opportunity to profile expression of thousands of genes in nonmodel organisms is becoming rapidly accessible through complementary techniques, some of which have actually been implemented in a turtle system (Kuraku et al., 2005; Storey, 2005).

Complementary DNA (cDNA) and Expressed Sequence Tags (ESTs)

Full-length cDNAs are DNA copies of messenger RNA (mRNA) transcripts created by a process called Reverse-Transcriptase PCR (**RT-PCR**; capable of reverse transcription up to about 20kb; Fig. 5). As DNA is inherently more stable than RNA, cDNA provides a way to keep a "library" of the organism's tissue/condition-specific transcriptome cloned into **plasmid** vectors (circular pieces of bacterial or phage DNA; detailed in Becker et al. [2003]). Expressed Sequence Tags (ESTs)

are generally created by one sequencing reaction from a cDNA clone, range between 200-800 nucleotides long, and provide a snippet of data with which one can identify genes that are being expressed in a certain treatment (Holloway et al., 2002). This technique allows for the relatively cheap, fast generation of large amounts of transcript data which can be an invaluable resource for studies of evolution and development. As of August 2007, over 45 million of these snippets from a variety of organisms and treatments were available through the national EST repositories (dbEST and Unigene databases from the National Center for Biotechnology Information [NCBI]).

Applications of cDNA and ESTs

Gene Discovery and Identification. — First developed in 1991 for use in human gene discovery, ESTs are one of the most useful tools for gene identification (Adams et al., 1991; Wolfsberg and Landsman, 1997). Since ESTs represent functional mRNA, they provide a gene expression profile from the treatment from which the mRNA was extracted (McCarter et al., 2000). Homologs and functional groups can be identified by comparing novel EST data to data created by other sequencing efforts (Ton et al., 2000). Full sequences of informative cDNAs can then be retrieved by sequencing the entire clone. For example, cDNA library screens were used to identify anoxia responsive genes in *Trachemys scripta elegans* and freeze responsive genes in *Chrysemys picta marginata* (Storey, 2005).

ESTs can also identify similar but unique transcripts of the same gene (i.e., isoforms). When aligned with genomic DNA, ESTs can illuminate splice variants, exon boundaries, and polymorphisms in untranslated regions (Wolfsberg and Landsman, 1997; Ulrich, 2000; Gemünd et al., 2001).

Marker Development. — Phylogenetic and phylogeographic studies are enhanced by the use of multiple, unlinked markers and existing EST projects as well as turtle specific EST projects, can generate primers to accomplish this (Brumfield et al., 2003). Because ESTs are copies of
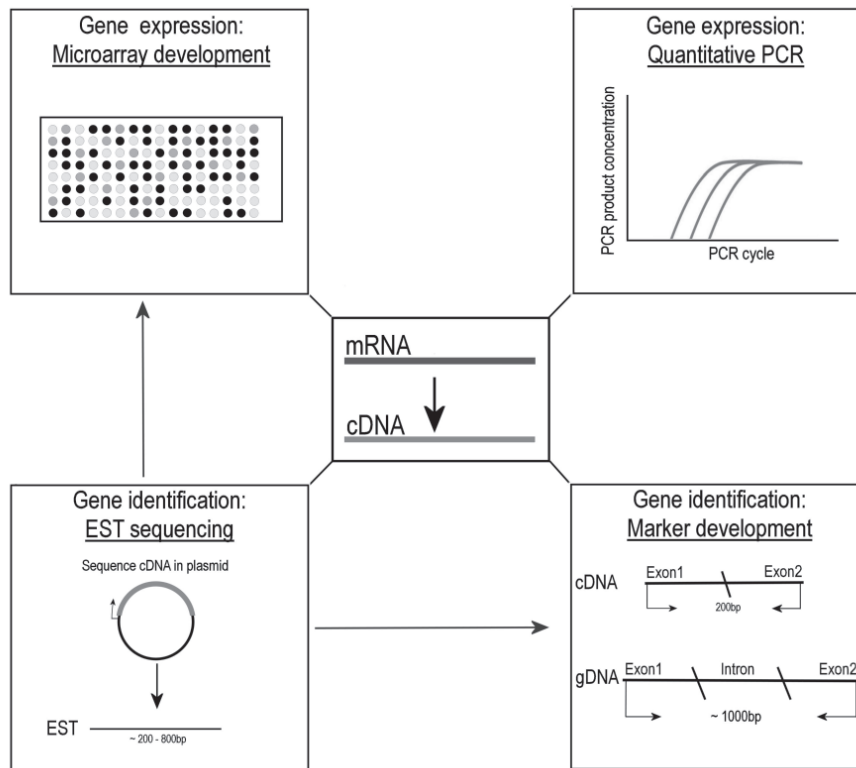
Figure 5. Messenger RNA (mRNA) is specific to the tissue and time it was taken from an organism. mRNA can be reverse transcribed into complementary DNA (cDNA) for a variety of uses. **Bottom Left:** cDNA can be cloned into a plasmid and sequenced to produce an expressed sequence tag (EST). An EST is one sequence read from an end of the cloned cDNA. When many ESTs are sequenced one can obtain a good estimate of which genes are expressed. **Top Left:** EST collections and cDNA clones can both be utilized to create a microarray. Thousands of these DNA seqments, called "probes," are printed on a specially treated glass slide. Each dot on the example slide represents a probe. Shown here is the result of an experiment using two conditions, for instance, warm and cold temperatures during sex determination. The lightest dots represent those probes that are over-expressed in the cold treatment relative to the warm treatment. The darkest dots represent those probes that are over-expressed in the warm treatment relative to the cold treatment. Dots of medium brightness represent genes that are relatively evenly expressed in both treatments. **Top Right:** Quantitative PCR assays gene expression for a limited number of genes. The relative starting concentrations of genes are measured by surveying the quantity of PCR product at each PCR cycle, leading to this method also being called real-time PCR. **Bottom Right:** EST sequencing can provide thousands of potential markers. One way to identify variable markers is to develop primers in two exons of a cDNA and use the same primers to amplify the gDNA. These primers will span an intron, an often variable nuclear region.

mRNA and do not include intron sequences, conserved primers can be anchored in ESTs that might amplify variable introns across disparate turtle groups (Fig. 5; similar to the strategy employed by Fujita et al. [2004] to discover the nuclear intron R35). SNPs can also be revealed by comparing ESTs between closely related species. In fact, empirical data suggest that each EST will contain *at least* one SNP (Brumfield et al., 2003). Lastly, ESTs can serve as probes for BAC libraries and isolate a gene of interest even when used from related species (McCarter et al., 2000).

Evolutionary Genetics. — The shear volume of ESTs generated by gene discovery projects provides a resource for surveys of genomic variation for evolutionary studies. For example, in a gene discovery project for chicken skeletal system development, over 6000 ESTs were generated (Jorge et al., 2004).

Jaramillo-Correa et al. (2001) used quantitative traits and markers developed from polymorphic ESTs to assay for signatures of population differentiation and compared these measures to investigate adaptive evolution (Q_{ST}-F_{ST} comparison) in white spruce, *Picea glauca*. Likewise, 95 microsatellite loci in noncoding regions of transcripts from Atlantic salmon, *Salmo salar,* were identified in EST databases and tested for signatures of selection, despite the fact that microsatellites are thought to evolve in a neutral fashion (Vasemägi et al., 2005). These authors also showed that some microsatellites displayed non-neutral patterns of evolution because they were tightly linked to genes under selection. Following loci with non-neutral patterns of evolution may be especially useful in identifying genes affected by selection in taxa such as turtles that lack extensive genetic resources or linkage maps.

Limitations. — Unlike genomic DNA, which will be relatively uniform in nearly every somatic cell in an organism's body, specific mRNAs will only be found in the specific tissues and during times when the gene is being expressed. This can make the acquisition of specific mRNAs difficult because RNA must be taken from the proper tissue during the treatment or developmental stage of interest, often requiring that specimens be sacrificed in the process. Further, RNA molecules are inherently more unstable than DNA, and in this respect, special care must be taken when handling samples in order to avoid contamination by somewhat ubiquitous RNA degraders called RNAses. Historically, tissue samples would be flash frozen in liquid nitrogen to preserve the molecule's integrity, but preservation products (e.g., RNAlater) have been developed and may provide better results if optimal harvest and storage conditions are not met. In addition, some studies have succeeded in extracting EST quality data from formalin-fixed, paraffin-embedded tissue that may enable the use of preserved specimens for gene identification, but not quantification (Lewis et al., 2001). Difficulties of formalin-fixed nucleic acid extraction have also been prohibitive, however, a successful extraction protocol could help researchers achieve a much deeper phylogenetic sampling, as the extensive turtle collections in museums are often times much easier to access than fresh field specimens.

Once the RNA is extracted and analyzed, researchers must take into account that the transcription of a gene into mRNA does not necessarily mean the mRNA is translated into a protein product. Several molecular mechanisms for silencing and stability reduction of transcripts are known (detailed in RNAi section below), and so upregulation of a gene's transcription is not sufficient to demonstrate that a gene is responsible for a certain phenotype. Functional assays are usually required to confirm that increased mRNA transcription is responsible, or partly responsible, for the phenotype exhibited. Most commonly, functional assays include inoculating the organism with a recombinant viral vector to over-express the gene of interest (detailed in Smith and Sinclair, 2004) or employing RNAi to turn off a gene's expression (see below).

Further, cDNA library construction is subject to contamination by bacteriophages, bias toward smaller, more abundant mRNAs, and is only truly relevant for the tissue, time, and development stage from which it is made (detailed in Becker et al., 2003). These problems are compounded in future applications like EST generation. The use of kits or contracting experienced companies can ameliorate contamination and biases while not contributing excessively to the cost of an experiment (Lucigen, AmpliconExpress, GATC Biotech [typical cost is approximately $6000 to supply a tissue sample and receive a complete library in return]). Also, although as many or as few ESTs can be generated from a cDNA library, typically several thousand sequence reads may be needed for the EST collection to have much utility and justify the cost of a relatively pricey cDNA construction. In such mass sequencing missions, ESTs are typically not checked for sequencing errors because minor mistakes usually do not prevent the matching of the EST to sequences of other organisms for identification (McCarter et al., 2000). This tolerance for inaccuracies may pose a problem if the ESTs are used for applications like SNP detection or protein sequence prediction. Lastly, sequencing of ESTs will result in redundant data. Although over 6000 ESTs were sequenced in the chicken development study mentioned above, only 2329 were unique after clustering (Jorge et al., 2004).

Future Directions. — Large amounts of EST data may be expected for future projects in chelonian gene discovery. Subsequently, thousands of potential phylogenetic markers will be generated by these large EST projects. Bapteste et al. (2002) illustrated the power of ESTs by identifying 123 orthologous genes which helped to resolve important, but previously unclear phylogenetic relationships in amoeboid lineages.

Real-Time PCR

Real-Time PCR (RT-PCR) or Quantitative PCR (Q-PCR) is capable of tracking the amount of amplified DNA produced at each cycle with the use of fluorescent dyes, thus

allowing the quantification of the initial RNA template (Fig. 5). The acronym RT-PCR is also commonly used for Reverse Transcriptase PCR, where an RNA template is converted to cDNA. These are simple, sensitive techniques for quantifying the relative number of gene transcripts in a particular tissue sample and are especially good for use with small sample volumes and discerning between related transcripts. The method is performed by measuring the PCR cycle at which the fluorescently labeled product can first be detected above background fluorescence. If, for instance, more copies of a particular gene are present in condition A than in condition B, then condition A's product will be detected at an earlier cycle, and it may be concluded that the gene being investigated is being up-regulated or over-expressed in condition A.

Gene Discovery and Comparative Gene Expression. — Comparing transcript levels is necessary in gene expression studies and is useful in understanding differences across treatments, individuals, species, etc. Differences in gene expression, along with knowledge of the suspected gene function in other species, can help infer the gene's role. For example, Kettlewell et al. (2000) used Q-PCR to assess expression levels of *Dmrt-1* in developing male and female embryos of a turtle with temperature-dependent sex determination (TSD), *Trachemys scripta,* and discovered that *Dmrt-1* shows higher levels of expression in males than in females. Because *Dmrt-1* also performs male specific functions in a range of taxa, and even has functional and sequence homologs in *Drosophila* and *Caenorhabditis elegans*, these authors suggested that *Dmrt-1* is important for sex determination in *T. scripta*. Further, comparative gene expression profiling has also provided insight into the evolutionary divergence of the developmental network underlying sex determination in turtles, helping identify candidate genes (e.g., *Sf1*) for the role of master TSD switch (Valenzuela et al., 2006; Valenzuela and Shikano, 2007). The utility of quantifying transcripts and making cDNAs for gene discovery can be expanded to identify genetic signatures of local and clinal adaptation, or to understand physiological processes, environmental response, ontogeny, and phylogenetic relationships (Gibson, 2002).

Limitations. — Because QRT-PCR is used to measure difference is transcript number, and because mRNA is so unstable, the QRT-PCR method is extremely sensitive to investigator error. For example, if one sample is fresher or bigger, then it might yield far more transcript copies than another (Wong and Medrano, 2005). To account for some of these issues, investigators should employ a normalization method. A conservative normalization method is to measure multiple housekeeping genes (i.e., genes that are constitutively 'on' and relatively evenly expressed across tissues and individuals along with each sample; Wong and Medrano, 2005).

Further, many genes are modified by transcription and translation machinery differently. That is to say, the same genomic DNA may make multiple mRNAs by using different translation or transcription start and stop sites and different intron splice sites, resulting in different 'isoforms' of the same gene (Weaver, 2005). All of these can contribute to functional differences in the mRNA's role. Therefore, when measuring the amount of mRNA with Q-PCR, primers should be designed that only amplify the functional isoform of interest. Otherwise, the number of transcripts being measured may be artificially inflated because investigators are actually measuring many related but functionally nonsynonymous transcripts.

Microarrays

Since their introduction in 1991 (Fodor et al., 1991), microarrays have been employed successfully to explore relative gene expression in many systems in a high-throughput way. As with Q-PCR and QRT-PCR, microarrays can tell researchers what genes are being expressed, when they are being expressed, where they are being expressed, and to some degree, how much they are being expressed. Microarrays, however, are not limited to small sets of genes of known sequence as in QRT-PCR. Furthermore, microarrays can be adapted to scan tens of thousands of genes, sometimes without knowing their sequence (anonymous cDNA microarrays). This high-throughput ability gives researchers enormous possibilities in understanding phenotypes and interactions between the genotype and the environment.

Traditional microarrays attach 'probes' (Fig. 5; cDNAs, oligonucleotides made from ESTs, genomic sequence, or even BACs, also called 'features') to a pretreated glass slide. These probes then hybridize to 'targets,' which are fluorescently labeled cDNAs made from the mRNA of the treatment of interest, in order to assay gene expression in that treatment. Stoughton (2005) and Holloway et al. (2002) offered comprehensive reviews of this technology, but multitudes of variations on this theme are present in the literature (for alternative microarray techniques see Brenner et al., 2000; Hegarty et al., 2005).

Applications of Microarrays

Gene Discovery. — Due to the large amount of sequence information required to construct the probes for oligonucleotide microarrays, such arrays hold potential for substantial gene discovery. Hybridization of the targets to the probes helps identify genes expressed in a particular treatment as well. For example, microarrays helped to identify a suite of genes responsible for a shift in worker to foraging behavior by honey bees (*Apis mellifera*; Whitfield et al., 2003) and nearly 100 genes that are candidates in social status modifications of cichlids (Renn et al., 2004). Further, a variation on the microarray, microbeads (detailed below), allowed researchers to identify genes involved in shell formation by targeting the carapacial ridge of the Chinese soft-shelled turtle, *Pelodiscus sinensis* (Kuraku et al., 2005). Another variation on the microarray technology, employing microarrays made from model organisms in-

stead of from the species of interest, was recently used to discover genes associated with cold tolerance in turtles (Storey, 2005).

Evolutionary Genetics. — Population level applications of microarrays can help uncover unique genetic variation or variable responses to environmental pressures in populations that may be extremely difficult to discover via candidate gene approach or other traditional DNA sequencing methods. Because phenotypic diversity without large DNA sequence divergence can still signify local adaptation, changes in gene expression and regulation may be illustrative of overall disparity between species (Schlötterer, 2002). Therefore, gene expression can greatly contribute to unique evolutionary trajectories of populations and species. Microarrays can help uncover these local, possibly adaptive differences in gene expression, thus identifying unique populations that warrant conservation (Turgeon and Bernatchez, 2003).

Limitations. — Microarrays are powerful tools, and require comparable levels of statistical and bioinformatic strength in analyzing the results (Stoughton, 2005). However, even with the help of a strong bioinformatics resource, extracting biological meaning from such a large and complex dataset is an arduous, on-going process (Butte, 2002). For quantification between two treatments, typically multi-chip experiments are required and statisticians are needed to design experiments with maximum power, as factors such as the day the chips were hybridized to the scanner used to view the fluorescence can add greatly to the variability of results. Interesting gene expression results are typically confirmed using QRT-PCR, because variation in microarray output data may be due to these experimental inconsistencies and not genuine gene expression differences (Pinhasov et al., 2004).

In addition, microarrays are expensive in terms of time and money. A start to finish project (i.e., development of an array from EST construction to confirmation of results) may take a lab studying a non-model organism two to four years, even when collaborating with high-throughput labs and computational specialists (detailed Holloway et al., 2002; Bowtell and Sambrook, 2003; Stoughton, 2005).

Future Directions. — A current alternative to using turtle-specific microarrays is to hybridize turtle mRNA to prefabricated microarrays from other model species such as chicken or human. Using nonspecific microarrays can provide an invaluable starting point in gene discovery. In fact, more than twelve genes involved in freeze tolerance and anoxia in *C. p. marginata* were identified by hybridizing turtle mRNA to human microarrays (Storey, 2005). Other cheaper and quicker alternatives to typical microarrays include anonymous cDNA microarrays, focused microarrays, and macroarrays. These arrays usually provide information of similar quality and may be viable alternatives for turtle investigators (Becker et al., 2003; Wurmbach et al., 2003; Hegarty et al., 2005).

Another alternative to using species-specific microarrays is the microbead library. Kuraku et al. (2005) used this technology for gene discovery in shell formation in *P. sinesis*. Here, cDNAs from the carapacial ridge (the region

of interest for shell formation) and the thoracic region (a negative control) were "cloned" separately onto microbeads to create two libraries (Brenner et al., 2000). The two libraries were then hybridized together, automatically sorted, and ones that showed higher signals (i.e., higher expression) for the carapacial ridge were sequenced and further identified. Microbeads do not require *a priori* knowledge of sequences or chip layout design and therefore can circumvent the common prohibitive problems of cost, time, and limited tissue samples which may plague other turtle researchers interested in the microarray technology.

RNA Interference (RNAi)

RNA interference (RNAi), a type of gene silencing, can shed light on developmental and adaptive processes by "knocking down" or "knocking out" the expression of particular genes and allows observation of the effects that turning a specific gene down or off has on particular phenotypes (see Mello and Cante, 2004, for more technical information). RNAi takes advantage of an innate defense system used by the organism which degrades double stranded RNA in a sequence specific fashion (Guo and Kempheus, 1995; Fire et al., 1998; reviewed in Cogoni and Macino, 2000; Guru, 2000; Hammond et al., 2001). By introducing foreign dsRNA with sequence identical or nearly identical to the gene of interest, the cell machinery naturally converts them into small RNA (siRNA or microRNA [miRNA]), which target mRNA similar in sequence for degradation and reduced gene expression. Thus, the silenced gene is transcribed but rapid degradation of the transcripts prevents their accumulation and associated function. Small RNA can also down-regulate gene expression by transcriptional silencing, or translational inhibition of mismatched targets (Morris et al., 2004).

Applications of RNAi

Gene Function. — RNAi techniques are well suited for developmental and physiological studies to determine gene function, genetic pathway analysis, and to examine gene redundancy. As such, this technique can be extended to investigate fitness consequences associated with particular genes and gene functions, and thus to examine the genetic variability underlying adaptive variation and adaptive potential in particular taxa. Its main strength derives from being an experimental rather than a correlative approach to identifying genetic variation underlying target phenotypes with important fitness consequences. Although still incipient in its application to vertebrates *in vivo*, this and related techniques hold promise as tools to experimentally study target gene regulation and loss-of-function screening (Cullen, 2005). This derives from the fact that natural miRNA play a key role in regulating vertebrate differentiation and development and thus, RNAi loss of function screening can shed light on the connections and biological functions of biochemical pathways (Silva et al., 2005, Wienholds and Plasterk, 2005). Important functions that have been targeted

for study by this approach in vertebrates include DNA repair, apoptosis, cancer, and response to drugs among many others (Silva et al., 2005; Dickins et al., 2005) that may have significant therapeutic applications. Similar experimental analysis is plausible for biological phenomena relevant to turtles, such as temperature tolerance, courtship and nesting behavior, sex determination, and aging, among others, as this technique allows the experimental identification of those genes that are necessary and sufficient for particular phenotypes.

Limitations and Future Directions. — The main limitation of these methods is the high level of technical expertise and associated costs in time and money, making them unsuitable for the average ecological genetics laboratory (Mello and Cante, 2004, for technical information). RNAi is an increasingly powerful tool to determine gene function and its fitness consequences such that collaborative work should be considered to solve their logistic limitations to answer questions in developmental biology with significant implications for ecology, evolutionary biology, and conservation. A rising number of companies offer RNAi products and services (e.g., Ambion, Integrated DNA Technologies, Invitrogen) that parallel the expanding use of these techniques by research laboratories and derived publications, including a dedicated journal (Journal of RNAi and Gene Silencing) that can be found online.

Conclusion

In conjunction with ecological and behavioral studies, genetic and genomic data offer exciting possibilities for valuable insight into the evolution and biology of chelonians. The techniques presented here have been successful in other systems, and will help to explore how turtles fit into their ecological communities and are affected by their environment. With this understanding, we will be able to more fully appreciate the complexity of these animals and their unique biological interactions, ultimately ensuring more successful conservation efforts.

Currently 37% of the world's 309 turtle species are provided protection under the Convention on International Trade of Endangered Species, and of 181 species listed by the IUCN Red List, 69% are identified as threatened, endangered, or vulnerable (IUCN, 2004). As turtles are species of great conservation concern, additional information gleaned from the fields of molecular ecology and evolutionary biology can be incorporated directly and rapidly into conservation programs. Although this review has provided only a brief description of new technologies, the future implementation of molecular markers will provide great insights into the fundamental biology of turtles and potentially how best to ensure their survival.

LITERATURE CITED

Abzhanov A., Protas M., Grant B.R., Grant P.R., Tabin C.J. 2004. Bmp4 and morphological variation of beaks in Darwin's finches. Science 305:1462-1465.

Abzhanov, A., Kuo, W.P., Hartmann C., Grant B.R., Grant P.R., Tabin C.J. 2006. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. Nature 442:563-567.

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos M.H., Xiao H., Merril C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., Venter., J.C. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252:1651-1656.

Aitken, N., Smith, S., Schwarz, C., Morin, P.A. 2004. Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. Molecular Ecology 13:1423-1431.

Akey, J.M., Zhang, G., Jin, L., Shriver, M.D. 2002. Interrogating a high-density SNP map for signatures of natural selection. Genome Research 12:1805-1814.

Albertson, R.C., Markert, J.A., Danley, P.D., and Kocher, T.D. 1999. Phylogeny of a rapidly evolving clade: the cichlid fishes of Lake Malawi, East Africa. Proceedings of the National Academy of Sciences, USA 96:5107–5110.

Altshuler, D., Pollara, V.J., Cowles, C.R. Van Etten, W.J., Baldwin, J., Linton, L., and Lander, E.S. 2000. An SNP map of the human genome generated by reduced representation shotgun sequencing. Nature 407:513-516.

Amos, W., Hutter, C.M., Schug, M.D. and Aquadro, C.F. 2003. Directional Evolution of Size Coupled with Ascertainment Bias for Variation in Drosophila Microsatellites. Molecular Biology and Evolution 20:660-662.

Avila, L.J., Morando, M.M., and Sites, J.W., Jr. 2006. Congeneric phylogeography: hypothesizing species limits and evolutionary processes in Patagonian lizards of the *Liolaemus boulengeri* group (Squamata: Liolaemini). Biological Journal Linnean Society 89:241-275.

Avise, J.C. 2000. Phylogeography. The History and Formation of Species. Harvard University Press, Cambridge, MA.

Avise, J.C. 2004. Molecular Markers, Natural History, and Evolution. Sinauer, Sunderland, MA.

Bachem, C.B.W., Oomen, R.J.F.J. and Visser, R.G.F. 1998. Transcript Imaging with cDNA-AFLP: A Step-by-Step Protocol. Plant Molecular Biology Reporter 16:157.

Bachem, C.W.B., van der Hoeven, R.S., de Bruijn, S.M., Vreugdenhil, D., Zabeau, M. and Visser, R.G.F. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. The Plant Journal 9:745–753.

Ballard, J.W.O. and Whitlock, M.C. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13:729–744.

Balloux, F. H., Brunner, N. Lugon-Moulin, Hausser, J., Goudet, J. 2000. Microsatellites can be misleading—an empirical and simulation study. Evolution 54:1414-1422.

BAPTESTE, E., BRINKMANN, H., LEE, J.A., MOORE, D.V., SENSEN, C.W., GORDON, P., DURUFLÉ, L., GAASTERLAND, T., LOPEZ, P., MÜLLER, M. AND PHILIPPE, H. 2002. The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium, Entamoeba,* and *Mastigamoeba*. Proceedings of the National Academy of Sciences 99:1414-1419.

BATZER M.A., ARCOT S.S., PHINNEY, J.W., ALEGRIA-HARTMAN, M., KASS, D.H., MILLIGAN, S.M., KIMPTON, C., GILL, P., HOCHMEISTER, M., IOANNOU, P.A., HERRERA, R.J., BOUDREAU, D.A., SCHEER, W.D., KEATS, B.J., DEININGER, P.L., STONEKING, M.1996. Genetic variation of recent Alu insertions in human populations. Journal of Molecular Evolution 42:22-29.

BECKER, K.G., WOOD III, W.H. AND CHEADLE, C. 2003. Membrane-based spotted cDNA arrays. In: DNA Microarrays, eds. Bowtell, D. and J. Sambrook. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.

BELFIORE, N.M., HOFFMAN, F.G., BAKER, R.J., DEWOODY, J.A. 2003. The use of nuclear and mitochondrial single nucleotide polymorphisms to identify cryptic species. Molecular Ecology 12:2011-2017.

BENNETT, P. 2000. Microsatellites. Journal of Clinical Pathology – Molecular Pathology 53:177-183.

BENSASSON, D. ZHANG, D.X., HARTL, D.L., HEWITT, G.M. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in Ecology and Evolution 16:314-321.

BENSCH, S., AND AKESSON., M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? Molecular Ecology 14:2899-2914.

BENSCH, S., AKESSON, S., IRWIN, D.E. 2002. The use of AFLP to find an informative SNP: genetic differences across a migratory divide in willow warblers. Molecular Ecology 11:2359-2366.

BERGSTEN, J. 2005. A review of long-branch attraction. Cladistics 21:163-193.

BERRY, O., TOCHER, M.D., AND SARRE, S.D. 2004. Can assignment tests measure dispersal? Molecular Ecology 13:551-561.

BOORE, J. L. 1999. Animal mitochondrial genomes. Nucleic Acids Research 27:1767-1780.

BOORE, J. L., COLLINS, T.M., STANTON, D., DAEHLER, L.L., AND BROWN, W.M. 1995. Deducing the pattern of arthropod phylogeny from mitochondrial gene rearrangements. Nature 376:163-165.

BORODULINA, O.R. AND KRAMEROV, D.A. 2005. PCR-based approach to SINE isolation: Simple and complex SINEs. Gene 349:197–205.

BOWEN, B.W., MEYLAN, A.B., AND AVISE, J.C. 1989. An odyssey of the green sea turtle: Ascension Island revisited. Proceedings of the National Academy of Sciences 86:573–576.

BOWTELL, D. AND SAMBROOK, J. 2003. DNA Microarrays: A Molecular Cloning Manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.

BRANDLEY, M., SCHMITZ, A., AND REEDER, T.W. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Systematic Biology 54:373-390.

BRENNER, S., WILLIAMS, S.R., VERMAAS, E.H., STORCK, T., MOON, K., MCCOLLUM, C., MAO, J.I., LUO, S., KIRCHNER, J.J., ELETR, S., DUBRIDGE, R.B., BURCHAM, T., ALBRECHT, G. 2000. In vitro cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. Proceedings of the National Academy of Sciences U S A. 97:1665-1670.

BRUGMANS, B., VAN DER HULST, R.G.M., VISSER, R.G.F., LINDHOUT, P., VAN ECK, H.J. 2003. A new and versatile method for the successful conversion of AFLP markers into simple single locus markers. Nucleic Acids Research 31:e55.

BRUMFIELD, R., BEERLI, P., NICKERSON, D., EDWARDS, S. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology and Evolution 18:249-256.

BURNS, C.E., CIOFI, C., BEHEREGARAY, L.B., FRITTS, T.H., GIBBS, J.P., MÁRQUEZ, C., MILINKOVITCH, M.C., POWELL, J.R., AND CACCONE, A. 2003. The origin of captive Galápagos tortoises based on DNA analysis: implications for the management of natural populations. Animal Conservation 6: 329-337.

BUTTE, A. 2002. The use and analysis of microarray data. Nature Reviews Drug Discovery 1:951-960.

CACCONE, A., GENTILE, G., BURNS, C.E., SEZZI, E., BERGMAN, W., RUELLE, M., SALTONSTALL, K., AND POWELL, J.R. 2004. Extreme difference in rate of mitochondrial and nuclear DNA evolution in a large ectotherm, Galápagos tortoises. Molecular Phylogenetics and Evolution 31:794-798.

CAMPBELL, D., AND BERNATCHEZ, L. 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. Molecular Biology and Evolution 21:945–956.

CARVAJAL, J.J., COX, D., SUMMERBELL, D., AND RIGBY, P.W. 2001. A BAC transgenic analysis of the Mrf4/Myf5 locus reveals interdigitated elements that control activation and maintenance of gene expression during muscle development. Development 128:1857-1868.

CASSENS, I., VAN WAEREBEEK, K., BEST, P.B., CRESPO, E. A., REYES, J., AND MILINKOVITCH, M.C. 2003. The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): A critical examination of network methods and rooting procedures. Molecular Ecology 12:1781-1792.

CASSENS, I., MARDULYN, P., AND MILINKOVITCH, M.C. 2005. Evaluating intraspecific network construction methods using simulated sequence data: do existing algorithms outperform the global maximum parsimony approach? Systematic Biology 54:363-372.

CHILTON, N.B. 2004. The use of nuclear ribosomal DNA markers for the identification of bursate nematodes (order Strongylida) and for the diagnosis of infections. Animal Health Research Reviews 5:173–187.

CHIU, C.-H., DEWAR, K., WAGNER, G. P., TAKAHASHI, K., RUDDLE, F., LEDJE, C., BARTSCH P., SCEMAMA, J-L., STELLWAG, E., FRIED, C., PROHASKA, S.J., STADLER, P.F. AND AMEMIYA, C.T. 2004. Bichir HoxA cluster sequence reveals surprising trends in ray-finned fish genomic evolution. Genome Research 14:11-17.

COGONI, C., AND MACINO, G. 2000. Post-transcriptional gene silencing across kingdoms. Genes and Development 10:638-643.

COHN, M.J. AND TICKLE, C. 1999. Developmental basis of limblessness and axial patterning in snakes. Nature 399:474-9.

COLSON, I. AND GOLDSTEIN, D.B. 1999. Evidence for complete mutations at microsatellite loci in Drosophila. Genetics 152:617-627.

CRANDALL, K.A., AND BUHAY, J.H. 2004. Genomic databases and the Tree of Life. Science 306:1144-1145.

CUMMINGS, M.P., OTTO, S.P., AND WAKELEY, J. 1995. Sampling properties of DNA sequence data in phylogenetic analysis. Molecular Biology and Evolution 12:814-822.

CULLEN, B.R. 2005. RNAi the natural way. Nature Genetics 37:1163-1165.

CUROLE, J.P., AND KOCHER, T.D. 1999. Mitogenomics: digging deeper with complete mitochondrial genomes. Trends in Ecology and Evolution 14:394-398.

DALMA-WEISZHAUSZ, D.D. AND MURPHY, G.M. 2002. Single nucleotide polymorphisms and their characterization with oligonucleotide microarrays. Psychiatric Genetics 12:97-107.

DE LA VEGA, F.M., DAILEY, D., ZIEGLE, J., WILLIAMS, J., MADDEN, D., GILBERT, D.A. 2002. New generation pharmacogenomic tools: A SNP linkage disequilibrium map, validated SNP assay resource, and high-throughput instrumentation system for large-scale genetic studies. Biotechniques 48-50, 52, 54.

DEAN, F.B., NELSON, J.R., GIESLER, T.L., AND LASKEN, R.S. 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA

polymerase and multiply-primed rolling circle. Genome Research 11:1095-1099.

DEL POZZO, G. AND GUARDIOLA, J. 1990. A SINE insertion provides information on the divergence of the HLA-DQA1 and HLA-DQA2 genes. Immunogenetics 31:229-232.

DESALLE, R. AND AMATO, G. 2004. The expansion of conservation genetics. Nature Reviews Genetics 5:702-712.

DEWOODY, J.A. 1999. Nucleotide variation in the p53 tumor suppressor gene of voles from Chernobyl, Ukraine. Mutation Research 439:25-36.

DEWOODY, J.A.D. AND AVISE, J.C. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. Journal of Fisheries Biology 56:461-473.

DICKINS, R.A., HEMANN, M.T., ZILFOU, J.T., SIMPSON, D.R., IBARRA, I., HANNON, G.J., AND LOWE, S.W. 2005. Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. Nature Genetics 37:1289-1295.

DOIRON, S., BERNATCHEZ, L. AND BLIER, P.U. 2002. A Comparative mitogenomic analysis of the potential adaptive value of arctic charr mtDNA introgression in brook charr populations (*Salvelinus fontinalis* Mitchill). Molecular Biology and Evolution 19:1902-1909.

EDWARDS, S. V., JENNINGS, W.B., AND SHEDLOCK, A.M. 2005. Phylogenetics of modern birds in the era of genomics. Proceedings of the Royal Society B-Biological Sciences 272:979-992.

EDWARDS, S.V. AND BEERLI, P. 2000. Perspective: gene divergence, and the variance in coalescent time in phylogeographic studies. Evolution 54:1839-1854.

ELLEGREN, H. 2000. Heterogeneous mutation processes in human microsatellite DNA sequences. Nature Genetics 24:400–402.

ENGSTROM, T.N., SHAFFER, H.B. AND MCCORD, W.P. 2004. Multiple data sets, high homoplasy and the phylogeny of softshell turtles (Testudines: Trionychidae). Systematic Biology 53:693-710.

ENGSTROM, T.N., EDWARDS, T., OSENTOSKI, M.F., AND MYERS, E.M. 2007. A compendium of PCR primers for mtDNA, microsatellite, and other nuclear loci for freshwater turtles and tortoises. Chelonian Research Monographs 4:124-141.

ESTOUP, A., JARNE, P. AND CORNUET, J-M. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Molecular Ecology 11:1591-1604.

EVRON, T., GREENBERG, D., MOR, T.S., AND SOREQ, H. 2006. Adaptive changes in acetylcholinesterase gene expression as mediators of recovery from chemical and biological insults. Toxicology 233:97-107.

FELDMAN, C.R. AND PARHAM, J.F. 2002. Molecular phylogenetics of Emydine turtles: taxonomic revision and the evolution of shell kinesis. Molecular Phylogenetics and Evolution 22:388-398.

FELDMAN, C.R. AND J.F. PARHAM. 2004. Molecular systematics of Old World stripe-necked turtles (Testudines: *Mauremys*). Asiatic Herpetological Research 10:28-37.

FIRE, A., XU, S., MONTGOMERY, M.K., KOSTAS, S.A., DRIVER, S.E., AND MELLO, C.C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391:806-811.

FISCHER, D., AND BACHMANN, K. 1998. Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). BioTechniques 24:796–802.

FISHER, M.C., KOENIG, G., WHITE, T.J. AND TAYLOR, J.W. 2000. A test for concordance between the multilocus genealogies of genes and microsatellites in the pathogenic fungus *Coccidioides immitis*. Molecular Biology and Evolution 17:1164-1174.

FITZPATRICK, B.M. AND SHAFFER, H. B. 2004. Environment-dependent admixture dynamics in a tiger salamander hybrid zone. Evolution 58:1282–1293.

FITZSIMMONS, N.N. AND HART, K.M. 2007. Genetic studies of freshwater turtles and tortoises: a review of the past 70 years. Chelonian Research Monographs 4:15-46.

FITZSIMMONS, N.N. LIMPUS, C.J., NORMAN, J.A., GOLDIZEN, A.R., MILLER, J.D., AND MORITZ, C. 1997a. Philopatry of male marine turtles inferred from mitochondrial DNA markers. Proceedings of the National Academy of Sciences 94: 8912-8917.

FITZSIMMONS, N.N., MORITZ, C., LIMPUS, C.J., POPE, L., AND PRINCE, R. 1997b. Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. Genetics 147:1843-1854.

FODOR, S.P.A., READ, J.L., PIRRUNG, M.C., STRYER, L., LU, A.T., SOLAS, D. 1991. Light-directed, spatially addressable parallel chemical synthesis. Science 251:767-773.

FRANKHAM, B., BALLOU, J.D., AND BRISCOE, D.A. 2002. Introduction to Conservation Genetics. Cambridge University Press, New York, NY.

FRASER, D.L. AND BERNATCHEZ, L. 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Molecular Ecology 10:2741-2752.

FRIESEN, V.L., CONGDON, B.C., WALSH, H.E., AND BIRT, T.P. 1997. Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. Molecular Ecology 6:1047-1058.

FRITZ, U., SIROKY, P., KAMI, H., AND WINK, M. 2005. Environmentally caused dwarfism or a valid species—is *Testudo weissingeri* Bour, 1996 a distinct evolutionary lineage? New evidence from mitochondrial and nuclear genomic markers. Molecular Phylogenetics and Evolution 37: 389-401.

FUJITA, M.F., ENGSTROM, T.N., STARKEY, D.E., AND SHAFFER, H.B. 2004. Turtle phylogeny: insights from a novel nuclear intron. Molecular Phylogenetics Evolution 31:1031-1040.

FUNK, D.J. AND OMLAND, K.E. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution and Systematics 34:397–423.

GEORGES, A., ADAMS, M. AND MCCORD, W. 2002. Electrophoretic delineation of species boundaries within the genus *Chelodina* (Testudines: Chelidae) of Australia, New Guinea and Indonesia. Zoological Journal of the Linnean Society 134: 401-421.

GEMÜND, C., RAMU, C., ALTENBERG-GREULICH, B., GIBSON, T.J. 2001. Gene2EST: a BLAST2 server for searching expressed sequence tag (EST) databases with eukaryotic gene-sized queries. Nucleic Acids Research 29:1272-1277.

GERBER, S., MARIETTE, S., STREIFF, R., BODENES, C., AND KREMER, A. 2000. Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. Molecular Ecology 9:1037–1048.

GIBSON, G. 2002. Microarrays in ecology and evolution: a preview. Molecular Ecology 11:17-24.

GILBERT, S.F., LOREDO, G.A., BRUKMAN, A., AND BURKE, A.C. 2001. Morphogenesis of the turtle shell: the development of a novel structure in tetrapod evolution. Evolution and Development 3:47-58.

GIRALDO, P., AND MONTOLIU, L. 2001. Size matters: use of YACs, BACs and PACs in transgenic animals. Transgenic Research 10:83-103.

GLAUBITZ, J.C., RHODES, O.E., DEWOODY, J.A. 2003. Prospects for inferring pairwise relationships with single nucleotide polymorphisms. Molecular Ecology 12:1039-1047.

GRAVES, J. A. M., AND SHETTY, S. 2001. Sex from W to Z: Evolution of vertebrate sex chromosomes and sex determining genes. Journal of Experimental Zoology 290:449-462.

GRIFFITHS, R. AND ORR, K. 1999. The use of amplified fragment length polymorphism (AFLP) in the isolation of sex-specific markers. Molecular Ecology 8: 671-674.

GRISARU, D., PICK, M., PERRY, C., SKLAN, E.H., ALMOG, R., GOLDBERG, I., NAPARSTEK, E., LESSING, J.B., SOREQ, H., AND DEUTSCH, V. 2006.

Hydrolytic and nonenzymatic functions of acetylcholinesterase comodulate hemopoietic stress responses. Journal of Immunology 176:27-35.

Guo, S., and Kempheus, K.J. 1995. Par-1, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell 81:611-620.

Guru, T. 2000. A silence that speaks volumes. Nature 404:804-808.

Hammond, S.M., Caudy, A.A., Hannon, G.J. 2001. Post-transcriptional gene silencing by double-stranded RNA. Nature Reviews Genetics 2:110-119.

Hare, M. 2001. Prospects for nuclear gene phylogeography. Trends in Ecology and Evolution 16:700-706.

Harris, D.E., and Murphy, L. 2001. Sequencing Bacterial Artificial Chromosomes. In: Methods in Molecular Biology (eds. Starkey MP, Elaswarapu R), pp. 217-234. Humana Press, Totowa, NJ.

Harris, D.J. 2003. Codon bias variation in C-*mos* between squamate families might distort phylogenetic inferences. Molecular Phylogenetics Evolution 27:540-544.

Hawkins, T.L., Detter, J.C. and Richardson, P.M. 2002. Whole genome amplification-applications and advances. Current Opinions Biotechnology 13:65-67.

Hedges, S. B. 1994. Molecular evidence for the origin of birds. Proceedings of the National Academy of Science 91:2621-2624.

Hedges, S.B., and Poling, S. 1999. A molecular phylogeny of reptiles. Science 283:998-1001.

Hegarty, M.J., Jones, J.M., Wilson, I.D., Barker, G.L., Coghill, J.A., Sanchez-Baracaldo, P., Liu, G., Buggs, R. J. A., Abbott, R.J., Edwards, K.J., and Hiscock, S.J. 2005. Development of anonymous cDNA microarrays to study changes to the *Senecio* floral transcriptome during hybrid speciation. Molecular Ecology 14:2493-2510.

Heintz, N. 2000. Analysis of mammalian central nervous system gene expression and function using bacterial artificial chromosome-mediated transgenesis. Human Molecular Genetics 9:937-943.

Heller, M. 2002. DNA microarray technology:devices, systems and applications. Annual Review of Biomedical Engineering 4:129-153.

Hennig, W. 1966. *Phylogenetic Systematics*. Urbana, IL: Univ. of Illinois Press.

Hickerson, M.J. and Cunningham, C. 2005. Contrasting Quaternary histories in an ecologically divergent sister pair of low-dispersing intertidal fish (*Xiphister*) revealed by multilocus DNA analysis. Evolution 59:344-360.

Hillis D.M. 1990. SINEs of the perfect character. Proceedings of the National Academy of Sciences 96:9979-9981.

Holloway, A.J., R. K. van Laar, R. W. Tothill and D. D. L. Bowtell. 2002. Options available—from start to finish—for obtaining data from DNA microarrays II. Nature Genetics Supplement 32:481-489.

Howell, W.M., Jobs, M., Gyllensten, U., and Brookes, A.J. 1999. Dynamic allele-specific hybridization - A new method for scoring single nucleotide polymorphisms. Nature Biotechnology 17:87-88.

Hsieh, H-M, Huang, L-H, Tsai, L-Ch, Liu, C-L, Kuo, Y-C, Hsiao, C-T, Linacre, A, Linacre, L., James C-I. 2006. Species identification of *Kachuga tecta* using the cytochrome b gene. Journal of Forensic Sciences 51:52-56.

Hudson, R.R. and Turelli, M. 2003. Stochasticity overrules the three-times rule: genetic drift, genetic draft, and coalescent times for nuclear loci versus mitochondrial DNA. Evolution 57:192-190.

Hutter, C.M., Schug, M.D., and Aquadro, C.F. 1998. Microsatellite variation in *Drosophila melanogaster* and *Drosophila simulans*: a reciprocal test of the ascertainment bias hypothesis. Molecular Biology and Evolution 15:1620-1636.

IUCN 2004. 2004 IUCN Red List of Threatened Species. <www.redlist.org>.

Iwabe, N., Hara, Y., Kumzazwa, Y., Shibamoto, K., Saito, Y., Miyata, T., and Katoh, K. 2005. Sister group relationship of turtles to the bird-crocodilian clade revealed by nuclear DNA-coded proteins. Molecular Biology and Evolution 22:810-813.

Jaramillo-Correa, J.P., Beaulieu, J. and Bousquet, J. 2001. Contrasting evolutionary forces driving population structure at expressed sequence tag polymorphisms, allozymes and quantitative traits in white spruce. Molecular Ecology 10:2729-2740.

Jenkins, S. and Gibson, N. 2002. High-throughput SNP genotyping. Comparative and Functional Genomics 3:57-66.

Jennings, W.B. and Edwards, S.V. 2005. Speciational history of Australian grass finches (*Poephila*) inferred from 30 gene trees. Evolution 59:2033-2047.

Jorge, E.C., Monteiro-Vitorello, C.B., Alves, H.J., Silva, C.S., Balan, R.G., Patrício, M. and Coutinho, L.L. 2004. EST analysis of mRNAs expressed during embryogenesis in *Gallus gallus*. International Journal of Developmental Biology 48:333 – 337.

Kajikawa, M. and Okada, N. 2002. LINEs mobilize SINEs in the eel through shared 3' sequence. Cell 111:433-444.

Kalinowski, S.T. 2002. How many alleles per locus should be used to estimate genetic distances? Heredity 88:62-65.

Karl, S.A., Bowen, B.W., and Avise, J.C. 1992. Global population genetic structure and male-mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analyses of anonymous nuclear loci. Genetics 131:163-173.

Kazazian, H.H.Jr, and Moran J.V. 1998. The impact of L1 retrotransposons on the human genome. Nature Genetics 19:19–24.

Kettlewell, J., Raymond, C. S. and Zarkower, D. 2000. Temperature-dependent expression of turtle *Dmrt1* prior to sexual differentiation. Genesis 26:174–178.

Kidwell, M.G. 2002. Transposable elements and the evolution of genome size in eukaryotes. Genetica 115: 49-63.

Knowles, L.L. 2004. The burgeoning field of statistical phylogeography. Journal of Evolutionary Biology 17:1–10.

Koopman, W. 2005. Phylogenetic signal in AFLP data sets. Systematic Biology 54:197 - 217.

Krawczak, M. 1999. Informativity assessment for biallelic single nucleotide polymorphisms. Electrophoresis 20:1676-1681.

Krenz, J.G., Naylor, G.J., Shaffer, H.B., and Janzen, F.J. 2005. Molecular phylogenetics and evolution of turtles. Molecular Phylogenetics and Evolution 37:178-91.

Kuhner, M.K., Beerli, P, Yamato, J., and Felsenstein, J. 2000. Usefulness of single nucleotide polymorphism data for estimating population parameters. Genetics 156:439-447.

Kumazawa, Y. and Nishida, M. 1999. Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: Statistical evidence for Archosaurian affinity of turtles. Molecular Biology and Evolution 16:784-792.

Kuraku. S., Usuda, R., and Kuratani, S. 2005. Comprehensive survey of carapacial ridge-specific genes in turtle implies co-option of some regulatory genes in carapace evolution. Evolution and Development 7:3-17.

Kuo, C.-H. and Janzen, F.J. 2004. Genetic effects of a persistent bottleneck on a natural population of ornate box turtles (*Terrapene ornata*). Conservation Genetics 5:425-437.

Kwok, P.Y. and Gu, Z.J. 1999. Single nucleotide polymorphism libraries: why and how are we building them? Molecular Medicine Today 5:538-543.

LaJeunesse, T.C. 2001. Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the gnus *Symbiodinium* using the ITS region: In search of a "Species" level marker. Journal of Phycology 37:866-880.

Lamb, T., Avise, J.C., and Gibbons, J.W. 1989. Phylogeographic patterns in mitochondrial DNA of the desert tortoise (*Xerobates*

agassizi), and evolutionary relationships among the North American gopher tortoises. Evolution 43:76-87.

LEMAIRE, C., VERSINI, J.-J., AND BONHOMME, F. 2005. Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. Journal of Evolutionary Biology 18:70-80.

LEWIS, F., MAUGHAN, N.J., SMITH, V., HILLAN, K., AND QUIRKE, P. 2001. Unlocking the archive – gene expression in paraffin embedded tissue. Journal of Pathology 195:66–71.

LI, W.-H. 1997. Molecular Evolution. Sinauer Assoc., Sunderland, MA.

LI, J., BUTLER, J.M., TAN, Y.P., LIN, H., ROYER, S., OHLER, L., SHALER, T.A., HUNTER, J.M., POLLART, D.J., MONFORTE, J.A., BECKER, C.H. 1999. Single nucleotide polymorphism determination using primer extension and time-of-flight mass spectrometry. Electrophoresis 20:1258-1265.

LI, Y.C., KOROL, A.B., FAHIMA, T., BEILES, A., AND NEVO, E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Molecular Ecology 11:2453–2465.

LOFFLER, K.A., ZARKOWER, D., AND KOOPMAN, P. 2003 Etiology of ovarian failure in blepharophimosis ptosis epicanthus inversus syndrome: FOXL2 is a conserved, early-acting gene in vertebrate ovarian development. Endocrinology 144:3237-43.

LOREDO, G.A., BRUKMAN, A., HARRIS, M.P., KAGLE, D., LECLAIR, E.E., GUTMAN, R., DENNEY, E., HENKELMAN, E., MURRAY, B.P., FALLON, J.F., TUAN, R.S., AND GILBERT, S.F. 2001. Development of an evolutionarily novel structure: fibroblast growth factor expression in the carapacial ridge of turtle embryos. Journal of Experimental Zoology 291:274-81.

LYAMICHEV, V.I., KAISER, M.W., LYAMICHEVA, N.E., VOLOGODSKII, A.V., HALL, J.G., MA, W.-P., ALLAWI, H.T., AND NERI, B.P. 2000. Experimental and theoretical analysis of the invasive signal amplification reaction. Biochemistry 39:9523-9532.

MADDISON, W.P. 1997. Gene trees in species trees. Systematic Biology 46:523 536.

MANEL, S., GAGGIOTTI, O.E., AND WAPLES, R.S. 2005. Assignment methods: matching biological questions with appropriate techniques. Trends in Ecology and Evolution 20:136-142.

MANIATIS, T., FRISTCH, E.F. AND SAMBROOK, J. 1982. Molecular cloning; a laboratory manual. Cold Spring Harbor Publication, Cold Spring Harbor, New York.

MARTH, G., YEH, R., MINTON, M., DONALDSON, R., LUI, Q., DUAN, S., DAVENPORT, R., MILLER, R.D., KWOK, P-Y. 2001. Single-nucleotide polymorphisms in the public domain: how useful are they? Nature Genetics 27:371-372.

MATTHEE, C.A., BURZLAFF, J.D., TAYLOR, J.F., AND DAVIS, S.K. 2001. Mining the mammalian genome for artiodactyl systematics. Systematic Biology 50:367-90.

McCARTER, J., ABAD, P., JONES, J.T., AND BIRD, D. 2000. Rapid gene discovery in plant parasitic nematodes via Expressed Sequence Tags. Nematology 2:719-731.

MELLO, C.C. AND CONTE, D. 2004. Revealing the world of RNA interference. Nature 431:338-342.

METZGAR, D., BYTOF, J., WILLS, C. 2000. Selection against frameshift mutations limits microsatellite expansion in coding DNA. Genome Research 10:72-80.

MILLER, L.M. 2000. Classifying genealogical origins in hybrid populations using dominant markers. Journal of Heredity 91:46–49.

MIYAKE, T. AND AMEMIYA, C.T. 2004. BAC libraries and comparative genomics of aquatic chordate species. Comparative Biochemistry and Physiology, Part C 138: 233–244.

MIYAMOTO, M.M. 1999. Perfect SINEs of evolutionary history? Current Biology 9:R816-819.

MOCK, K.E., THEIMER, T.C., RHODES, O.E., GREENBERG, D.L., AND KEIM, P. 2002. Genetic variation across the historical range of the wild turkey (*Meleagris gallopavo*). Molecular Ecology 11:643–657.

MOORE, W.S. 1995. Inferring phylogenies from mtDNA variation:mitochondrial gene trees versus nuclear-gene trees. Evolution 49:718-726.

MORANDO, M., AVILA, L.J., BAKER, J., SITES, J.W., JR. 2004. Phylogeny and phylogeography of the *Liolaemus darwinii* complex (Squamata: Liolaemidae): Evidence for introgression and incomplete lineage sorting. Evolution 58:842–861.

MORIN, P.A., LUIKART, G., WAYNE, R.K., AND THE SNP WORKSHOP GROUP. 2004. SNPs in ecology, evolution, and conservation. Trends in Ecology and Evolution 19:208–216.

MORITZ, C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology 51:238-254.

MORRIS, K.V., CHAN, S.W., JACOBSEN, S.E., LOONEY, D.J. 2004. Small interfering RNA-induced transcriptional gene silencing in human cells. Science 305:1289-92.

MUELLER, U.G. AND WOLFENBARGER, L.L. 1999. AFLP genotyping and fingerprinting. Trends Ecology and Evolution. 14:389–395.

MURDOCK, C. AND WIBBELS, T. 2003a. Expression of Dmrt1 in a turtle with temperature-dependent sex determination. Cytogenetics and Genome Research 10:302-8.

MURDOCK, C. AND WIBBELS, T. 2003b. Cloning and expression of aromatase in a turtle with temperature-dependent sex determination. General Comparative Endocrinology 130:109-19.

NEAR, T., MEYLAN, P.A., AND SHAFFER, H.B. 2005. Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. American Naturalist 165:137-146.

NICHOLS, R.A. AND FREEMAN, K.L.M. 2004. Using molecular markers with high mutation rates to obtain estimates of relative population size and to distinguish the effects of gene flow and mutation: a demonstration using data from endemic Mauritian skinks. Molecular Ecology 13:775-787.

NICHOLSON, G., SMITH, A.V., JONSSON, F., GÚSTAFSSON, O., STEFÁNSSON, K., DONNELLY, P. 2002. Assessing population differentiation and isolation from single- nucleotide polymorphism data. Journal of the Royal Statistical Society Series B-Statistical Methodology 64:695-715.

NICOD, J.C., AND LARGIADER, C.R. 2003. SNPs by AFLP (SBA):a rapid SNP isolation strategy for non-model organisms. Nucleic Acids Research 31:e19.

NIELSEN, R. 2000. Estimation of population parameters and recombination rates from single nucleotide polymorphisms. Genetics 154: 931-942.

NIELSEN, R. 2005. Molecular signatures of natural selection. Annual Reviews of Genetics 39:197-218.

NOONAN, J. P., GRIMWOOD, J. DANKE, J. SCHMUTZ, J., DICKSON, M., AMEMIYA, C.T. AND MYERS, R.M. 2004. Coelacanth genome sequence reveals the evolutionary history of vertebrate genes. Genome Research 14:2397-2405.

OHYA, Y.K., KURAKU, S., AND KURATANI, S. 2005. Hox code in embryos of Chinese soft-shelled turtle *Pelodiscus sinensis* correlates with the evolutionary innovation in the turtle. Journal of Experimental Zoology. Part B. Molecular Developmental Evolution 304:107-118.

PAGEL, M. AND MEADE, A. 2004. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. Systematic Biology 53:571 – 581.

PARHAM, J.F., FELDMAN, C.R, AND BOORE, J. L. 2006a. The complete mitochondrial genome of the enigmatic bigheaded turtle (*Platysternon*): Description of unusual genomic features and the reconciliation of phylogenetic hypotheses based on mitochondrial

and nuclear DNA. BioMed Central Evolutionary Biology 6:11.

Parham, J.F., Macey, J.R., Papenfuss, T.J., Feldman, C.R., Türkozan, O., Polymeni, R. and J.L. Boore. 2006b. The phylogeny of Mediterranean tortoises and their close relatives (*Testudona* tax. nov.) based on complete mitochondrial genome sequences from museum specimens: with comments on mitochondrial genomic features, phylogenetic taxonomy, and paleobiogeography. Molecular Phylogenetics Evolution 38:50-64.

Parham, J.F., Simison, W.B., Kozak, K.H., Feldman, C.R., and Shi, H. 2001. New Chinese turtles: endangered or invalid? A reassessment of two species using mitochondrial DNA, allozyme electrophoresis and known-locality specimens. Animal Conservation 4:357-367.

Pearse, D.E. and Avise, J.C. 2001. Turtle mating systems: behavior, sperm storage, and genetic paternity. Journal of Heredity 92:206-211.

Pearse, D.E. and Crandall, K.A. 2004. Beyond F_{ST}: population genetic analysis for conservation. Conservation Genetics 15:585-602.

Pearse, D.E., Arndt, A.D., Valenzuela, N., Miller, B.A., Cantarelli, V.H., Sites, J.W. Jr. 2006a. Estimating population structure under non-equilibrium conditions in a conservation context: continent-wide population genetics of the giant Amazon river turtle *Podocnemis expansa* (Chelonia: Podocnemididae). Molecular Ecology 15:985-1006.

Pearse, D.E., Dastrup, R.B., Hernandez, O. and Sites, J.W. Jr. 2006b. Paternity in an Orinoco population of endangered Arrau river turtles, *Podocnemis expansa* (Pleurodira; Podocnemididae), from Venezuela. Chelonian Conservation and Biology 5:232-238.

Pearse, D.E., Eckerman, C.M., Janzen, F.J., and Avise, J.C. 2001. A genetic analogue of mark-recapture methods for estimating local population size: an approach based on molecular parentage assessments. Molecular Ecology 10:2711-2718.

Pearse, D.E., Janzen, F.J. and Avise, J.C. 2002. Multiple paternity, sperm storage, and reproductive success of female and male painted turtles (*Chrysemys picta*) in nature. Behavioral Ecology and Sociobiology 51:164-171.

Pinhasov, A., Mei, J., Amaratunga, D. Amato, F.A., Lu, H., Kauffman, J., Xin, H. Brenneman, D.E., Johnson, D.L., Andrade-Gordon, P., Ilyin, S.E., 2004. Gene expression analysis for high throughput screening applications. Combinatorial Chemistry and High Throughput Screening Applications 7:133-140.

Place, A. R. and V. A. Lance. 2004. The temperature-dependent sex determination drama: same cast, different stars. In: Temperature-dependent sex determination in vertebrates. Eds. N. Valenzuela and V. Lance. Smithsonian Institution.

Place, A.R., Lang, J., Gavasso, S., and Jeyasuria, P. 2001. Expression of P450(arom) in *Malaclemys terrapin* and *Chelydra serpentina*: a tale of two sites. Journal of Experimental Zoology 290:673-90.

Platz, J.E. and Conlon, J.M. 1997. ...and back again. Nature 389:246.

Pollock, D. D., Eisen, J.A., Doggett, N.A., and Cummings, M.P. 2000. A case for evolutionary genomics and the comprehensive examination of sequence biodiversity. Molecular Biology and Evolution17:1776-1858.

Posada, D. and Crandall, K.A. 2002. The effect of recombination on the accuracy of phylogeny estimation. Journal of Molecular Evolution 54:396-402.

Primmer, C.R., Borge, T., Lindell, J., Saetre, G.P. 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. Molecular Ecology 11:603-612.

Ray, D.A., Xing, J., Hedges, D.J., Hall, M.A., Laborde, M.E., Anders, B.A., White, B.R., Stoilova, N., Fowlkes, J.D., Landry, K.E., Chemnick, L.G., Ryder, O.A. and Batzer, M.A. 2005. *Alu* insertion loci and platyrrhine primate phylogeny. Molecular Phylogenetics and Evolution 35:117-126.

Renn, S.C.P., Aubin-Horth, N. and Hofmann, H.A. 2004. Biologically meaningful expression profiling across species using heterologous hybridization to a cDNA microarray. BMC Genomics 5: 1–13.

Roman, J. and Bowen, B.W. 2000. The mock turtle syndrome: genetic identification of turtle meat purchased in the southeast United States. Animal Conservation 3:61-65.

Ronaghi, M. 2001. Pyrosequencing sheds light on DNA sequencing. Genome Research 11: 3-11.

Roy-Engel, A. M. 2002. Non-traditional *Alu* evolution and primate genomic diversity. Journal of Molecular Biology 316:1033-1040.

Saetre, G.P., Borge, T., Lindell, J., Moum, T., Primmer, C.R., Sheldon, B.C., Haavie, J., Johnsen, A. and Ellegren, H. 2001. Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. Molecular Ecology 10:737-749.

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

Sarkar, N. and Kashyap, V.K. 2003. SNP markers in forensic testing- a preliminary study. Forensic Science International 136:52-53.

Sasaki, T., Takahashi, K., Nikaido, M., Miura, S., Yasukawa, Y. and Okada, N. 2004. First application of the SINE (Short Interspersed Repetitive Element) method to infer phylogenetic relationships in reptiles: An example from the turtle superfamily Testudinoidea. Molecular Biology and Evolution 21:705-715.

Sanger, F., Micklen, S., and Coulson, A.R. 1977. DNA sequencing and chain-terminating inhibitors. Proceedings of the National Academy of Sciences 74:5463-5467.

Schlötterer, C. 2002. Towards a molecular characterization of adaptation in local populations. Current Opinion in Genetics and Development 12:683–687.

Schlötterer, C., and Harr, B. 2002. Single nucleotide polymorphisms derived from ancestral populations show no evidence for biased diversity estimates in *Drosophila melanogaster*. Molecular Ecology 11:947-950.

Seddon, J.M., Parker, H.G., Ostrander, E.A., Ellegren, H. 2005. SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. Molecular Ecology 14:503-511.

Shedlock, A.M. and Okada, N. 2000. SINE insertions: powerful tools for molecular systematics. Bioessays 22:148-160.

Shedlock, A.M., Takahashi, K. and Okada, N. 2004. SINEs of speciation: Tracking lineages with retroposons. Trends in Ecology and Evolution 19:545-553.

Shepherd, M., Cross, M., Maguire, T., Dieters, M., Williams, C., Henry, R. 2002. Transpecific microsatellites for hard pines. Theoretical and Applied Genetics 104:819–827.

Silva, J.M., Li, M.Z., Chang, K., Ge, W., Golding, M.C., Rickles, R.J., Siolas, D., Hu, G., Paddison, P.J., Schlabach, M.R., Sheth, N., Bradshaw, J., Burchard, J., Kulkarni, A., Cavet, G., Sachidanandam, R., McCombie, W.R., Cleary, M.A., Elledge, S.J., and Hannon, G.J. 2005. Second-generation shRNA libraries covering the mouse and human genomes. Nature Genetics 37: 1281-1288.

Sites, J.W., Jr., FitzSimmons, N.N., da Silva, N.J., and V.H. Cantarelli. 1999. Population genetic structure in the giant Amazon river turtle (*Podocnemis expansa*): inferences from two classes of molecular markers. Chelonian Conservation Biology 3:454-463.

Sites, J.W. and Marshall, J.C. 2003. Delimiting species: a Renaissance issue in systematic biology. Trends in Ecology and Evolution 18:462-470.

Sites, J.W. and Marshall, J.C. 2004. Operational criteria for delimiting species. Annual Review of Ecology, Evolution, and Systematics. 35: 199-227.

Slate, J., Visscher, P.M., MacGregor, S., Stevens, D., Tate, M.L.,

PEMBERTON, J.M. 2002. A genome scan for quantitative trait loci in a wild population of red deer (*Cervus elaphus*). Genetics 162: 1863-1873.

SLATE, J. 2005 Quantitative trait locus mapping in natural populations: progress, caveats and future directions. Molecular Ecology 14: 363-379.

SMITH, C.A., AND SINCLAIR, A.H. 2004. Sex determination: insights from the chicken. BioEssays 36: 120-132.

SMITH, C.T., TEMPLIN, W.D., SEEB, J.E., SEEB, U.W. 2005. Single nucleotide polymorphisms provide rapid and accurate estimates of the proportions of US and Canadian Chinook salmon caught in Yukon River fisheries. North American Journal of Fisheries Management 25:944-953.

SONG, L.N., YANG, Y.L., AND XU, R.B. 1991. Stress-induced regulation of glucocorticoid receptor gene expression in the rat. Receptor 1:207-16.

SPINKS, P.Q., AND SHAFFER, H.B. 2005. Range-wide molecular analysis of the western pond turtle (*Emys marmorata*): cryptic variation, isolation by distance, and their conservation implications. Molecular Ecology 14:2047-2064.

SPINKS, P.Q., AND SHAFFER, H.B. 2007 Conservation phylogenetics of the Asian box turtles (Geoemydidae, *Cuora*): mitochondrial introgression, numts, and inferences from multiple nuclear loci. Conservation Genetics 8:641-657.

SPINKS, P.Q., SHAFFER, H.B., IVERSON, J.B., McCORD, W.P. 2004. Phylogenetic hypotheses for the turtle family Geoemydidae. Molecular Phylogenetics and Evolution 32:164-182

SPOTILA, L.D., SPOTILA, J.R., AND HALL, S.E.1998. Sequence and expression analysis of WT1 and Sox9 in the red-eared slider turtle, *Trachemys scripta*. Journal of Experimental Zoology 281:417-27.

STARKEY, D.E., SHAFFER, H.B., BURKE, R.L., FORSTNER, M.R.J., IVERSON, J.B., JANZEN, F.J., RHODIN, A.G.J., ULTSCH, G.R. 2003. Molecular systematics, phylogeography, and the effects of Pleistocene glaciation in the painted turtle (*Chrysemys picta*) complex. Evolution 57:119-128.

STEPHENS, P.R., AND WIENS, J.J. 2003. Explaining species richness from continents to communities: the time-for-speciation effect in emydid turtles. American Naturalist 161:112-128.

STEPHENS, P.R., AND WIENS, J.J. 2004. Convergence, divergence, and homogenization in the ecological structure of emydid turtle communities: the effects of phylogeny and dispersal. American Naturalist 164:244-254.

STICKNEY, H.L., SCHMUTZ, J., WOODS, I.G., HOLTZER, C.C., DICKSON, M.C., KELLY, P.D., MYERS, R.M. AND TALBOT, W.S. 2002. Rapid mapping of zebrafish mutations with SNPs and oligonucleotide microarrays. Genome Research 12: 1929-1934.

STOREY, K. B. 2005. Reptile freeze tolerance: metabolism and gene expression. Cryobiology 52:1-16.

STOUGHTON, R.B. 2005. Applications of DNA microarrays in biology. Annual Review of Biochemistry 74:53–82.

STUART, B.L. AND PARHAM, J.F. 2004. Molecular phylogeny of the critically endangered Indochinese box turtle (*Cuora galbinifrons*). Molecular Phylogenetics and Evolution 31:164-77.

TAKAHASHI, R., ITO, K., FUJIWARA, Y., KODAIRA, K., HIRABAYASHI, M., AND UEDA, M. 2000. Generation of transgenic rats with YACs and BACs: preparation procedures and integrity of microinjected DNA. Experimental Animals 49:229-233.

TEMPLETON, AR. 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. Molecular Ecology 13:789–809.

THOMAS, J.W., TOUCHMAN, J.W., BLAKESLEY, R.W., BOUFFARD, G.G., BECKSTROM-STERNBERG, S.M. MARGULIES, E.H., BLANCHETTE, M., SIEPEL, A.C., THOMAS, P.J., McDOWELL, J.C., MASKERI, B., HANSEN, N.F., SCHWARTZ, M.S., WEBER, R.J., KENT, W.J., KAROLCHIK, D., BRUEN, T.C., BEVAN, R., CUTLER, D.J., SCHWARTZ, S., ELNITSKI, L., IDOL, J.R., PRASAD, A.B., LEE-LIN, S.-Q., MADURO, V.V.B., SUMMERS, T.J., PORTNOY, M.E., DIETRICH, N.L., AKHTER, N., AYELE, K., BENJAMIN, B., CARIAGA, K., BRINKLEY, C.P., BROOKS, S.Y., GRANITE, S., GUAN, X., GUPTA, J., HAGHIGHI, P., HO, S.-L., HUANG, M.C., KARLINS, E., LARIC, P.L., LEGASPI, R., LIM, M.J., MADURO, Q.L., MASIELLO, C.A., MASTRIAN, S.D., McCLOSKEY, J.C., PEARSON, R., STANTRIPOP, S., TIONGSON, E.E., TRAN, J.T., TSURGEON, C., VOGT, J.L., WALKER, M.A., WETHERBY, K.D., WIGGINS, L.S., YOUNG, A.C., ZHANG, L.-H., OSOEGAWA, K., ZHU, B., ZHAO, B. SHU, C.L., DE JONG, P.J., LAWRENCE, C.E., SMIT, A.F., CHAKRAVARTI, A., HAUSSLER, D., GREEN, P., MILLER, W. AND GREEN, E.D. 2003. Comparative analyses of multi-species sequences from targeted genomic regions. Nature 424:788-793.

TON, C., HWANG, D.M., DEMPSEY, A.A., TANG, H.-C., YOON, J., LIM, M., MABLY, J.D., FISHMAN, M.C., LIEW, C.-C. 2000. Identification, characterization, and mapping of expressed sequence tags from an embryonic zebrafish heart cDNA library. Genome Research 10: 1915-1927.

TURGEON, J. AND BERNATCHEZ, L. 2003. Reticulate evolution and phenotypic diversity in North American ciscoes, *Coregonus* ssp. (Teleostei: Salmonidae):implications for the conservation of an evolutionary legacy. Conservation Genetics 4:67-81.

ULRICH, C. M., BIGLER, J., VELICER, C.M., GREENE, E.A., FARIN, F.M., AND POTTER, J.D. 2000. Searching Expressed Sequence Tag databases: discovery and confirmation of a common polymorphism in the thymidylate synthase gene. Cancer Epidemiology, Biomarkers and Prevention 9:1381–1385.

VALENZUELA, N. 2000. Multiple paternity in sideneck turtles *Podocnemis expansa*: evidence from microsatellite DNA data. Molecular Ecology 9:99-106.

VALENZUELA, N. AND SHIKANO, T. 2007. Embryological ontogeny of *Aromatase* gene expression in *Chrysemys picta* and *Apalone mutica* turtles: comparative patterns within and across temperature-dependent and genotypic sex-determining mechanisms. Development, Genes and Evolution 217:55-62.

VALENZUELA, N., LeCLERE, A., AND SHIKANO, T., 2006. Comparative gene expression of steroidogenic factor 1 in *Chrysemys picta* and *Apalone mutica* turtles with temperature-dependent and genotypic sex determination. Evolution and Development 8:424-432.

VASEMÄGI, A., NILSSON, J., AND PRIMMER, C.R. 2005. Expressed Sequence Tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar L.*). Molecular Biology and Evolution 22:1067–1076.

VEKEMANS, X., BEAUWENS, T., LEMAIRE, M., AND ROLDAN-RUIZ, I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. Molecular Ecology 11:139–151.

VIGNAL, A., MILAN, D., SAN CRISTOBAL, M., AND EGGEN, A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution 34: 275-305.

VINCENT, C., BONTOUX, M., LE DOUARIN, N.M., PIEAU, C., MONSORO-BURQ, A.H.2003. Msx genes are expressed in the carapacial ridge of turtle shell: a study of the European pond turtle, *Emys orbicularis*. Development, Genes, and Evolution 213:464-9.

VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., VANDELEE, T., HORNES, M., FRITJERS, A., POT, J. PELEMAN, J., KUIPER, M. AND ZABEAU, M. 1995. AFLP- a new technique for DNA-fingerprinting. Nucleic Acids Research 23:4407-4414.

WALDICK R.C., KRAUS S., BROWN M., AND WHITE B.N. 2002. Evaluating the effects of historic bottleneck events: an assessment of microsatellite variability in the endangered North Atlantic right whale. Molecular Ecology 11:2241-2249.

Wakeley, J., Nielsen, R., Liu-Cordero, S.N., and Ardlie, K. 2001. The discovery of single-nucleotide polymorphisms-and inferences about human demographic history. American Journal of Human Genetics 69: 1332-1347.

Watkins, W.S., Rogers, A.R., Ostler, C.T., Wooding, S., Bamshad, M.J., Brassington, A.-M.E., Carroll, M.L., Nguyen, S.V., Walker, J.A., Prasad, B.V.R., Reddy, P.G., Das, P.K., Batzer, M.A., Jorde, L.B. 2003. Genetic variation among world populations: inferences from 100 *Alu* insertion polymorphisms. Genome Research 13:1607-1618.

Weaver, R.F. (Ed.). 2005. Molecular Biology. McGraw Hill. New York, NY.

Weiner, A. M. 2002. SINEs and LINEs: the art of biting the hand that feeds you. Current Opinion in Cell Biology 14: 343-350.

Weiner, A., Deininger, P. L. and Efstratuadis, A. 1986. Non-viral retrotransposons:Genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annual Review of Biochemistry 55:631-661.

Weinig, C. and Schmitt, J. 2004 Environmental effects on the expression of quantitative trait losi and implications for phenotypic evolution. Bioscience 54: 627-635.

Whitfield, C.W., Cziko, A.-M., Robinson, G.E. 2003.Gene expression profiles in the brain predict behavior in individual honey bees. Science 302:296-299.

Wienholds, E. and Plasterk, R.H.A. 2005. MicroRNA function in animal development. Febs Letters 579: 5911-5922.

Wilding, C.S., Butlin, R.K., and Grahame, J. 2001. Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. Journal of Evolutionary Biology 14:611–619.

Wolfsberg, T.G. and Landsman, D. 1997. A comparison of expressed sequence tags (ESTs) to human genomic sequences. Nucleic Acids Research 25:1626-1632.

Wong, M.L. and Medrano, J.F. 2005. Real-time PCR for mRNA quantification. Biotechniques 39:1-11.

Wright, T.F., P.M. Johns, J.R. Walters, Lerner, A.P., Swallow, J.G., and Wilkson, G.S. 2004. Microsatellite variation among divergent populations of stalk-eyed flies, genus *Cyrtodiopsis*. Genetical Research 84:27-40.

Wurmbach, E., Yuen, T., and Sealfon, S.C. 2003. Focused microarray analysis. Methods 31: 306–316.

Yang, Z. 1996. Maximum likelihood models for combined analyses of multiple sequence data. Journal of Molecular Evolution 42:587-596.

Zane, L., Bargelloni, L., and Patanello, T. 2002. Strategies for microsatellite isolation: a review. Molecular Ecology 11:1–16.

Zardoya, R. and Meyer, A. 1996. Phylogenetic performance of mitochondrial genes in resolving relationships among vertebrates. Molecular Biology Evolution 13:933-942.

Zardoya, R. and Meyer, A. 1998. Complete mitochondrial genome suggests diapsid affinities of turtles. Proceedings of the National Academy of Sciences 95:14226–14231.

Zenger, K.R., Eldridge, M.D.B., Pope, L.C. and Cooper, D.W. 2003. Characterization and cross-species utility of microsatellite markers within kangaroos, wallabies and rat kangaroos (Macropodoidea: Marsupialia). Australian Journal of Zoology 51: 587–96.

Zhang, D.X. and Hewitt, G.M.1996. Nuclear integrations: challenges for mitochondrial DNA markers. Trends in Ecology and Evolution 11: 247-251.

Zhang, D.X. and Hewitt, G.M. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems, and prospects. Molecular Ecology 12:563-584.

Zhao, Z.M., Fu, Y.X., Hewett-Emmett, D., and Boerwinkle, E. 2003. Investigating single nucleotide polymorphism (SNP) density in the human genome and its implications for molecular evolution. Gene 312:207-213.

Glossary

AFLP – A genome–wide screen for dominant molecular markers through restriction enzyme digests, followed by selective PCR.

Allele – Different forms of the same gene (AA, aa), or if at a noncoding locus, this can refer to variation in DNA sequence.

Ascertainment bias – Systematic distortion in measuring the true frequency of a phenomenon due to the way in which the data are obtained. One example is illustrated by the empirical observation that microsatellite alleles found in a focal species may not amplify as well in related species and result in null alleles.

Autapomorphic – A derived characteristic exclusive to a given taxon or monophyletic group.

BAC – **B**acterial **A**rtificial **C**hromosome, an *E. coli* plasmid used as a vector to hold large inserts (up to 300,000 bp) of foreign DNA.

cDNA – A DNA copy complementary to a mRNA sequence made by the enzyme reverse transcriptase.

Cloning – A technique which refers to one of two things: 1) inserting a gene from one organism into another via a vector for propagation and investigation, or 2) identification of the location and sequence of a gene which is correlated with a certain phenotype.

Codominant marker – A locus whose alleles are co–dominant, i.e., the genotype of heterozygotes is readily recognizable from that of homozygotes.

Codon – A set of three nucleotides that specifies either termination of translation or a certain amino acid to be incorporated into a growing polypeptide (protein) during translation.

Demography – The study of size, structure, and distribution of populations, and their change over time due to births, deaths, migration, and ageing.

DNA – The material from which genes are made; deoxyribonucleotides linked with phophodiester bonds.

DNA fingerprinting – The use of multiple markers that provide unique DNA profiles for individual identification.

Dominant marker – A gene whose alleles are dominant, i.e., the genotype of heterozygotes is indistinguishable from that of the dominant homozygotes thus impeding the estimation of the heterozygote frequency.

Duplication – Doubling of a DNA sequence such as a dinucleotide repeat within a microsatellite, or as much as an entire gene, chromosome, or genome.

Electromorph – An allele identified by its unique mobility through gel electrophoresis, due to the specific molecular weight and conformation of the allele (e.g., DNA fragment, isozyme).

EST – **E**xpressed **S**equence **T**ags are a short cDNA sequence from one end of an expressed gene used to fish a gene out of the chromosomal DNA by matching base pairs.

Euchromatin – The less condensed part of the chromatin, as compared to heterochromatin; located away from the centromeres and telomeres of chromosomes.

Evo-Devo – A relatively new field called evolutionary developmental biology which takes a comparative look at the genetics behind developing organisms across all taxonomic levels.

Evolutionarily Significant Unit (ESU) or Management Unit (MU) – A group which has reciprocal monophyly in a mitochondrial marker and divergent allele frequencies at a nuclear marker; this designation should be assessed by genealogical concordance within and across genes within the species. This term should designate populations, species, or subspecies considered to have an independent evolutionary legacy. The definition of a manage-

ment unit is similar but does not typically require a large phylogenetic distance and instead only requires that the alleles frequencies be diverging.

FISH – **F**luorescent **in S**itu **H**ybridization, a technique of hybridizing a fluorescently labeled DNA probe to whole chromosomes to determine the physical location of that marker.

Frame shift – A mutation that causes the reading frame of the codons to change; most commonly indels of 1 or 2 bases.

F_{ST} – A genetic measure of population subdivision that describes the variation in allele frequencies among different populations; typically an F_{ST} value of 0.25 is taken as evidence of substantial population differentiation.

Gene – A segment of DNA which performs a specific function such as coding for a protein, specifying a functional RNA molecule, or regulating other functions as in the case of DNA replication, chromosome segregation, or maintenance of chromosome integrity.

Gene tree – Contained within a species tree, it represents a branching pattern of evolution as the gene is passed on to more than one progeny per generation. Processes such as horizontal transfer, deep coalescence, and gene duplication or extinction can result in discordance between gene trees and species trees.

Genome – The complete genetic information contained in an organism.

Genotype – The particular allelic combinations found at a specific locus or loci of an individual (i.e., AA, Aa, aa).

Heterozygous – An individual with two different alleles for the same gene.

Homology – Sharing of characters because of their common ancestry.

Homoplasy – Characters that evolved more than once (e.g., as by convergent evolution) and were not present in the most recent common ancestor of the species sharing them.

Homozygous – An individual carrying two identical alleles of a given gene.

Indel – An insertion or deletion of nucleotides in a DNA segment.

Intron – A segment of noncoding DNA that separates coding parts (exons) within a gene.

Isochore – A region of genomic DNA sequence in which G+C compositions are relatively uniform.

Karyotype – The total set of all chromosomes of a cell of any living organism, displayed in pairs, and arranged by size, such that chromosomal aberrations and sex can be detected.

Locus – A delimited section chromosome housing a particular gene or other marker.

Marker – A gene, mutation, or other sequence that serves as an indicator of a known location in the genome.

Mendelian – Markers that are inherited under Mendel's laws of equal, random segregation and independent assortment during gamete production; examples include autosomal dominant, autosomal recessive, and sex–linked recessive and dominant genes.

Microarray – DNA sequences spotted on a microscope slide to which a labeled DNA pool of interest is hybridized in search for matching sequences.

Microsatellite – A DNA motif (2-6 bp long) repeated many times in tandem.

mtDNA – DNA of the mitochondria, typically about 16.5 kilobases (kb) for the entire genome. In animals, sequence evolution occurs more quickly than in most nuclear DNA. One exception includes nuclear microsatellites.

Neutral processes – Genetic processes which are not governed by selection (i.e., most commonly random genetic drift and random mutation).

Numt – Transferred pieces of mtDNA to nuclear chromosomal regions.

Ortholog – Homologous sequences where sequence divergence follow speciation.

PCR – **P**olymerase **C**hain **R**eaction is the exponential increase of DNA fragments *in vitro* using an enzyme (polymerase) that copies the DNA in between primers annealed to the flanking regions of the desired sequence.

Paralog – Homologous sequences that have arisen by a duplication event (i.e., hemoglobin and myoglobin). Each of the two duplicates are then on different evolutionary trajectories and are no longer comparable for phylogenetic analysis.

Phylogeny – The evolutionary relationships of groups of organisms, typically arranged in a branching diagram.

Phylogeography – The study of the patterns and processes responsible for the geographic distribution of genealogical lineages, particularly closely related species.

Plasmid – A double stranded piece of DNA that is separate from the chromosomal DNA; typically circular, ranging from 1–400 kb, and varying from one copy to several hundreds of copies in the cell.

Primer – An RNA or DNA fragment about 20 bp long that supplies the initial free end needed for DNA replication.

Pseudogene – A previously active gene which has accumulated a series of inactivating mutations.

QTL analysis – **Q**uantitative **T**rait **L**oci analysis, a statistical way to estimate the potential location on the genome coding for a complex or quantitative trait (i.e., height).

Restriction sites – A DNA sequence that is recognized by restriction enzymes which then cut the DNA molecule at or near that sequence.

RNA – A copy of DNA made into a polymer of ribonucleotides linked by phosphodiester bonds.

RT-PCR – **R**everse **T**ranscription-**P**olymerase **C**hain **R**eaction is a technique in which an RNA strand is reverse transcribed into its DNA complement, followed by amplification of the resulting DNA by PCR. **Real-Time PCR** – A PCR method capable of tracking the amount of amplified DNA produced at each cycle with the use of fluorescent dyes, thus allowing the quantification of the initial template (also called **Q**uantitative PCR or RT-PCR [QPCR or QRT-PCR]).

Saturation – Multiple nucleotide substitutions at a site that erase phylogenetic signal because conserved nucleotides cannot be distinguished from nucleotide sites that have independently mutated back to the same state (creating homoplasy).

SINE and LINE – Retrotransposons with utility as phylogenetic markers. SINEs (Short Interspersed Nuclear Elements) are nonautonomous, while LINEs (Long Interspersed Nuclear Elements) are autonomous (i.e., they can support their own transposition).

SNP – **S**ingle **N**ucleotide **P**olymorphism, a single nucleotide difference between two or more individuals at a particular locus.

Species tree – A phylogenetic tree representing the branching pattern among species lineages.

Transcriptome – The total set of mRNA transcripts produced in an individual at any given time.

Translocation – Movement of a section of DNA from its current location in a chromosome to a different chromosome.

Transposon – Sequences of DNA that can move around to different positions within the genome of a single cell and, in the process, may cause mutations and change the amount of DNA in a genome. They are also called jumping genes or mobile genetic elements.

Vector – A small DNA construct used in cloning, capable of carrying a foreign DNA fragment of interest into a host cell (such as *E. coli* bacteria) and facilitating its replication in that cell.