CHAPTER 3

Genetic concepts and tools to support wildlife population biology

> *For those not studying biology at the time in the early 1950s, it is hard to imagine the impact the discovery of the structure of DNA had on our perception of how the world works. Reaching beyond the transformation of genetics, it injected into all of biology a new faith in reductionism. The most complex of processes, the discovery implied, might be simpler than we had thought. It whispered ambition and boldness to young biologists and counseled them: Try now: strike fast and deep at the secrets of life.*
>
> **E.O. Wilson (1995), *Naturalist***

INTRODUCTION

Until fairly recently the term "genetics" would have hardly been mentioned in a book on applied wildlife population biology. Now, however, genetic issues related to wildlife populations seem to paint the newspapers almost every day. Is the red wolf taxonomically distinct enough to warrant special management? Are northern hairy-nosed wombats genetically impoverished? Are red-cockaded woodpeckers suffering low survival or reproductive rates due to inbreeding? Does a sample from meat in a freezer match that of an illegally harvested deer?

Because genetic concepts and tools are increasingly at the heart of many wildlife management issues, a variety of genetic applications will be considered throughout this book. This chapter will explain what genetic variation is and how it can be described using several common genetic markers, and will describe a few of the insights into wildlife populations that can be gained by genetic analysis. In short, this chapter will build the foundation for applying genetic approaches to other topics throughout the book.

WHAT IS GENETIC VARIATION?

The **phenotypic expression** of almost all individual traits (ranging from body weight to camouflage pattern to sprint speed to metabolic efficiency) is a function of the genetic makeup, or **genotype**, coupled with the environment to which the individual is exposed.[1] The **genes** that make up the genotype are stretches of DNA along chromosomes in the nucleus of all cells; the location of a gene on a chromosome is called a **locus** (or **loci** for >1 locus). Different forms of a gene, known as **alleles**, vary in the specific sequences of DNA **nucleotides** (adenine (A), guanine (G), thymine (T), and cytosine (C)). Genes are **transcribed** by messenger RNA (mRNA) to provide the template for building amino acid chains that comprise a protein product

[1] Actually, it turns out that heritable variation in phenotypes can also arise from **epigenetic** variation, whereby molecular processes activate or disable activity of particular genes that may be inherited by future generations. It is unknown how much variation in phenotypes is explained by epigenetic versus genetic variation (Bossdorf et al. 2008).

Conservation of Wildlife Populations: Demography, Genetics, and Management, Second Edition. L. Scott Mills.
© 2013 John Wiley & Sons, Ltd. Published 2013 by John Wiley & Sons, Ltd.

Box 3.1 The Hardy–Weinberg principle and describing heterozygosity

The Hardy–Weinberg (HW) principle forms the cornerstone of population and conservation genetics. The idea behind HW is that allele and genotype frequencies remain constant over time at equilibrium; that is, if they are unaffected by evolutionary forces such as natural selection, genetic drift, mutation, and gene flow. Populations are of course affected by these processes, but HW equilibrium genotype frequencies provide a tremendously useful benchmark; if the population is out of HW equilibrium we can ask why, thereby taking the first step toward elucidating mechanisms acting on a population's genetic composition.

Here is how to determine the HW frequencies. If two alleles (A_1 and A_2) at a locus have frequency p and q respectively, then after one generation of random mating the frequencies of the three possible genotypes (A_1A_1, A_1A_2, and A_2A_2) are p^2, $2pq$, and q^2 respectively. Since these are the only possible genotypes with just two alleles, then $p^2 + 2pq + q^2 = 1$.

Extending this idea to multiple alleles, the Hardy–Weinberg frequency of the homozygote genotype for any allele i is p_i^2. Because an individual that is not homozygous must be heterozygous at a locus, the expected Hardy–Weinberg frequency of heterozygotes, given k alleles at a locus, is

$$1-\sum_{i=1}^{k} p_i^2$$

Heterozygosity at the population level is typically described as expected or observed. Expected heterozygosity is that expected under Hardy–Weinberg equilibrium. By contrast, observed heterozygosity is the actual proportion of individuals observed to be heterozygous, averaged across loci. Deviations between observed and expected heterozygosity can be useful for inferring processes that are acting upon wildlife populations, such as genetic drift, selection, and gene flow.

Let us work through a simple example for just one locus based on data from the endangered Hawaiian Laysan finch (Tarr et al. 1998, Frankham et al. 2010); real studies would use multiple loci. At this one locus there are three alleles, with the following frequencies: $p_1 = 0.364$, $p_2 = 0.352$, and $p_3 = 0.284$. (Notice that the three allele frequencies sum to 1.0.) Using the Hardy–Weinberg principle, the expected heterozygosity is

$$1-\sum_{i=1}^{k} p_i^2 = 1-\left(0.364^2 + 0.352^2 + 0.284^2\right) = 0.663$$

In this case, 29/44 sampled finches are heterozygotes, so observed heterozygosity at this locus (0.659) was very close to expected heterozygosity (0.663).

(phenotype). Sometimes a phenotype is a **single-gene trait**, determined from combinations of alleles at just one gene – for example, whether or not you can roll your tongue or the dark coat color in the red fox (Våge et al. 1997) – but most often traits are determined by complex combinations of genes at many loci, often interacting with the environment.

Most vertebrates are diploid, meaning that each individual contains two sets of chromosomes and offspring inherit one allele at each locus from each parent (a few amphibian, lizard, and fish species are polyploid, where each individual contains >2 sets of chromosomes). If the two alleles at a locus are the same the individual is **homozygous** for that gene; if they are different then the individual is **heterozygous**. At the population level, the description of heterozygosity relies on the concept of Hardy–Weinberg equilibrium (Box 3.1). While heterozygosity describes variation in how genes are packaged at each locus, several other terms describe the variation in number of alleles at

each locus. A gene is considered **polymorphic** if >1 allele is detected at a locus across all individuals sampled; otherwise the gene is **monomorphic**. **Allelic diversity** or **allelic richness** describes the average number of alleles per locus.[2] As a practical aside for wildlife management, allelic diversity is more likely to be lost following a severe population contraction (**bottleneck**) than is heterozygosity, because heterozygosity is not much affected (at least initially) by the changes in frequencies of rare alleles lost during a bottleneck.

In addition to measures of nuclear genetic variation based on loci with distinct, identifiable alleles in the nucleus of the cell, another form of genetic variation occurs in **mitochondrial DNA** (mtDNA). Mitochondria are organelles often referred to as the cell's powerhouse because they produce energy. The genes in mitochondria are different from those in the nucleus, with mtDNA coding for cell machinery functions and not for phenotypes that we can observe (with a few exceptions). Because mtDNA is haploid (having one form of the gene, not two as in nuclear genes), heterozygosity cannot be measured. However, mtDNA has some important features that make it very useful for applied population biology. First, in contrast to the one copy of each nuclear gene within each cell, mtDNA is present as multiple identical copies – thousands within most mammalian body cells – allowing analysis of mtDNA from very small or poor-quality samples (e.g. single hairs). Second, in vertebrates mtDNA is maternally inherited and does not recombine, meaning that sons and daughters inherit mtDNA from their mother only. As a result of the maternal inheritance of haploid mtDNA molecules, one breeding pair of parents contains only one transmittable copy of the mtDNA genome, in contrast to the four possible copies of nuclear genes. These features make mtDNA a sensitive marker for detecting hybridization, reductions in population size, tracing maternal lineages and sex-specific dispersal, and inferring mating systems.

Finally, we must distinguish between neutral genetic variation and adaptive variation. **Neutral** genetic variation is unconnected from morphology, behaviors, disease resistance, or other attributes that determine fitness. As we will see, neutral variation is exactly what we want to quantify for many applications that focus on genetic drift (including population structure or connectivity, loss of heterozygosity due to drift, and historical effective population size), because drift by definition drives the random change in allele frequencies due to sampling small populations. Also, individual identification (e.g. in forensic application) is typically based on neutral variation.

However, neutral markers do not tell us about **adaptive variation** in genes that influence fitness. Described more in Chapter 6, **fitness** refers to the relative contribution by individuals of a certain genotype to future population growth. Therefore, adaptive variation is that which defines the present and future potential for individuals to respond to changing environmental conditions. Some of the greatest approaches in genetics in the last decade have been in identifying particular genes that underlie inherited adaptive traits. These open the doors to know, for example, how animals may adapt to high elevations, deal with novel diseases, or confront changes in temperatures. In many cases adaptive variation involves continuous phenotypes (e.g. sprint speed, body size, horn size, date of emergence from hibernation). These **quantitative traits** have complex inheritance with many genes involved, interacting with the environment.

GENETIC MARKERS USED IN WILDLIFE POPULATION BIOLOGY

How do we actually measure genetic variation (heterozygosity and allelic diversity or polymorphism) and use genetic markers to resolve questions in applied wildlife population biology? Prior to the 1960s, genetic variation was measured almost entirely by observing breeding patterns and phenotypic variants, or antigen–antibody reactivities following injection of purified protein into rabbits. These procedures had obvious limitations for tracking genetic variation in wild populations of elusive vertebrates.

The development of protein electrophoresis in the 1960s provided – for the first time – a direct way to measure genetic variation. The protein products of different alleles, called **allozymes**, could be visualized to identify homozygotes and heterozygotes. Although allozyme analysis is relatively inexpensive, requires less training than DNA-based methods, and builds off a rich source of comparative data for hundreds of vertebrate species studied since the 1960s, it has several disadvantages for most wildlife applications. First, the animal must be killed to obtain protein-rich tissue from

[2] Technically, a gene is polymorphic if the most common allele has a frequency of less than 95% (or 99%). Also, allelic diversity is adjusted for sample size because fewer samples will tend to have fewer different alleles.

an organ (e.g. brain, liver, or heart). In addition, tissue must be analyzed when fresh or after quick-freezing, confronting researchers with the prospect of transporting liquid nitrogen or freezers into the field. Third, allozymes have low resolution: the products of only a few genes can be examined, only about a quarter of loci are polymorphic, and heterozygosity tends to be less than 0.1 (Hartl & Clark 1997). The low resolution means that for questions related to recent, subtle changes in genetic variation – for example, due to human-caused population fragmentation – allozymes typically have low power to detect differences.

The next revolutionary wave splashed in the 1980s, with the coupling of the direct analysis of DNA with the polymerase chain reaction (PCR; see Box 3.2). One benefit of DNA-based genetic markers is that they typically have high resolution for distinguishing individuals and populations.[3] Another huge benefit is that PCR facilitates detection of a genetic signal even from samples that are poor in quality or tiny in quantity. This allows genetic sampling of ancient DNA from animals in museums, archives, or archaeological digs, such as 20,000-year-old saber-toothed cats (Janczewski et al. 1992), 50,000-year-old Siberian mammoths (Gilbert et al. 2007), or 120 million-year-old weevils (Cano et al. 1993)!

Perhaps most importantly, PCR-based DNA analysis opens the door to sampling wild animals in ways that are **nondestructive** (where a biopsy or other tissue is obtained but the animal is not killed) and even **noninvasive** (where the genetic sample is collected without having to catch or otherwise disturb the animal) (Beja-Pereira et al. 2009, Kelly et al. 2012). The possibilities for noninvasive DNA sampling are limited only by creativity of the researcher and can include DNA extracted from hair, feathers, feces, urine, blood, ear punches, toe clips, eggshells, carcasses, and antlers (Morin & Woodruff 1996, Taberlet et al. 1999). Freezing of field-collected samples for DNA analysis is not immediately necessary if the sample has been stored in a proper container and dried or preserved (common options include silica gel and alcohol; see Beja-Pereira et al. 2009, Oyler-McCance & Leberg 2012).

The rest of this section will describe some of the most commonly used genetic tools, or markers, in wildlife population studies. The first several mostly describe neutral genetic variation, while the final section describes ways to quantify adaptive variation.

[3] Because not all DNA changes result in amino acid changes and not all amino acid changes alter the protein structure, direct assays of DNA will pick up many genetic changes missed by protein electrophoresis, which detects only changes in the protein products.

Fragment analysis

A wide array of techniques fall into the category of fragment analysis, whereby DNA fragments (mtDNA or nuclear DNA) are run through a gel to separate pieces of DNA by size. Often restriction enzymes are used to break (or "restrict") fragments of PCR-amplified DNA at specific base pairs (often 4–8) to produce **restriction fragment length polymorphism** (RFLP) markers. Different-sized fragments are produced depending on whether and how mutations have changed the DNA sequences recognized by the restriction enzyme. If the sizes of the fragments are characteristic for a species or individual, they can be used for diagnostic identification. RFLPs are widely used to differentiate species from noninvasively collected samples such as scat or hairs; examples include identifying San Joaquin kit foxes and other co-occurring canid species (Paxinos et al. 1997), and differentiating various mustelids in the northwest US (Fig. 3.1).

If PCR primers and restriction enzymes have not been developed to produce diagnostic restriction fragment lengths for a particular species of interest, amplified fragments can be sequenced (determining the nucleotides in order) and compared against an international DNA database (e.g. GenBank). A rapidly expanding variant of sequencing mtDNA fragments for species identification is the "**barcode of life initiative**" (Savolainen et al. 2005). For animals, DNA barcoding targets a 648 bp region of the mitochondrial Cytochrome c oxidase subunit I (referred to as *cox1* or COI). For comparison purposes, the standardized procedure of barcoding has many advantages for species identification from noninvasively collected samples. As just one example, barcodes of *cox1* allowed researchers to distinguish among 24 vertebrates hunted in Africa and South America, facilitating monitoring of poaching and commercial trade in endangered species (Eaton et al. 2010).

Microsatellite DNA

Microsatellite markers are actually a form of fragment analysis, but their properties are different enough and their use in wildlife population biology widespread

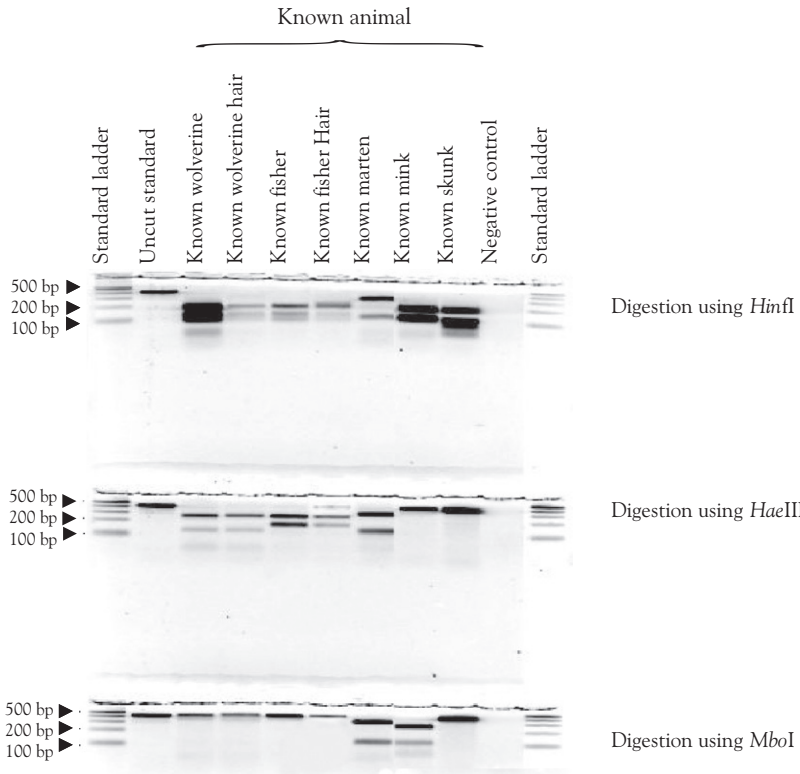

Fig. 3.1 An example of RFLP fragment analysis of mtDNA to distinguish different forest mustelids of the northern USA using single hairs from noninvasive snags (from Riddle et al. 2003, with kind permission of Springer Science and Business Media). After amplifying the cytochrome *b* region of mtDNA with PCR, the DNA was digested with three different restriction enzymes, creating species-specific fragments that collectively distinguish among different species. The first and last lanes are a molecular ladder that helps to determine the size of the bands, and the uncut standard contains a PCR product from a wolverine not subjected to the restriction digests; the negative control is pure water to check for contamination. An example for practice: the first restriction digest (*Hin*fI) distinguishes between marten (with two fragments, of 329 and 113 bp in size) and wolverine (with three fragments, of 212, 132, and 98 bp), but wolverine has exactly the same bands as fisher. Therefore, the next digest (*Hae*III) distinguishes between wolverine (259, 140, and 43 bp) and fisher (259 and 183 bp).

enough to give them their own section heading. Each microsatellite locus contains short (1–10 bp, usually 2–5 bp) sequences of nuclear DNA repeated between 5 and 100 times (for example, the two nucleotide bases cytosine and adenine, or C and A, are repeated 17 times).[4] Microsatellite loci are amplified using PCR, with the size of the amplified alleles determined by the size and number of repeats (so an allele with CA repeated 17 times will be 4 bp smaller than an allele with CA repeated 19 times). As with other fragment analyses, smaller alleles run further down the gel. A homozygote individual displays only one band, whereas a heterozygote displays two (Fig. 3.2).

Microsatellites are well suited to traditional population genetic models because each locus is codominant (distinguishable), with alleles displaying Mendelian inheritance. In this sense, microsatellites produce similar sorts of information (including heterozygosity and allelic diversity) to allozyme electrophoresis. Unlike allozymes, however, microsatellites have very high levels of variation and are PCR-based, facilitating their use with nondestructive and noninvasive sampling (Luikart & England 1999).

[4] The nature of this marker explains its other names: simple sequence repeats (SSRs) and variable number of tandem repeats (VNTRs).

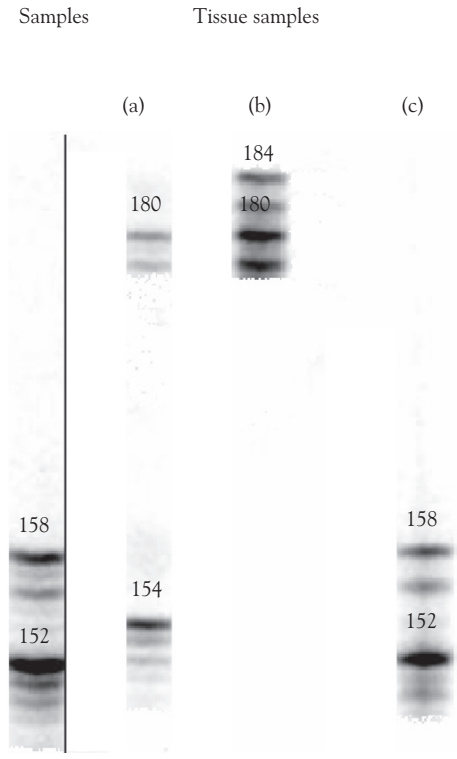

Fig. 3.2 An example of a forensic application using microsatellite DNA (modified from Blanchong et al. 2002, copyright The Wildlife Society). Which deer did the antler come from? Shown here is one microsatellite locus analyzed from an antler sample and from tissues (a–c) from three different white-tailed deer. The size of each allele (in base pairs) in each sample is written above the allele (larger alleles have a higher number of the repeat). In this case each individual sample is a heterozygote. Notice that the antler sample matches tissue sample c. Although just one locus is shown for demonstration, actual applications use multiple loci to minimize the likelihood that twoindividuals share the same genotype.

Single nucleotide polymorphisms (SNPs)

SNPs (pronounced "snips") are nuclear markers especially useful for poor quality samples typical of noninvasive sampling and with potential to explore loci with fitness consequences (Morin et al. 2009). As their name implies, SNPs are regions where different species or individuals within a species have single nucleotide differences (say, a "G" instead of an "A"). In poor-quality samples, SNPs are more easily amplified than microsatellites because the region is shorter (50–70 bp) than that of microsatellites (80–300 bp). Also, SNPs are more widespread through the genome (perhaps every 200–500 bp for many species), so studies may use 50–100 or more SNP loci (or using new genomics approaches described in the next section, many thousands of loci). SNP databases for many speices are rapidly becoming available. Finally, SNPs can be screened in regions containing expressed genes, providing measures of variation for genes with fitness consequences (see the next section). SNPs may eventually replace microsatellites for some applications, such as determining population structure and connectivity using noninvasive sampling.

Genes that affect fitness: functional genomics, adaptive variation, and transcriptomics

The past decade has seen almost unbelievable advances in techniques to analyze the genomes of both model (e.g. humans and domestic animals and plants) and wild species. **Genomics** refers to analysis of large numbers (hundreds or thousands) of loci, across the genome and including sequencing of the entire genome. One application of genomics is greatly to increase the numbers of markers available for assaying

Box 3.2 How PCR turns tiny samples of DNA into larger samples

PCR is a process of amplifying low quality or low quantity DNA samples. The specific steps for conducting PCR can be simplified and summarized as shown in Fig. 3.3 (modified from Frankham et al. 2010:58, after Avise 1994):

Step **a** includes extracting (i.e. isolating) DNA from the sample and preparing the DNA for the PCR reaction. To the isolated DNA are added: **primers** (DNA fragments of about 20 bp) that attach to each strand of the DNA at specific places outside the locus to be amplified; synthetic nucleotides

(bases), the building blocks of life; and finally, Taq polymerase, an enzyme that attaches the nucleotides to the synthesized DNA strand.

Next (step **b**), for each PCR cycle the thermal cycler machine first heats the DNA up to about 94 °C to separate (denature) the two strands of DNA, and then cools to 55–65 °C to bind (anneal) the primers to their target sequences on each strand. Next the temperature increases slightly to about 72 °C to extend the primer into double-stranded DNA, with the Taq polymerase attaching the nucleotides to their respective complementary bases (adenine with thymine, guanine with cytosine) on the other strand. In one PCR cycle you have just doubled the number of DNA strands for your target region!

In step **c** the PCR cycles about 30 more times, nearly doubling the DNA each cycle, ending with millions of copies of each original DNA strand.

Without Taq polymerase, PCR would not happen. Importantly, Taq polymerase was synthesized from *Thermus aquaticus*, a bacterium that lives in hot springs in Yellowstone National Park in the US. The fact that prior to the 1960s Taq was just unknown slime in a hot pool reminds us of a benefit of biodiversity (Varley 1993:14):

> *Here in the world's most popular geothermal region, an obscure, primitive, hot spring bacterium is discovered that contains an even more obscure enzyme that in turn establishes a procedure that promises to change the world for the better . . . The fact is that [Taq] was available for discovery there in Mushroom Pool because the feature and its basin were not available for more destructive, short-term uses . . . Our celebration of Taq is thus tinged with a vague sense of waste: what else, around the world, have we lost already, and how much more can we afford to lose?*

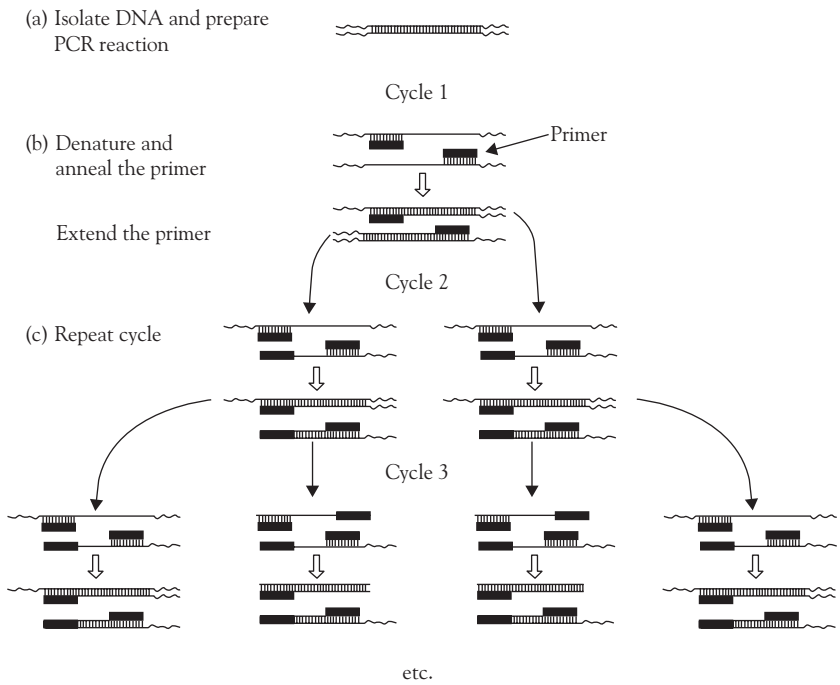

Fig. 3.3 A graphic of the steps involved in PCR. Modified from Frankham et al. 2010:58, after Avise 1994.

neutral genetic variation, such as SNPs or microsatellites. For example, genomic consequences of severe population reductions in the last wild populations of European bison (*Bison bonasus*), plains bison (*B. bison bison*), and wood bison (*B. bison athabascae*) were evaluated by genotyping 52,978 SNPs (Pertoldi et al. 2010)!

Perhaps, most importantly, genomics has also facilitated an explosion of techniques to investigate adaptive variation, a frontier mostly inaccessible to the traditional neutral markers described above. The remarkable technology underlying much genomic research is generally referred to as **high throughput sequencing** or **next generation sequencing** (Davey et al. 2011). Because genomics using next generation sequencing both increases the numbers of neutral markers used in traditional conservation genetics applications and also underlies methods to identify adaptive variation (**functional genomics**), genomics strengthens inferences from classical population genetics while simultaneously allowing us to ask new questions (Allendorf et al. 2010).

Several approaches are currently being used to identify adaptive variation. If something is already known from other species about gene function for a trait of interest, those **candidate loci** can be screened to see if and how they may affect the trait in the new target species. One profitable use of candidate genes has been in considering adaptation to high elevation, where hypoxia (chronic oxygen deprivation) can become a strong agent of selection for any living species. Because hemoglobin is the oxygen transport molecule, the hemoglobin genes are obvious candidates for studies of adaptation to high elevation. Adaptive amino acid changes in the hemoglobin molecule have been well described for deer mice living at different elevations (Fig. 3.4) (Storz et al. 2007, 2009). Another good example of candidate genes being used to great effect is in coat color in mammals, a trait that has critical importance in camouflage for both prey and predators (Box 3.3).

If no candidate genes are available, or to understand functional mechanisms that involve other genes

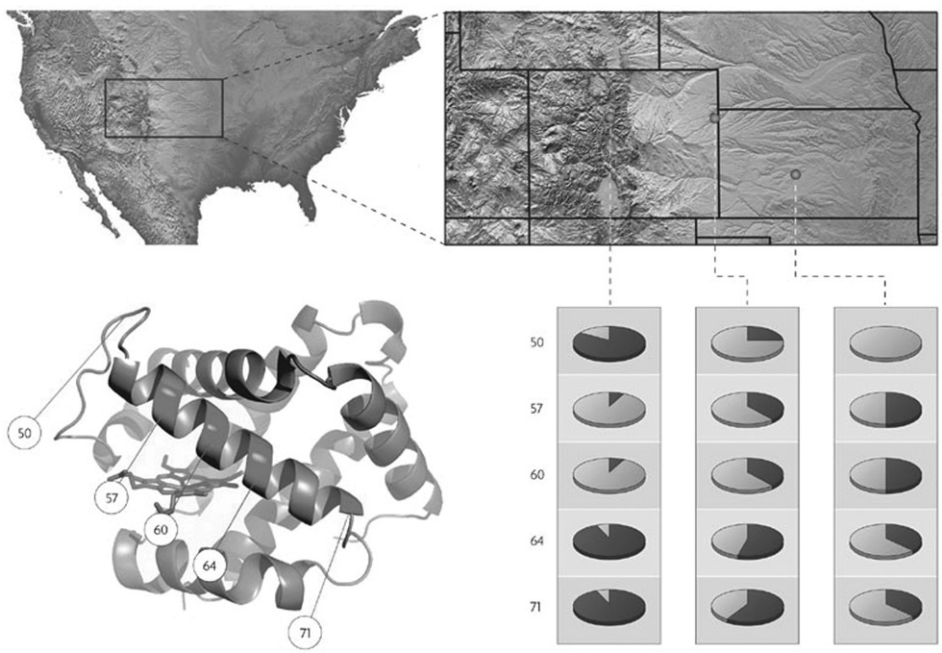

Fig. 3.4 Molecular basis of local adaptation of deer mice to oxygen availability at different elevations. Mice sampled at three locations of different elevations (and therefore oxygen availability) showed different frequencies for five polymorphisms (arising from two interacting genes) in the α-globin protein that makes up part of hemoglobin molecules. Figure from Mitchell-Olds et al. 2007, adapted from Storz et al. 2007. (See Color Plate 1)

Box 3.3 Adaptive genetic variation in mammal coat color

Mammal coat color determines background matching, or camouflage, and camouflage is obviously under strong selection (if you are a tasty prey and you stand out, you die; if you are an ambush predator and you do not blend in, you do not eat). Indeed, camouflage is one of the most important evolutionary forces explaining coloration in mammals (Caro 2005, Stevens & Merilaita 2009). Virtually all color in mammal coats come from variation in quantity, quality, and distribution of just two types of pigments, dark eumelanin and light (yellow to red) pheomelanin. How these pigments get expressed is a fascinating example of how the basis of local adaptation is being revealed by genomics, transcriptomics, and candidate gene approaches.

Many genes contribute in complex ways to the timing, distribution, and shading of pigment; they turn each other on and off, cause hormone cascades, and block and counter-block hormone receptors in incredibly complex ways to lead to coat color (e.g. more than 100 loci in mice affect coloration) (Protas & Patel 2008). These complexities include **epistatic** effects, where one gene controls the expression of others. Through this complexity, two genes have emerged as especially prominent candidate genes that direct coat color for many species: *Melanocortin 1 receptor* (*Mc1r*) and *Agouti* (the gene and the gene product have the same name but the gene is italicized). Mc1r has been linked to intraspecies color differences in many mammal species (from pocket mice and deer mice to jaguars), as well as at least five bird species and some reptiles (Protas & Patel 2008). Simply put, Mc1r is a receptor expressed in melanocytes (pigment-producing cells) to produce eumelanin and dark coats, and Agouti is a repressor, or antagonist of Mc1r, that shifts pigment production toward lighter pigmented pheomelanin. In many cases other genes such as *α-melanocyte stimulating hormone* (*αMSH*) or (in dogs) *β-defensin* act against Agouti to lead to dark coats (see Candille et al. 2007 for a nice explanation of why your Labrador retriever is yellow versus black).

For *Peromyscus* mice, selection on background matching of coat color is strong (Kaufman 1974) and recently the genetic basis of multiple *Peromyscus* populations has been well studied. Natural selection has, for example, led to coastal beach mice on light and bright sandy environments having a lighter coat color compared to nearby mainland mice living on darker inland soils; changes in the Mc1r candidate gene can explain much, but not all, of this adaptive variation (Fig. 3.5; see Hoekstra et al. 2006; Mullen et al. 2009). Interestingly, even though selection has led to phenotypic convergence so that some of the beach subspecies are hard to tell apart based on coat color, neutral markers (microsatellites) show strong differentiation among subspecies, implying that they are "good" subspecies that should be considered for status as separate species (Mullen et al. 2009).

Finally, to address a related topic that often comes up when talking about coat or skin color: albinism in humans can be traced to several genes that do not enter the pigmentation process. Rather, these genes control (are epistatic to) all or part of the pigmentation pathway. These are recessive genes, meaning that as long as an individual has at least one dominant albinism allele, the skin pigments do their thing and express their color. However, an individual homozygous for the recessive albinism alleles expresses the albino condition, overriding the pigmentation pathways.

(*Continued*)

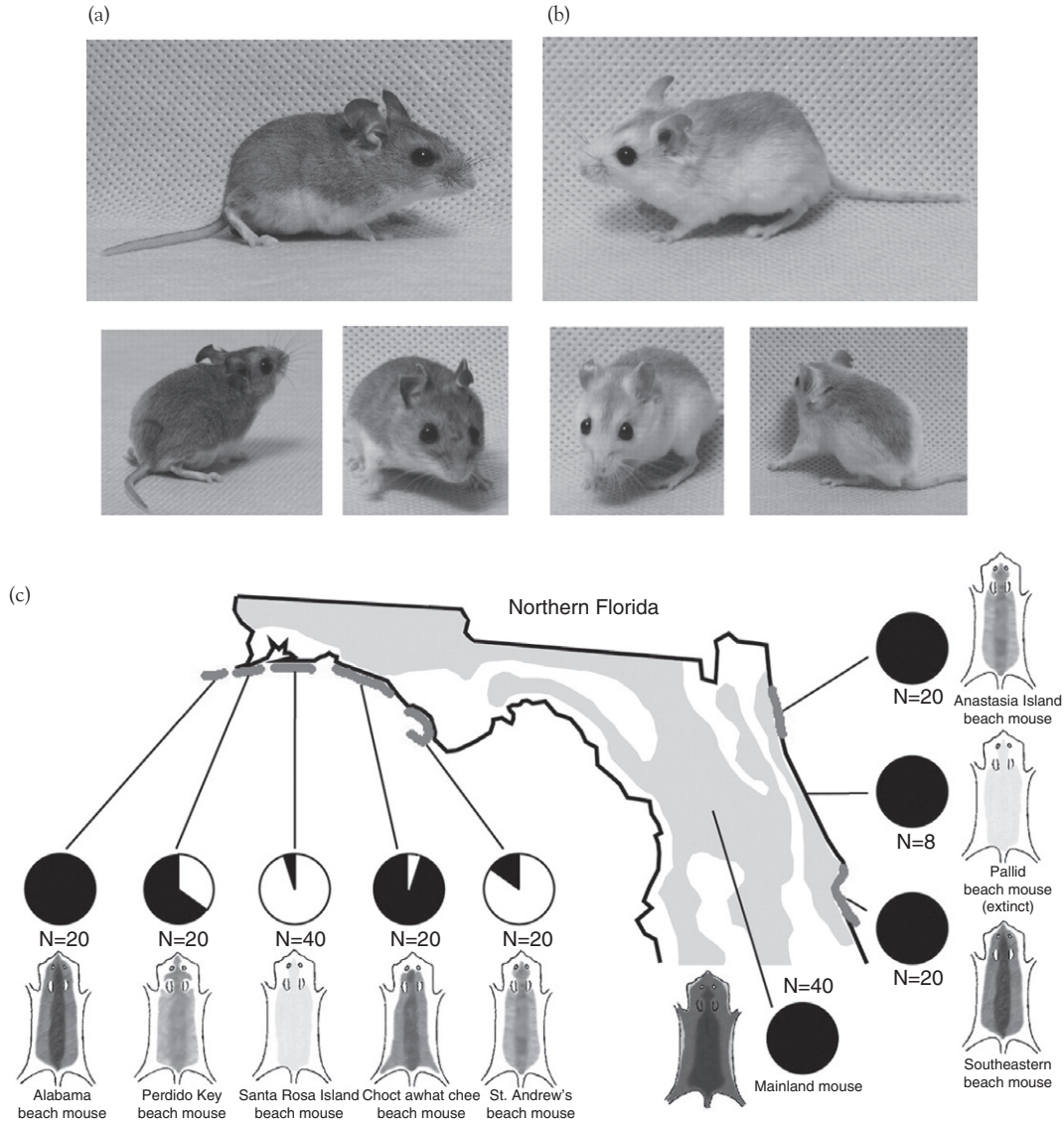

Fig. 3.5 The genetic basis for camouflage in wild mice (*Peromyscus polionotus*). Panels (a) and (b) show typical coloration of (a) a mainland mouse found on dark soils throughout the grayed area on map and (b) a beach mouse from the Santa Rosa Island subspecies, which is, like other coastal subspecies (red areas in Panel (c)) found on light-colored sandy soils. (Pelage lightness correlates closely with sand brightness at that location for the different subpopulations; Kauffman 1974, Mullen et al. 2009.) Circles in (c) show frequencies of the light and dark alleles of Mc1r, the candidate coding gene for coat color. Notice that the three beach mouse populations on the east coast of Florida have converged on the light phenotype through a different mechanism that does not involve the Mc1r light allele. Panel (c) modified from and Panels (a) and (b) directly from Hoekstra et al. (2006). (See Color Plate 2)

beyond previously identified candidates, large numbers of markers or extensive sequencing can be used to identify where and how selection is operating. One way to determine regions under selection is called **outlier analysis**. Based on the massive numbers of neutral markers such as SNPs or microsatellites generated by next generation sequencing, markers that deviate in frequencies relative to neutral expectations provide a signal of past selection at the marker or at loci linked to the marker.

High throughput sequencing also facilitates whole-genome sequencing. Though unthinkable for any species until recently, as of 2010 over 60 complete animal genomes have already been published, and an ambitious project called the "1,000 Plant and Animal Reference Genomes Project" plans to coordinate sequencing of 150 animal genomes in 2011 alone (Ekblom & Galindo 2011). Sequencing throughput of all species is doubling every 5 months (Davey et al. 2011).

A more common approach, one step short of whole genome sequencing, uses restriction enzymes to break DNA into manageable sizes for sequencing as genetic markers. Many types of these **reduced-representation sequencing** methods are available (Davey et al. 2011, Ekblom and Galindo 2011) and are being applied to wild species that do not necessarily require a **reference genome** (same or closely related species, already sequenced).

A fundamentally different approach to identify adaptive genetic variation directly assays gene expression by mapping the RNA that arises as the transcript products of genes. This approach, known as **transcriptomics** has the advantage of assaying not only changes in gene sequences but also the regulation of gene expression, which may explain as many differences in phenotypes and evolutionary changes as the gene sequences themselves. One rapidly developing way to do transcriptomics, known as RNA-seq (Wang et al. 2009), relies on high throughput sequencing of the mRNA transcribed from a stretch of DNA.[5] The sequence reads are then compared to reference genomes to produce a transcription map.

The biggest disadvantage of genomics via high throughput sequencing is cost (but that is dropping fast[6]) and the potentially overwhelming amount of data that can be generated when reading perhaps millions of base pairs daily. Genomics and transcriptomics will continue to revolutionize how we study and manage wildlife populations.

INSIGHTS INTO WILDLIFE POPULATION BIOLOGY USING GENETIC TOOLS

Molecular biology and noninvasive sampling have truly awesome potential for population analysis (Table 3.1). Most of the chapters in this book will contain some application of genetic tools to wildlife population biology questions, for example in quantifying connectivity and isolation among populations, estimating abundance, and solving forensics cases. Instead of elaborating all of these myriad applications in this chapter, I will introduce just a few uses of genetic tools that relate to some of the most basic tasks in population biology: identifying important taxonomic units and distinguishing among species and individuals.

Taxonomy and hybridization

How individuals are assigned to taxonomic groups determines the fundamental units of conservation and management. Genetic characteristics supplement morphology and other information (e.g. life history, geographic range) to determine taxonomic affiliation. In so doing, genetic information may reveal that groups historically lumped into one species actually have distinct evolutionary legacies, with potentially different conservation needs. Conversely, multiple species or subspecies actually may not be distinct; recognizing the similarity may release resources that could be spent on taxa with more critical needs. Box 3.4 describes case studies of each of these scenarios. Genetic markers can also supplement other information to help resolve important taxonomic affiliations below the species or subspecies level, including **evolutionary significant units, designatable units,** and **management units** (Box 3.5).

[5] Technically, the RNA is converted to cDNA (DNA complementary to the RNA base pairs) in a reaction catalyzed by the enzyme reverse transcriptase; it is the cDNA that is sequenced.

[6] Want to sequence yourself (Davey et al. 2011)? A diploid human genome consisting of two 3-gigabase sequences costs about 5000 British pounds or roughly 8000 US dollars. For about 1/35 of that amount, you could sequence a sample 200,000 markers from restriction-site associated fragments.

Table 3.1 Some insights into wildlife population ecology that can be gained from the use of genetic markers and analysis. In most cases nondestructive or even noninvasive sampling can be used with both nuclear and mtDNA markers. Most examples are detailed in this chapter or elsewhere in this book.

Application	Examples
(a) Taxonomic relationships	Tuatara across New Zealand should be managed as multiple taxonomic units instead of just one; seaside sparrows should be managed as two taxonomically important units instead of nine
(b) Hybrid detection	Lynx–bobcat hybrids, barred owl–spotted owl hybrids, coyote–red wolf hybrids
(c) Species identification and distribution	Carnivore species distribution based on hair-rub pads or scats; what prey species are eaten based on remains in owl pellets
(d) Individual identity and gender as a basis for estimating abundance and vital rates	Estimation of the abundance of humpback whales in the North Atlantic, grizzly bears in the US rocky mountains and wombats in Australia; survival estimation of tigers in India
(e) Connectivity among populations	Number of skinks moving between rocky outcrops; sex-specific dispersal rates of white-toothed shrews
(f) Levels of genetic variation and size of historical populations	Heterozygosity loss in the small and endangered Florida panthers; historic size of northern elephant seal populations before being decimated by hunting
(g) Costs of inbreeding and benefits of genetic rescue	Wild bighorn sheep suffer inbreeding depression that sometimes does and sometimes does not affect population dynamics; multiple fitness traits recover and Florida panther numbers increase following genetic rescue
(h) Parentage	Which wolves in a pack produced which pups; which male rhinoceros breed
(i) Forensics	Which species were killed to produce so-called whale meat in restaurants; what species and which individuals were killed by a poacher using a confiscated bloody gun

Box 3.4 Genetic information clarifying taxonomy and improving wildlife management

Case study 1: a group managed as a single species is actually multiple subspecies with distinct conservation requirements

The New Zealand tuatara (*Sphenodon*) (Fig. 3.6) is the only surviving genus of one order of reptiles and is probably the most distinctive surviving reptile genus in the world, with a morphology nearly unchanged over the last 100 million years. Tuatara have been protected fully since 1895, with the focus on a single species (*Sphenodon punctatus*) throughout New Zealand. Subsequent genetic and morphological analyses, however, have determined at least two different regional taxonomic groups that warrant separate management (Daugherty et al. 1990).

Fig. 3.6 The Stephens Island tuatara, an as-yet-unnamed subspecies of *Sphenodon punctatus.*

Neglecting these distinctions could lead to the extinction of evolutionarily distinct groups or lead to inappropriate mixing during translocations.

Case study 2: multiple taxonomic units are recognized to actually be just two (from Avise & Nelson 1989)

Historically, nine subspecies of the seaside sparrow (*Ammodramus maritimus*) were recognized based on plumage and subtle morphological characteristics. One of these subspecies, the dusky seaside sparrow (*A. maritimus nigrescens*) was listed as endangered in 1966, as it dwindled in number due to habitat change. In 1980, only six males remained (demographic stochasticity in action; see Chapter 5) and the subspecies was considered extinct by June 1987. However, subsequent mtDNA analysis indicated that the dusky seaside sparrow was not a unique subspecies after all: there was no basis for phylogenetic distinction of *A. maritimus nigrescens* from other Atlantic coastal populations of *A. maritimus*. Because all Atlantic populations shared one mtDNA genotype and all Gulf Coast populations shared another (see Fig. 3.7), the major conservation focus should be on two subspecies – Atlantic coastal populations and Gulf Coast populations – instead of nine. Recognizing that a taxonomic revision is strongest when supported by a combination of approaches, Avise and Nelson (1989) also found morphological and ecological support for their thesis. In short, there is no question that the habitat loss in this case has been disastrous and that it wiped out the local population formerly known as the dusky seaside sparrow. However, in retrospect the dusky seaside sparrow probably did not warrant the conservation attention that we might provide for more unique lineages, and conservation would be better served by focusing on just two forms: Atlantic and Gulf Coast.

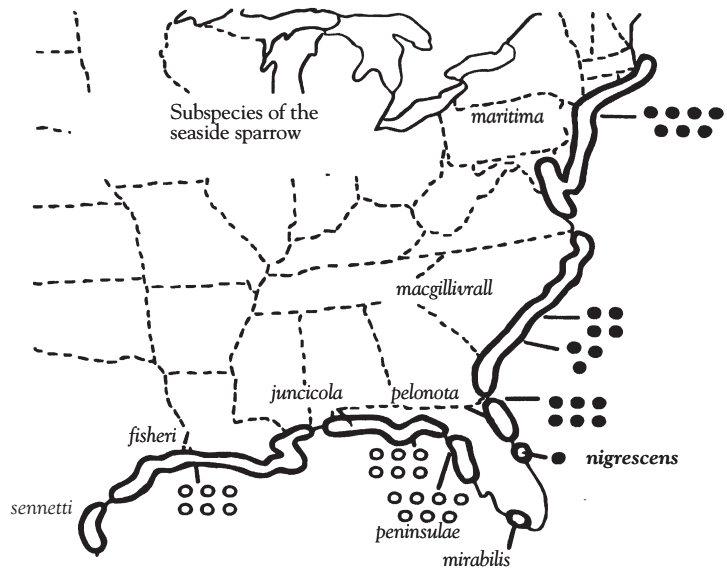

Fig. 3.7 Geographic distributions in the eastern US of the nine originally recognized subspecies of the seaside sparrow. Open and closed circles represent birds carrying distinctive Gulf Coast and Atlantic Coast mtDNA genotypes respectively. From Avise and Nelson (1989). Copyright (1989) AAAS.

Box 3.5 Identifying population units for conservation management below the species or subspecies level

In many cases, conservation policy needs higher resolution distinctions than "species" or "subspecies". An **evolutionary significant unit** (ESU) has been defined as "a population (or group of populations) that (1) is substantially reproductively isolated from other conspecific population units, and (2) represents an important component in the evolutionary legacy of the species" (Waples 1995:9). In other words, ESUs are considered to be the ecological and evolutionary building blocks of the species, whose conservation will allow the continued evolution of the species. ESUs are relevant to legal and policy frameworks, including the Species at Risk Act (2003) in Canada, the Environment Protection and Biodiversity Conservation Act (1999) in Australia, and the Endangered Species Act (1973) in the US (where the Act allows for listing not only species and subspecies but also **distinct population segments**, which are considered equivalent to ESUs; Waples 2006).

The scientific criteria for defining ESUs are evolving, with much discussion on the appropriate role of genetic markers. For example, some scientists argue for an operational definition based strictly on molecular phylogenies to discern historical isolation and evolutionary potential: ESUs are reciprocally monophyletic for mtDNA alleles (i.e. all members are descended from a single common ancestor unique for each ESU) and differ significantly in allele frequency at nuclear loci (Moritz 1994, 1995). In contrast, others call for incorporation of geographical, life-history, habitat, behavioral, and morphological differences to help determine whether the population represents an important part of the species' evolutionary legacy (Waples 1995, Crandall et al. 2000).

Sometimes these different approaches will not point toward the same ESU designation. For example, one small group of populations of Cryan's buckmoths is geographically separated from other populations in North America. A strictly molecular ESU criterion would suggest that this is not an ESU, as there are no significant differences in allele or haplotype frequency between these isolated populations and others. However, host plant performance experiments indicated that Cryan's buckmoth larvae consume and grow on a unique plant host compared to other populations, implying an ESU based on local adaptation not detected by neutral genes (Fraser & Bernatchez 2001). In other cases, molecular criteria will provide insights – particularly about the historical legacy of populations – when there are neither resources nor time to conduct experiments of adaptive differences among population groups. In summary, differing ESU approaches can complement each other, focusing on the common goal of addressing the protection of evolutionary potential.

The Committee on the Status of Endangered Wildlife in Canada has adopted **designatable units** (DUs) to identify population units below the species level for separate analysis of risk status (Green 2005). DUs are first based on whether populations are distinguishable (this embraces the ESU idea) and then on whether they have differing conservation status (based perhaps on regulations or specific stressors in that area that affect vulnerability).

Similarly, **management units** (MU) refer to population groupings based on restricted demographic interchange (analogous to the stock concept in fisheries). The meaning of "restricted" depends on management objectives, so MUs may have little resemblance to ESUs. Although MUs may have diverged allele frequencies in nuclear DNA, they are not expected to show reciprocal monophyly for mtDNA alleles (Moritz 1994). Thus MUs are unlikely to have different evolutionary potentials: the target level of distinctiveness will be driven largely by policy or management needs and explicit consideration of risks (Taylor & Dizon 1999). For example, the relatively well-connected populations that make up an MU may collectively sustain a higher harvest rate than if the harvest focused on a more isolated target (Brook & Whitehead 2005). In contrast to ESUs, where translocations are generally avoided, translocations among MUs will generally not be detrimental, and may even be advantageous for maintaining genetic variation.

Genetic information can also help in detecting and interpreting the consequences of **hybridization**. Hybridization – defined broadly as the interbreeding of individuals from genetically distinct populations – is an enormously complex topic with difficult biological issues and perplexing management implications (for excellent overviews see Rhymer & Simberloff 1996, Allendorf et al. 2001). Prior to 1990, interpretation of the US Endangered Species Act 1973 reflected the widespread view that hybrids were impure, so that protection under the US Endangered Species Act should be discouraged for hybrids between species or subspecies.

In an excellent example of biological information directly influencing policy, a paper by O'Brien and Mayr (1991) helped overturn the hybrid policy. Although there is currently no formal policy on hybrids, federal agencies recognize the following (US Department of the Interior, Department of Commerce 1996):

- Occasional hybrids are to be expected between species, and "natural occurrences of hybrid individuals or hybrid zones between recognizable species do not disintegrate the genetic integrity of the species" (O'Brien & Mayr 1991:1187–8). For example, occasional hybrids with the bobcat should not influence the threatened status of Canada lynx in the contiguous US. As long as the hybrid offspring more closely resemble the listed species – based on morphological, behavioral, ecological, and molecular data – the US Endangered Species Act protection extends to those offspring.
- Hybrid lineages between species usually die out, but they will sometimes establish themselves as a breeding population with their own adaptations and evolutionary history worthy of conservation (assuming the lineage was developed outside of confinement and is self-sustaining and naturally occurring). Thus red wolves should receive protection under the US Endangered Species Act, even if they originated as wolf–coyote hybrids (see Box 3.6).
- At the subspecies level, hybridization occurs naturally and may have adaptive benefits. In cases where genetic variation is low and new genetic variation is brought in to combat inbreeding depression, offspring should receive protection under the US Endangered Species Act and the population status should not be compromised. The breeding of Texas panthers with highly endangered and inbred Florida panthers (see Chapter 9) is an excellent example.
- By contrast, hybrid progeny (among species or distinct subspecies) arising haphazardly from human actions should be discouraged and potentially removed, especially when intercross progeny jeopardize the persistence of a listed species (Rhymer & Simberloff 1996). For example, mallard ducks have been introduced around the world, hybridize readily with narrowly distributed endemic species, and have been implicated in declines of New Zealand grey ducks, Hawaiian ducks, Florida mottled ducks, and Australian black ducks. To hint at the complexity of the hybrid issue, Allendorf et al. (2001) point out that when such introgression becomes nearly complete, as it has in the case of the New Zealand grey duck, conservation (instead of elimination) should be considered because there may be no other option to avoid complete loss of the hybridized species.

Because of its maternal inheritance, mtDNA can be used to help determine the direction of hybridization. For example, because coyote mtDNA is found in gray wolves but not vice versa, hybridization between coyotes and wolves occurs by way of a male wolf mating with a female coyote (Lehman et al. 1991). Similarly, lynx were the mothers of Canada lynx–bobcat hybrids (Schwartz et al. 2004), sage grouse the mothers of sage- and sharp-tailed grouse hybrids (Aldridge et al. 2001), and barred owls the mothers of barred owl–spotted owl hybrids (Haig et al. 2004).

In summary, the appropriate way to deal with hybrids once again invokes the nondemocratic adage trumpeted throughout this book: not all hybrids are created equal. The hybrids most important to eliminate can also be those that are hardest to detect: for example, hybrids derived from domestic animals or from human-induced habitat changes (e.g. barred owls moving west with logging and other habitat changes and mating with threatened spotted owls; Haig et al. 2004). For small populations, hybridization from human-induced changes may be an underappreciated threat, as sterile hybrids can lead to demographic dead ends for population growth of the species of concern, and fertile hybrids can lead to hybrid swarms that also threaten persistence of the pure species (Chapter 11).

Determining species identity and distribution

Many species are hard to catch and hard to see, and identifying species solely from morphology of hairs or

Box 3.6 The red wolf (*Canis rufus*) as a case study in detecting and interpreting hybridization

Fig.3.8 A red wolf pup being held by a US Fish and Wildlife Service biologist. Photograph courtesy of US Fish and Wildlife Service.

The red wolf was once distributed throughout the southeastern US, but numbers plummeted in the early 1900s due to predator control, habitat destruction, and hybridization with coyotes (Phillips et al. 2003). The species was listed as endangered in 1967. Because the few remaining free-ranging red wolves were becoming hybridized out of existence by coyotes (which were expanding eastward), the last red wolves were removed from the wild to use as breeding stock for eventual reintroduction. In 1973 a captive breeding program was begun with 14 of the most "pure" red wolves out of 400 animals captured in southwestern Louisiana and southeastern Texas. Although the wild population was considered extinct in 1980, reintroductions began in 1987 into a 680,000 ha peninsula in eastern North Carolina (M.K. Phillips et al. 2003). The reintroduction appears to be a success: by 2002, all red wolves in the population were wild-born, and the 2011 population consists of at least 100–20 animals distributed in 20 packs.

Within the scientific community, red wolf conservation has been controversial, largely because it touches on so many of the vexing issues related to hybridization. Hybridization with coyotes (which expanded into the eastern US only in the 1990s) is the biggest threat to red wolf persistence, so management requires that hybrids be identified using molecular methods (Adams & Waits 2007) and then eliminated or sterilized. Although some argue that red wolves may have originated as hybrids between coyotes and gray wolves, such an event would have pre-dated modern human activities. As Dowling et al. (1992:602–3) note: "Genetically distinct taxa of hybrid origin must not be denied protection [under the US Endangered Species Act 1973] due to mixed ancestry. If the red wolf proves to represent an historically stable entity generated by long past (maybe even ancient) hybridization between gray wolf and coyote, then it is a taxon of hybrid origin that clearly should be protected." Indeed, this plea for species of natural hybrid origin to be protected under the ESA is current US Fish and Wildlife Service practice (Haig & Allendorf 2006:156).

scats is notoriously unreliable (Piggott & Taylor 2003). Now, however, species can be reliably identified through DNA analysis of evidence that they leave behind. mtDNA is the usual marker of choice for species identification based on small or degraded samples, primarily due to the multiple mtDNA copies in each cell (typically 100–1,000 or more) compared to the one copy of nuclear DNA. The researcher identifies regions of mtDNA – perhaps RFLPs or sequences from the PCR product – that are variable among species but conserved (constant) within species, providing diagnostic markers for species identification.

For example, the distribution of Canada lynx and other forest carnivores across the entire northern US was evaluated noninvasively by sampling hairs left behind on rub pads (Box 3.7). Similarly, "molecular scatology" (Kohn & Wayne 1997) can be used to determine species in an area. For example, in a study to determine food habits of carnivores in Venezuala, scat sizes overlapped for sympatric species (puma, jaguar,

Box 3.7 The National Lynx Survey as a case study for species identification using noninvasive genetic sampling

In Chapter 2, I described the rationale behind the National Lynx Survey. The use of noninvasive sampling to assess the distribution of lynx across 16 states proved to be a reliable and informative approach for conducting the survey.

The device for sampling elusive and low density lynx was a 10 cm × 10 cm carpet pad with nails sticking out, smothered in a beaver castoreum and catnip oil scent lure. Lynx (and other species) rub against it and leave hairs behind (McDaniel et al. 2000). At each sampling site, 125 rub pads were placed in a systematic grid: 25 transects 3.2 km apart, with each transect consisting of five rub pads 100 m apart. Pads were checked after 2 weeks.

Species identification of the collected hairs relied on PCR amplification of short (about 400 bp) segments of mtDNA, coupled with the use of restriction enzymes, to produce species-specific fragments of DNA (Mills et al. 2000a). These fragments are consistent across the range of a species and are not shared by other species (see Fig. 3.9). Importantly for identification of species of political concern, exhaustive tests to validate the species-identification protocol were conducted prior to initiating the survey (Mills 2002).

After 3 years of sampling, more than 21,000 pads had been placed in the field and from these approximately 7000 samples were processed (McKelvey & Mills, unpublished data). About 67% of the hair samples – including single hairs or fragments of hair – could be identified to species. Although the sampling method was designed to target lynx, and 96 rubs from lynx were recorded (mtDNA only identifies species, not number of individuals), similar approaches facilitated the identification of other forest carnivores that happened to rub on the pads (Riddle et al. 2003; see Fig. 3.1). For example, 2040 rubs were from black bears, 414 from bobcats, 109 from cougar, 25 from domestic cats, and 383 from coyotes.

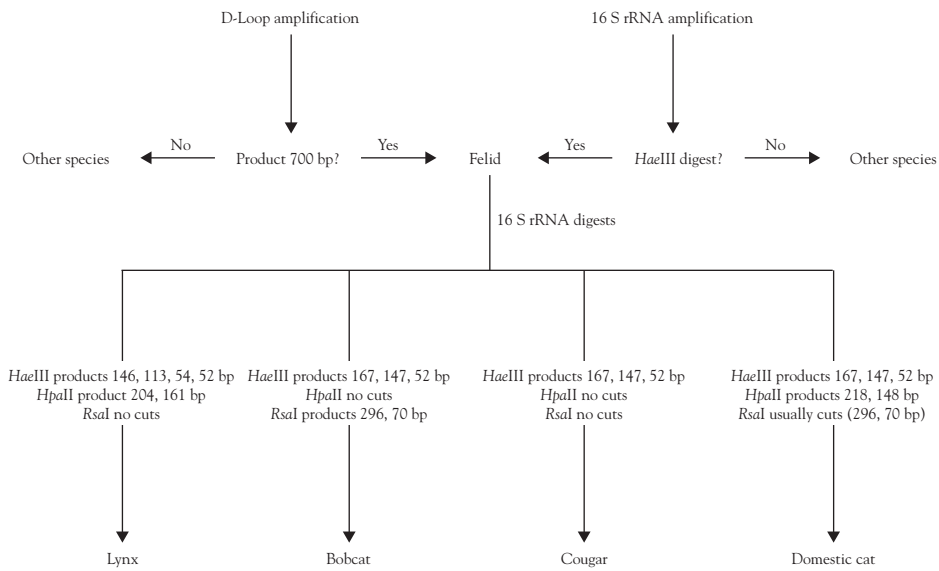

Fig. 3.9 The approach used to distinguish felids in the National Lynx Survey of the USA. The diagnostic test for felid species identification was a 360-bp section of the 16 S rRNA amplified from mtDNA and subsequently digested with restriction enzymes to produce species-specific fragment patterns. From Mills et al. (2000a). Reproduced with kind permission of Springer-Verlag.

ocelot, and crab-eating fox), so a diagnostic mtDNA test was required to distinguish the scats by species (Farrell et al. 2000; see Paxinos et al. 1997 for US carnivores). Likewise, early detection of red foxes, a devastating yet elusive invasive species into Tasmania, was greatly improved by an island-wide scat collection program that yielded 1160 georeferenced carnivore scats, of which 78% contained sufficient DNA for speces identification (Berry et al. 2007). Excreted material can even yield the identity of the species that were eaten, such as small mammals identified from material in owl pellets (Taberlet & Fumagalli 1996).

Of course, the plethora of new molecular techniques has not gone unnoticed by the wildlife law-enforcement community, with **forensic** applications identifying species or population of origin, sex, and individual identity. Although we will return to examples throughout the book, one of the earliest cases of DNA forensic work involved investigating whether whale meat was from species that could be legally harvested. Baker and Palumbi (1994) purchased whale meat from Japanese retail markets and restaurants, which ranged from unfrozen sliced meat to dried and salted strips marinated in sesame oil and soy sauce. Because international laws prohibited them from transporting tissue samples to their laboratories in New Zealand and Hawaii, they set up a mobile PCR laboratory in their hotel room, amplifying mtDNA so it could later be identified to species via sequencing.[7] By comparing the species identity and geographic origin obtained from DNA analysis with catch records, they were able to conclude that several species of protected whales were being hunted, processed, and imported illegally and sold openly on the Japanese retail market. Subsequent repeats of this approach in both Japan and South Korea (Lento et al. 2001) led to the purchase of over 1000 samples. Of these, 61 turned out to be from internationally protected whale species and more than 140 were not true whale meat at all, but rather were porpoise, dolphin, sheep, or horse!

[7] This approach is both technically innovative and assures compliance with international law under the Convention on International Trade in Endangered Species (CITES). No DNA from the original purchased meat was transported out of the country, because the researchers used a clever molecular approach (Palumbi & Cipriano 1998): the DNA synthesized during PCR attaches to magnetic beads that get pulled out of the tube with a magnet, separating the synthetic product entirely from native whale DNA, which is left behind.

Determining gender and individual identity

Once a species is identified, gender and individual identity may be determined with other genetic markers. In mammals, where male sex chromosomes are XY and females XX, males carry DNA markers associated with the Y chromosome (such as the testis determining factor, the SRY gene), whereas females do not; this means gender of the animal can be determined from a scrap of skin, a bundle of hairs, or feces (Woods et al. 1999, Shaw et al. 2003, Pilgrim et al. 2005). A different approach amplifies a portion of DNA with alleles of different size residing on both the X and Y chromosomes (Shaw et al. 2003). Gender-determination techniques are often used in **forensic** work, especially for ungulates, where game laws tend to be strongly sex-specific. The use of PCR with Y chromosome markers has successfully determined the sex of killed ungulates based on bloodstains (from knives and rifle bolts), hair, and meat (Gilson et al. 1998). In many other animal species – including birds, snakes, and some turtles and lizards – females have the heterogametic sex chromosomes (e.g. females are ZW and males are ZZ; Griffiths et al. 1998, Modi & Crews 2005), so DNA markers can again be used to identify gender. However, determining gender for some reptile, amphibian, and fish species using sex chromosomes is complicated because gender can be determined by other factors such as temperature, pH, or social conditions (Chapter 4).

Individual identity is usually determined based on genotypes derived from highly variable nuclear DNA, especially microsatellite DNA and SNPs. Once individuals are distinguished based on their genotype, many exciting applications follow. Obviously, forensics benefits from knowing not just species but also gender and identification (Box 3.8). A different type of forensic example involves identification of predators killing domestic sheep: salivary DNA from puncture wounds on sheep carcasses were swabbed and analyzed with DNA to determine not only the species but also the gender and individual identity of the attacking coyote (Blejwas et al. 2006).

Another use of individual genotypes is to estimate abundance and other vital rates in a noninvasive capture–mark–recapture framework (Box 3.9). In one of the first applications of these approaches, humpback whales in the North Atlantic Ocean were sampled noninvasively (from sloughed whale skin) and nondestructively (from biopsy darts). An abundance of 4894 (95% confidence interval, 3375–7123)

Box 3.8 DNA identifies a serial wolf killer in Italy

Wolves are slowly recovering in parts of Europe, but poaching remains a severe threat, with an estimated 20% of the population illegally killed each year in Italy. In 2008, police confiscated a tooth necklace from a man in Italy, then a few days later discovered a dead (poached) wolf with the muzzle missing. Were the teeth on the necklace from local wolves? If so, how many individuals were represented, and did the teeth include the dead specimen (Caniglia et al. 2010)? A genetic database of European dog and wolf samples was already in place. Analysis of mtDNA, 12 microsatellites, and several sex-specific genes extracted from the tooth necklace and compared against the wolf and dog genetic database delivered an irrefutable indictment against the owner of the necklace: the teeth came from six different individuals, three males and three females, all native Italian wolves (not dogs or wolves from other countries); and yes, one of the teeth completely matched the genotype of the macabre specimen found with the missing muzzle (Caniglia et al. 2010).

Box 3.9 Abundance estimation using noninvasive genetic sampling

Noninvasive genetic sampling has revolutionized the possibilities for estimating abundance and other vital rates (Chapter 4) of wildlife species formerly considered too elusive or expensive to sample using traditional approaches. In compelling words: "Relief from sampling despair has an unexpected source in feces" (Kohn & Wayne 1997:226). Without the animal knowing it, multiple samples can be collected and individually genotyped, just as individual animals are often marked in traditional mark–recapture studies (Chapter 4).

However, this revolutionary approach to marking animals noninvasively has challenges (Waits & Leberg 1999). Low quantities or quality of template DNA may indicate nonexistent individuals through several mechanisms, collectively known as **genotyping error**. Obviously, contamination of the sample can give false signals. Even without contamination, however, slippage of DNA polymerase during PCR can create **false alleles** where the size of an allele is scored incorrectly. Also, **allelic dropout** can occur when one or more alleles at a heterozygous locus fail to amplify during PCR, such that a heterozygote is scored as a homozygote (genotype AB is scored as either AA or BB). These genotyping errors will tend to cause a positive bias in abundance estimates, because the same animal captured multiple times may appear to be different animals (see Creel et al. 2003). On the other hand, an opposite problem, where genotypes are insufficient to distinguish among different individuals, will tend to create a negative bias in the abundance estimate. We have called this a **shadow effect** (Mills et al. 2000b), because multiple different animals could be indistinguishable genetic shadows of each other.

These concerns extend beyond abundance estimation and are applicable to most uses of noninvasive sampling, but they are not insurmountable. The shadow effect disappears as the number of markers is increased through genomics and high variation markers. Genotyping errors are decreasing with improved laboratory techniques and can now be formally dealt with in a way that maintains rigor of abundance estimates (e.g. Lukacs & Burnham 2005, Kalinowski et al. 2006, Wright et al. 2009).

male and 2804 (1776–4463) female whales was estimated, with local and migratory movements of up to 10,000 km and genetic mixing in winter breeding areas (Palsboll et al. 1997). Other classic examples include grizzly bear abundance determined from hairs left behind on barbed wire hair snares and on natural rub trees (>33,000 hair samples across a 31,000 km^2 study area in northwestern Montana resulted in an impressive estimate of 765 bears with a 95% CI of 715–831; Kendall et al. 2009), and highly endangered northern hairy-nosed wombat abundance calculated from single hairs collected at burrow entrances (Sloane et al. 2000). Likewise, the number of breeders or genetically effective population size can be directly estimated from genotypes from noninvasive genetic sampling (Chapter 9).

Markers such as microsatellites and SNPs can do more than just provide an individual-specific "barcoded" genotype. They also display Mendelian codominant expression, meaning that they identify alleles inherited by that individual from both parents. From this one can calculate allele frequencies, heterozygosity, and other measures of genetic variation. In turn, these can be used to estimate population structure, connectivity among populations, population of origin of immigrants (Chapter 10), loss of genetic variation that can lead to inbreeding depression (Chapter 9), and the occurrence of drastic population reductions (**bottlenecks**) in the past or present (e.g. Cristescu et al. 2010, Luikart et al. 2010).

Finally, individual genotypes derived from DNA can be used to assess **parentage** by comparing alleles of an individual to alleles of putative parents. Genetic analysis of parentage can help determine reproductive success, including the proportion of parents that breed, and mean and variance in number of offspring per parent. Genetic measures are most useful when direct behavioral observations are either impossible or potentially misleading. For example, little was known about the mating system of the highly endangered black rhinoceros before Garnier et al. (2001) used microsatellite analysis of feces to document strong polygyny and skewing of reproductive success (of 19 offspring, more than half were sired by one male, whereas seven of 11 adult males had no offspring over 10 years). These findings were important both for captive breeding and for deciding appropriate translocations among the small, scattered remnants of the wild population.

SUMMARY

Describing, measuring, and interpreting genetic variation has become a central component of modern wildlife population biology. A major reason for the explosion of genetic techniques and applications for wildlife has been the development of the PCR, which has allowed genetic information to be obtained from unfathomably old or tiny tissue samples obtained noninvasively. Another technological breakthrough, exploding into our field right now, is the development of high throughput sequencing that allows evaluation of large portions of the genome and idenfication of the genetic basis of traits under selection in a changing world (genomics and transcriptomics). A few of the applications of DNA markers to wildlife population biology include determination of the hybrid status or taxonomic affiliation, identifying cryptic species and individuals, estimating abundance and connectivity, quantifying genetic variation and parentage, and determining genes that may affect fitness in the real world (we saw examples of coat color and adaptation to high elevation).

With this background information on genetic markers and how they can help us understand wildlife population ecology, we are well poised to consider other genetic applications throughout this book, including estimating abundance and reproduction, measuring isolation and connectivity of populations, predicting inbreeding depression, detecting diseases, and solving forensic cases.

It would have been impossible 25 years ago to imagine the insights into wildlife population biology that genetic markers could provide. Surely the same will be true for the next 25 years. As a result, more than ever, wildlife biologists must understand basic genetic concepts and tools, and be dedicated to collect and archive genetic samples whenever studies involve handling animals. Because some challenges remain in the analysis of poor quality DNA inherent in noninvasive samples, the most prudent approach is to archive higher quality samples from nondestructive sampling (e.g. blood or ear tissue punches) when animals are handled, instead of just relying on noninvasive sampling. These archived samples will provide baselines, for example, in assaying future changes in abundance or connectivity (say, in parks), and for building databases crucial for forensics cases.

Many of these techniques and applications are in their infancy. The burgeoning of these applications

means that inevitable mismatches will occur between technique, analysis, and application. That is why the strongest applications of genetic tools continue to be those accompanied by demographic, ecological, and field data.

FURTHER READING

Allendorf, F.W., Luikart, G., and Aitken, S.N. (2012) *Conservation and the Genetics of Populations*, 2nd edn. Wiley Blackwell, Oxford. An outstanding textbook covering all facets of conservation genetics.

Avise, J.C. (1994) *Genetic Markers, Natural History, and Evolution*. Chapman and Hall, New York. Although markers have changed over the last decade, this book remains a classic explanation of fundamental genetic concepts and techniques.

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., and Blaxter, M.L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews: Genetics* **12**, 499–510. A nice entry into a set of techniques that are changing remarkably fast, with great potential application to wildlife population ecology.

Oyler-McCance, S.J. and Leberg, P.L. (2012) Conservation genetics and molecular ecology in wildlife management. In: *The Wildlife Techniques Manual: Research*, 7th edn (ed. N.J. Silvey), pp. 526–46. John Hopkins University Press, Baltimore, MD. An up-to-date and complete reference on the application of genetics in wildlife population ecology.

Palsbøll, P.J., Allen, J., Bérubé, M., Clapham, P.J., et al. (1997) Genetic tagging of humpback whales. *Nature* **388**, 767–769.