CHAPTER 3

Wildlife Genetics in Mountainous Rugged Asian Landscapes: Methods, Applications, and Examples

Jan E. Janečka¹, Paulo Alves², Dibesh Karmacharya³, Nicole Samsel⁴, Ellen Cheng⁵, David Tallmon⁶, and Michael Schwartz⁷

¹Texas A&M University, Texas, USA
²University of Porto & CIBIO, Porto, Portugal
³Center for Molecular Dynamics, Kathmandu, Nepal
⁴University of Montana, Montana, USA
⁵Ugyen Wangchuck Institute for Conservation and Environment, Bumthang, Bhutan
⁶University of Alaska Southeast, Alaska, USA
⁷USFS, Rocky Mountain Research Station, Montana, USA

Introduction

Over the last century, some of the greatest human achievements have been attained, unfortunately, often at the expense of Nature. We have tragically witnessed an ever-increasing number of extinctions, and many other species have rapidly declined (Vitousek 1997, Pimm et al. 2006). Despite these losses, there is also much to be hopeful for. In the past 50 years, dedicated conservationists have successfully recovered numerous species from the brink of extinction. Some of these, such as the American bison (Bison bison), gray wolf (Canis lupus), and golden lion tamarin (Leontopithecus rosalia), are now abundant in many areas (Kierulff and Rylands 2003, Freese et al. 2007, Wayne and Hedrick 2010). The paradigm of uninhibited progress has started to shift towards sustainable development. In addition, more tools and resources are available for conservation today than ever, making it possible to design realistic initiatives that enable preservation of our wild lands (Manel et al. 2003, Waits and Paetkau 2005, Mills 2013, Schwartz et al. 2007, Long et al. 2008, Frankham 2010).

It is important to have a balance between economic development and conservation, particularly in countries that are experiencing rapid growth, such as Bhutan, Nepal, and India (United Nations 2010). Knowledge is the foundation for such a balance. In order to guide development and effectively allocate conservation, we must first identify the species and populations most threatened by human activities, determine which factors are critical for their persistence, predict how our actions will influence them, design ways to mitigate any significant
negative effects, and monitor the success of our conservation efforts. That's a tall order. Only through carefully planned research initiatives can we hope to gain the knowledge we need.

Genetics has become a very important aspect of wildlife research and conservation. The genome is the blueprint for all life, and its interaction with the environment creates the remarkable biodiversity we see today. By exploring genomes, there are vast opportunities to understand the history of species, and in turn the evolution of life on Earth. Yet genetics is not only applicable to academic endeavors; there are also many practical applications. Because genetic variation is affected by population and ecological processes in a predictable manner, it is possible to understand life history and demographic parameters including population distribution and size, temporal fluctuations, connectivity, and social structure by interpreting patterns of genetic variation (see Table 3.1 for examples) [Kimura 1968, Schwartz et al. 2007, Hedrick 2011]. Genetics has therefore become an important component of wildlife science.

Wildlife studies are particularly difficult because of the inherent limitations imposed on research. Much of the work must be done in the field by observing animals in natural environments. The elusive nature of many species, often combined with rugged topography or thick vegetation, can make direct observations next to impossible, particularly in places like the mountainous regions of Himalayas in Asia, Alps in Europe, Rockies in North America, and Andes in South America. Complementary to the other methods discussed in this book (e.g., sign surveys, telemetry, and camera-trapping), genetics provides an additional framework to get around these challenges [Waits and Paetkau 2005, Mills 2013, Long et al. 2008, Rodgers and Janecka 2012].

One of the most useful advantages of wildlife genetic studies is that individuals can be sampled noninvasively (i.e., without direct handling). As animals move through the landscape they often leave tracks and sign behind, along with biological material including hair, feathers, feces, and urine, that were traditionally incorporated into wildlife studies with a certain level of uncertainty. Fortunately, these biological materials include remnants of cells; therefore with genetics they can be unambiguously attributed to a species, and often to a specific individual [Waits and Paetkau 2005, Kelly et al. 2011]. In this way, animals can be “directly” observed in their natural environment, without visual contact or capture. In contrast to visual detection by sight or via remote sensing cameras, a physical DNA sample is obtained to generate additional information on the population.
However, genetic information by itself is not very informative. The data generated on individuals and populations must be put into a biological context. The greatest benefit from genetics occurs when such studies are combined with more traditional methods including sign surveys, camera trapping, and telemetry. By using multiple complementary approaches we can better understand the status of species, how they relate to each other and their environment, and what factors may be influencing their persistence. Specifically, we can glean information on the biodiversity of an area, the distribution of species, the status of populations, landscape connectivity, and dispersal (Manel et al. 2003, Schwartz et al. 2007). More in-depth studies can even provide information on social structure and relatedness (Widdig et al. 2002, Jedrzejewksi et al. 2005, Janecka et al. 2006, Honer et al. 2007, Wagner et al. 2007). Every year, there are new developments in genetics along with a reduction in costs, facilitating such studies (Thomson et al. 2010).

As with all research approaches discussed in this book, the first and most critical step in a genetic study is to identify the specific questions to be addressed. The next step is to consider how genetics can be used effectively to answer these questions. Finally, the study design must be carefully considered so that the questions being asked can be realistically answered. Researchers have to decide how samples will be collected, if molecular markers are available for the species of interest, and how and where the genetic data will be generated. And finally, the entire project needs to be carefully reviewed in the context of the original objectives to confirm whether a genetic approach is indeed an effective allocation of resources.

This chapter is meant to provide a brief introduction to the world of wildlife genetics with a focus on noninvasive methods. We discuss where the information comes from, how it is generated, and what questions can be answered. Finally, we provide examples and additional information in appendices that can be used to develop a pilot population survey. We encourage the reader to explore other excellent books and review articles that focus on this topic (Waits and Paetkau 2005, Mills 2013, Long et al. 2008, Hedrick 2011). We hope that this chapter serves as a primer to generate additional interest for wildlife genetics in Asia, and other mountainous regions of the world.

**Overview of Genetics: from DNA to Molecular Markers**

**DNA structure**

All organisms, with the exception of some viruses, use DNA (deoxyribonucleic acid) as their genetic material. DNA has two essential roles: coding for and maintaining
the cell components via proteins and regulatory RNA molecules, and passing on the information to offspring.

A DNA strand is a relatively simple molecule composed of subunits called nucleotides. There are four types of nucleotides, each consisting of a phosphate molecule attached to one of 4 different bases: adenine (A), guanine (G), cytosine (C), and thymine (T). Within cells, these nucleotides are attached together in long chains, to form two strands that spiral round each other creating a double helix. The backbone of this helix consists of the phosphate molecules. The bases connect the opposing strands via hydrogen bonds (similar to the rungs of a ladder) in a specific pattern, with A only pairing to T, and C to G. Contiguous DNA strands are formed by millions of nucleotides that are usually arranged in structures called chromosomes.

Three fundamental DNA-related processes occur in cells: DNA replication, transcription of DNA into RNA, and translation of RNA into proteins. Replication is essential for passing on genetic material to other cells and to offspring, and transcription/translation for cellular development and processes. DNA replication occurs in the nucleus and mitochondria, and starts with the separation of the two strands (denaturation), which afterwards serve as templates that are also copied. With this process, the genetic information is maintained in each new cell.

The DNA in the nucleus is organized into chromosomes, which contain genes, noncoding regions, and structural elements. Genes were traditionally defined as segments of DNA that coded for proteins; however, this has been extended to include other regulatory molecules such as microRNAs. Genes ultimately determine how, when, and where proteins are made. Proteins are polymers of amino acids, the type and order of which define their form and function. These proteins are determined by the nucleotide sequence in the exons, which are composed of sets of three nucleotide bases (called a codon) that code for a specific amino acid. Proteins are synthesized during the process of transcription of DNA into RNA, followed by translation of the RNA into the respective amino acids that compose them.

Most genes consist of sequences of nucleotides that contain alternating sections of coding and non-coding regions (exons and introns, respectively). All cellular processes and structures are mediated, either directly or indirectly, through proteins and RNA molecules. However, the great majority of DNA within a genome does not code for proteins or regulatory molecules, and to date much of it has no known function. Only about 2% of genomes contain instructions for the synthesis of
proteins and other regulatory molecules. Genes are positioned in the chromosomes at specific physical locations called loci (locus in singular).

During DNA replication, transcription, and translation, occasional errors produce permanent changes in DNA or RNA sequences, these are known as mutations. If a mutation occurs during DNA replication in a gamete (a reproductive cell that fuses to form a zygote, i.e., egg and sperm) it is passed on to the offspring. There are several types of mutations including substitutions (called single nucleotide polymorphisms, SNPs) and insertion/deletion of nucleotides (called indels). Mutations accumulate over time leading to the variation (also called polymorphism) in genomes we see today (Kumar 2005, Schwartz et al. 2007).

Mutations can also occur on larger scales, with entire sections of chromosomes being duplicated, deleted, or translocated. Wildlife genetics typically examines SNPs or indels. A genetic variant arising from a mutation in a specific position on a chromosome (referred to as a locus), is called an allele. Loci that have variation (i.e., more than one allele) can be used as molecular markers for assessing species diversity, evolution, population structure, individual identification, and relatedness (Avise 2004).

Mutations can occur in either genomic regions that influence gene function or in areas that do not, and therefore are considered neutral. Mutations in functional regions of proteins can change the amino acid sequence or expression patterns: If these alterations are dramatic and negatively affect biological processes, they can be lethal and/or cause diseases. In rare cases, they can be beneficial to an individual. The majority of mutations are neutral and do not affect fitness (Kimura 1968). Distribution and frequency of neutral alleles (the different variants) within and among individuals is largely determined by population parameters and processes such as the number of individuals, population size and fluctuations, selection on nearby loci, and migration (Avise 2004, Hedrick 2011). Therefore in wildlife genetics neutral variation often is used to understand populations.

**Assaying genetic variation**

Genetic information is inherited in four different ways depending on where in the genome it is located. For mammals, the genes on the Y-chromosome (~58 million nucleotides in humans) are passed only from the father to the male offspring. Those on the X chromosome (~160 million nucleotides in humans) are passed through both males and females, although males have only one copy. The majority of all other genetic material is located on autosomes (non-sex chromosomes, ~3 billion nucleotides in humans) and has bi-parental transmission (i.e., both maternally and paternally). Finally, mitochondria have their own genetic material
(mtDNA) organized in a small circular genome (~16,000 nucleotides in humans) only inherited through females. Because of differences in the size, content, structure, and modes of inheritance between the X, Y, autosomes, and mtDNA, their molecular markers have different applications.

Making sense of all the variation present in a ~3 billion nucleotide genome is very challenging. Until the latter part of the 20th century there were many technological limitations for applying genetics to wildlife. A revolution occurred after the invention of the Polymerase Chain Reaction, or PCR, and the use of a thermostable DNA polymerase enzyme enabling direct amplification of specific molecular markers (Saiki et al. 1988). The DNA polymerase enzyme used in the PCR was first discovered in a bacterium *Thermophilus aquaticus* in the hot springs of Yellowstone National Park (USA) (Antunes et al. 2008). All such enzymes are generically referred to as *Taq* (Chien et al. 1976). A PCR yields millions of copies of a targeted segment of DNA, achieving sufficient quantities for downstream analyses such as visualization on an agarose gel, sequencing, and genotyping (see Box 3.1).

**BOX 3.1. Overview of the Polymerase Chain Reaction**

In addition to the *Taq* polymerase, a PCR requires two short DNA sequences (called **primers**) with a nucleotide sequence complementary to the beginning and end of the targeted DNA segment (the molecular marker). To perform a PCR, the DNA from an organism (referred to as template) is added to a tube that contains primers, nucleotides (building blocks of DNA), *taq* polymerase, MgCl₂, and other salts and additives. The mixture is then placed in a thermocycling machine that can be programmed to increase and decrease the temperature of the samples.

First, the mixture is heated to separate the double-stranded DNA template into single strands (denaturation). Then it is cooled to allow the primers to bind to the DNA template (annealing). After annealing, the polymerase begins to synthesize new strands of DNA starting from the primers (extension). At the end of the first cycle, each double-stranded DNA molecule consists of one new and one old DNA strand. The new copies are used as templates in subsequent cycles. There is an exponential increase in the targeted DNA segment; after 20-25 cycles yielding millions of new copies of the particular molecular marker.
**Molecular markers**

One of the first steps in undertaking a genetic study is identifying the appropriate molecular markers for the population of interest. Molecular markers fall into two broad categories: short tandem repeats (STRs) and single copy gene segments. Short tandem repeats are also called microsatellites, and consist of repeats of a DNA sequence motif (typically 2–4 bases) that are surrounded by conserved flanking regions (Järne and Lagoda 1996, Goldstein and Pollock 1997, Ellegren 2004). For instance, PUN100 microsatellite consists of 18–23 repeat units of “AC” in snow leopards (although additional alleles of different length may be found in unstudied populations) (Janecka et al. 2008b).

The number of repeats in a microsatellite is variable because DNA polymerase sometimes slips during replication of repetitive sequences, adding one more or one less repeat unit (Ellegren 2004). This type of mutation is more common than a substitution or indel in non-repetitive DNA segments (Järne and Lagoda 1996). Microsatellites therefore tend to be more polymorphic than regions with unique sequence. Alleles in microsatellites are defined by the number of repeats, which can be inferred from the size of the PCR amplicons. It is important to note that the size differences for most microsatellites are too small to resolve on an agarose gel and must be differentiated with a sequencing instrument (e.g., Applied Biosystems 3730xl DNA Analyzer).

If enough variable microsatellites are genotyped (i.e., the alleles present at each locus are identified) each individual in a population will have a different combination of alleles (Waits et al. 2001). Therefore, the composite genotype (i.e., the combined genotypes of multiple microsatellites) can be used as a DNA “tag” to identify individuals (Figure 3.1) (Taberlet et al. 1997, Kohn et al. 1999, Waits et al. 2001). This approach has been used in many wildlife studies (Tallmon et al. 2002, Tallmon et al. 2004, Bhagavatula and Singh 2006, Schwartz 2009, Schwartz et al. 2009, Janecka et al. 2011a). We can also use the allele frequencies in the population to describe variation, structure, connectivity, hybridization, and to estimate parameters such as effective population size (Järne and Lagoda 1996, Waits et al. 2001, Janecka et al. 2008c, Oliveira et al. 2008, Oliveira et al. 2010, Janecka et al. 2011b). Due to their high levels of variation and abundance in the genome microsatellites are currently among the most widely used class of molecular markers in wildlife research.
Figure 3.1. An example of the rationale behind individual identification of unknown scat samples using microsatellites. Allele sizes are obtained by analyzing electropherograms of PCR amplicons from a sequencer (e.g., ABI 3730). The inset shows an image of the genotype for microsatellite Locus 2 of Scat 09; note that the individual is heterozygous for 92 and 96 base pair alleles. Scats that have the same genotype at all loci (i.e., have the same genetic "tag") are grouped together and assigned to the same individual. For example, the "tag" 118/124-96/100-167/167 was assigned to Snow Leopard 1. Some scat samples will have missing data (e.g., Scat 08 and Scat 24) and these should be potentially removed from the dataset or re-analyzed.

The second broad class of molecular markers consists of single copy genes. Informative SNPs can be identified by DNA sequencing (Sambrook and Russel 2001, Avise 2004). For example, there are numerous SNPs in specific positions of the cytochrome b mtDNA that differentiate the snow leopard and common leopard, and other sympatric species (Figure 3.2) (Janecka et al. 2008b). If these differences are fixed between the species, they can be used for species identification. DNA sequencing of many gene segments across many animals can generate SNP information that can be analyzed to understand population structure and processes (Manel et al. 2003, Avise 2004, Kohn et al. 2006). This approach can be extended to understand the evolutionary history of species and their phylogenetic relationships (Kohn et al. 2006, Janecka et al. 2007, Janecka et al. 2008a, Davis et al. 2010, Eizirik et al. 2010).
Figure 3.2. Phylogenetic approach to species identification. A small section of cytochrome b or another gene segment can be sequenced, aligned, and a tree reconstructed. Sequences from unknown scat that closely group with reference sequences and show >97% similarity can be assigned to a species. However, results must be interpreted with caution particularly in areas such as Bhutan, Nepal, and Tibet, where many species occur for which there is currently no reference sequence. Data from Janečka et al. 2008b.

**Getting DNA – From the Field to the Laboratory**

**Sampling wildlife populations**

Before embarking on a wildlife genetics study one must critically assess what kind of information is sought. The goals of the project need to be carefully identified and clearly outlined. One must also contemplate the predicted end point, and consider whether the outcomes will contribute to conservation and management. As with all endeavors “Begin with the end in mind” (Covey 1989). Never is this more applicable than in wildlife research. Think about the potential results from a successful outcome. If they would make substantial contributions towards your conservation or management goals, then proceed to develop a strategy for executing the project.

There are many kinds of concepts that can be addressed using genetics. These include intra-population (e.g., status, distribution, abundance, and trends), inter-populations (e.g., dispersal, gene flow), ecological (e.g., habitat use, predator-prey interactions), and evolutionary (e.g., describing new species, relationship between
species, adaptation to environments) (Table 3.1). Are you trying to map the
distribution of a species in an area? Will you examine its taxonomic status? Do you
want to get abundance estimates? Are you focusing on specific populations, or do
you want regional information? What kind of sampling will be most effective for
the species? Answering these questions prior to your study is critical because it will
dictate not only the type and duration of sampling, but also the molecular markers
and analyses that you will use (Figure 3.3). In some cases, you may decide that
more traditional methods (e.g., mark-recapture, telemetry) may be more
appropriate.

Table 3.1. Examples of studies that illustrate the application of molecular markers to understanding
wildlife populations.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Species</th>
<th>Sampling</th>
<th>Molecular Marker</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot study</td>
<td>Snow leopard</td>
<td>scat</td>
<td>cytochrome b, microsatellites, AMELY</td>
<td>Janečka et al. 2008b</td>
</tr>
<tr>
<td>Diet</td>
<td>Snow leopard</td>
<td>scat</td>
<td>cytochrome b, microsatellites</td>
<td>Anwar et al. 2011</td>
</tr>
<tr>
<td>Occupancy</td>
<td>Fisher</td>
<td>scat</td>
<td>16s, microsatellites</td>
<td>Zielinski et al. 2006</td>
</tr>
<tr>
<td>Abundance</td>
<td>Snow leopard</td>
<td>scat</td>
<td>cytochrome b, microsatellites</td>
<td>Janečka et al. 2011a</td>
</tr>
<tr>
<td>Connectivity</td>
<td>Lynx</td>
<td>tissue</td>
<td>microsatellites</td>
<td>Schwartz et al. 2002</td>
</tr>
<tr>
<td>Effective population</td>
<td>Ocelot</td>
<td>blood</td>
<td>microsatellites</td>
<td>Janečka et al. 2008a</td>
</tr>
<tr>
<td>size</td>
<td></td>
<td></td>
<td>multiple nuclear and mtDNA gene segments</td>
<td></td>
</tr>
<tr>
<td>Species delineation</td>
<td>Colugos</td>
<td>hide</td>
<td>multiple nuclear and mtDNA gene segments</td>
<td>Janečka et al. 2008c</td>
</tr>
<tr>
<td>Species identification</td>
<td>Carnivores</td>
<td>various</td>
<td>nuclear SNPs</td>
<td>Oliveira et al. 2010</td>
</tr>
<tr>
<td>Hybridization</td>
<td>Wildcat</td>
<td>tissue</td>
<td>microsatellites</td>
<td>Oliveira et al. 2008</td>
</tr>
<tr>
<td>Relationship among</td>
<td>Felids</td>
<td>blood</td>
<td>multiple nuclear, Y, X, and mtDNA gene segments</td>
<td>Davis et al. 2010</td>
</tr>
<tr>
<td>species</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If genetic methods are indeed appropriate for achieving the goals of your project, the next important step is to assess whether the necessary molecular resources are available for your targeted species (Figure 3.3). The best way to do this is to search the literature for closely related species that have already been studied, ideally within the same genus, although many molecular markers will work across an entire family. For example, microsatellites developed in the domestic cat can be applied to other felids, such as snow leopards, tigers, and leopards (Menotti-Raymond et al. 2003, Bhagavatula and Singh 2006, Murphy et al. 2007, Janecka et al. 2008b). If markers are not available they can be designed. However, this requires additional time and expenses, along with close collaboration of a reputable genetic laboratory (Glenn and Schable 2005). At this stage, a laboratory where the genetic analysis will be completed should be identified.

The next critical aspect that will strongly influence the success of a project is the design of an appropriate survey scheme (see Appendix 3.1 for an example) (Long and Zielinski 2008, Schwartz and McKelvey 2009). In order to determine how many samples to collect from an area, and how they will be collected, one must again go
back to the goals of the project as they will influence the study design to be implemented. For example, different population state variables described in Chapter 2 (e.g., distribution, abundance, and trend) require different sampling protocols.

In most cases, the study area will be divided into a grid (Krebs 1998). This is then used to delineate sampling blocks. The blocks need to be of the appropriate size; it is common practice to make them roughly the same size as the mean home range of the target species (Long et al. 2008). Each block should be sampled if it is logistically feasible. However, if resources are limited, then the blocks can be selected based on a stratified sampling design (Quinn and Keough 2002). This means that they are randomly sampled in proportion to the available habitat in that area. Within each block, sites are chosen where the likelihood of detecting the target species is highest (Figure 3.4). The blocks sampled and the specific locations of transects will be largely influenced by physical access and the resources available for the survey, particularly in rugged, mountainous areas. The intensity of sampling is a very important variable to consider. It depends on capture probability, the objective (e.g., occupancy versus abundance), and the method used to estimate the parameters (MacKenzie et al. 2002, Wintle et al. 2004).

**Figure 3.4.** Examples of the type of sign used to identify sites that have high snow leopard activity within a sampling block. This information should be used to increase the detection probability during a noninvasive genetic survey. Photo credits J. Janečka.
In many regions of Asia, there is little information on the home range size and capture probability of species, making it difficult to design the most effective sampling strategies. In these cases, one must base the first survey on studies carried out in similar ecosystems and/or on closely related species, and also incorporate local knowledge. An initial population survey should be done, and subsequently the block sizes and sampling techniques adjusted based on the preliminary data. Even for well-studied species, survey strategies should be constantly reassessed and optimized to produce ever more reliable data (Boulanger et al. 2006).

The two primary methods for noninvasive genetic sampling of individuals include collection of scat and snagging of hairs (Waits and Paetkau 2005, Beja-Pereira et al. 2009, Kelly et al. 2011). Scat can be collected on wildlife trails and travel corridors for species including wolves, coyotes, bears, otters, elephants, bobcats, tigers, and snow leopards (Kohn et al. 1999, Eggert et al. 2003, Janecka et al. 2008b, Hajkova et al. 2009, Marucco et al. 2009, Mondol et al. 2009, Ruell et al. 2009). For some animals, such as snow leopards, which occur in dry, cold regions and leave scat in highly visible sites, it is an excellent method (Janecka et al. 2008b, Janecka et al. 2011a). Snow leopards and other carnivores often leave multiple scats at one site on distinct landscape features, such as saddles and outcrops, increasing the efficiency of sampling (Figure 3.5). Please refer to Appendices 3.1–3.5 for examples of scat survey methods, sampling techniques, and data collection sheets. Some species occupy habitats with dense vegetation, dung beetles, or hot, humid climates, and leave scat in sites with poor visibility, making it difficult to collect sufficient numbers of quality scat samples, therefore other techniques may be warranted.
Figure 3.5. A saddle in the Gobi Desert of Mongolia that was on a sampled snow leopard scat transect. There were a total of 7 scats (blue circles) and 3 scrapes (green arrows) observed on the saddle. Sites like this are important to focus on during a survey in order to obtain sufficient numbers of noninvasive genetic samples. Photo credit J. Janečka.

In some cases hair collection may be the preferred means of sampling populations. One way that hair can be collected in areas with snow cover is to backtrack on fresh pugmarks, carefully searching for sites where the animal bedded down or rubbed against rocks or vegetation. Hair snares are an alternate option for some species. These “traps” are often either constructed out of barbed wire surrounding a container of bait, or carpet pads with roofing nails attached to a tree marked with scent lures that elicit rubbing behavior. Effective hair traps have been designed for lynx, bears, foxes, martens, fishers, and numerous other species (McDaniel et al. 2000, Poole et al. 2001, Bremner-Harrison et al. 2006, Schmidt and Kowalczyk 2006, Zielinski et al. 2006, Downey et al. 2007, Pauli et al. 2008).

When considering optimal sampling strategies one must bear in mind that there are species, population, behavioral, and site differences that will impact their effectiveness (Gompper et al. 2006, Long et al. 2007, Ruell and Crooks 2007, Ebert et al. 2010, Ralls et al. 2010). If no work has been done on your target species in habitat similar to where you are surveying, then sampling methods must be carefully tested before resources are allocated to a large-scale survey. Focus
transects on sites where there is a relatively high probability of detecting the target species. These include wildlife trails that are at the intersection of travel routes and landscape features, and have a high number of sign on them (i.e., tracks, scent rocks, claw rakes and scats, Figures 3.4 & 3.5).

Noninvasive samples (e.g., hair and scat) often have low quality and quantity DNA, making them very susceptible to contamination and genotyping errors (Taberlet et al. 1999, Waits et al. 2001, McKelvey and Schwartz 2004b, a, Pompano et al. 2005, Waits and Paetkau 2005, Rodgers and Janecka 2012). To minimize downstream problems it is essential that samples in the field are collected in a clean manner and properly stored. Various preservation methods (e.g., silica desiccant, ethanol, freezing) have been used with varying success; no technique has been found to be greatly superior (Frantzen et al. 1998, Murphy et al. 2000, Bubb et al. 2010). Often, the most practical method for the field is to store samples dry on silica desiccant (Appendices 3.2 & 3.3). Regardless of the technique used, the key is to minimize handling of samples and to maintain them in a stable environment. It is also critical to record information associated with each sample at the time of collection, including geographic location (GPS preferred), approximate age, size, nearby animal sign, vegetation types, landscape features, and any noteworthy observations (Appendices 3.4 & 3.5).

Population studies often produce large numbers of noninvasive samples (100s to 1,000s) and genetic data. Excellent organization is one of the most important means of preventing data errors, lost samples, and missing information. As soon as samples are collected, they need to be shipped to a central location. A specific person should be assigned to log samples into a collection archive and enter information into a database to prevent samples and information from being lost. Many field projects and laboratories have a frequent turnover in technicians; therefore it is critical that the principle investigator makes certain that all samples and data are being properly stored in a secure area.

**Obtaining DNA from samples**

Once the samples are archived in a laboratory the first step in genetic analysis is the extraction of DNA free of impurities and any contaminants (from other samples or even field personnel) (Figure 3.6). A good quality DNA sample ensures the analysis yields accurate data on species, gender, and individual identification (Taberlet et al. 1999, McKelvey and Schwartz 2004a, Pompano et al. 2005). If the DNA is not properly extracted, downstream errors will jeopardize conservation initiatives and management plans, and will ruin the reputation of laboratories. The most common causes of error in the lab include poorly trained technicians, personnel that do not
have an invested interest in the project, rushing through protocols, and taking shortcuts to save money. It is critical that laboratory technicians care about the results, take their time, carefully follow the protocols, and practice excellent laboratory techniques.

Figure 3.6. Flow chart of genetic analysis of noninvasive samples.

Cross-contamination of samples (i.e., DNA from one sample being present in the extraction of another sample) is a major threat with non-invasive samples. The reason these samples (e.g., hair and scat) are prone to contamination is that the DNA from the target animal is of low quantity and quality, and therefore even trace contamination from another source interferes with the analysis (Taberlet et al. 1999). In contrast, when working with DNA from tissues, there is enough DNA from the source individual that trace contaminants are usually not detected during the analysis. The elimination of cross-contamination starts in the field with careful sample collection and continues in the lab with meticulous handling during the extraction process. The use of barrier tips and manipulation of noninvasive samples

The process behind different DNA extraction protocols is similar even though individual steps may differ (Sambrook and Russell 2001). In the beginning of an extraction, cells and their nuclei are lysed, releasing DNA into solution (Sambrook and Russell 2001). The cellular membranes are further broken down and proteins degraded (cut into small pieces with proteinase K). Subsequently, the DNA is separated from other cellular components and inhibitors. One efficient method (used by Qiagen) is to bind the DNA to a silica membrane, wash off the cellular debris, contaminants, and residual impurities, and then elute the DNA using a buffer. Extracted DNA is stable at 4°C for months, but it is best to maintain samples at -20°C for long-term storage. Several companies manufacture kits specifically for DNA extraction and tailor them to different types of samples (e.g., muscle, blood, buccal cells, etc.). To date, the QIAamp DNA Stool Mini Kit (Qiagen) is the most widely used product for scat. It is designed specifically for removing impurities in feces that could interfere with downstream analysis. Other techniques are available for hair samples (Walsh et al. 1991, Goossens et al. 1998, Vigilant 1999, Suenaga and Nakamura 2005, Björnerfeldt and Vila 2007).

The quality and quantity of the extracted DNA needs to be determined after the extraction. The simplest method is to use an agarose gel (Sambrook and Russell 2001). Agarose can be envisioned as a matrix of molecules with spaces (or “holes”) between them (similar to a sponge). When an electrical current is applied, the DNA, which is negatively charged, moves through these “holes” towards the positive electrode. The DNA can be visualized by staining the gel in ethidium bromide (or a comparable stain such as SYBR green), which binds to DNA and fluoresces (i.e., glows) when exposed to ultraviolet light (λ = 312 nm). The DNA quantity is proportional to the intensity of fluorescence, and the DNA fragment length to the rate it migrates through the gel. Therefore an agarose gel provides information on the amount and integrity of the DNA obtained from a sample. (CAUTION: EtBr is a very strong carcinogen and UV is damaging to the eyes. Safer options such as SYBR green are available).

There are several excellent manuals that provide a step-by-step guide to DNA extraction, gel electrophoresis, and other important protocols (Sambrook and Russell 2001, Ausubel et al. 2002, Barker 2005). These should be closely studied to further understand these methods. Many details could not be covered in this chapter due to space limitations.
Application, Use, and Interpretation of Genetic Data

Measuring genetic variation
Genetic research explores the vast amounts of variation present in the genomes of individuals, populations, species, and higher-level taxonomic groups (Avise 2004). The interaction of genes and the environment leads to phenotypes (or physical traits) that characterize an organism. The genetic component is the heritable portion of these phenotypes. The influence of both natural selection and neutral processes upon genomes leads to the evolution of species.

The ultimate sources of genetic variation are mutations that occur during meiosis, as discussed above. Genetic variation of an individual is manifested in the number of alleles (i.e., variants) present across all loci. In population genetics we describe this variation by sampling a subset of loci, and measuring the number, distribution, and type of alleles or haplotypes (a series of linked alleles) (Hedrick 2011). At the population level, variation is described as percent of polymorphic loci, the mean number of alleles per locus, mean heterozygosity, and allele and haplotype frequencies (Hedrick 2011). Models have been developed that can be used to infer population parameters (e.g., effective population size and gene flow) based on the patterns observed in the genetic variation within and among populations (Avise 2004, Hartl and Clark 2007). In this section we will discuss how DNA polymorphisms are used to make species, sex, and individual identification, and to obtain other types of population information.

Species identification
Species identification is essential for noninvasive studies of wildlife as many samples collected in the field are of unknown origin. Previous studies have found high error rates in species identification of scat in the field, even by experienced biologists (Reed et al. 1997, Farrell et al. 2000, Davison 2002, Janecka et al. 2008b). It is also difficult to differentiate hairs and other tissues collected in the field or confiscated from wildlife markets (e.g., meat, claws, hides, bones, and horns) (Baker et al. 2007, Alacs et al. 2010).

Mitochondrial loci, most commonly cytochrome b and cytochrome c oxidase subunit 1, are favored for species identification because they are easier to PCR amplify in poor quality DNA samples. Primers are available for PCR amplification of portions of these genes across a wide range of taxa (Farrell et al. 2000, Hebert et al. 2003). The PCR products generated for unknown samples can be sequenced, aligned, and compared with sequences from known species (Figure 3.2) (Farrell et al. 2000, Janecka et al. 2008b, Janecka et al. 2011a). If the locus used for species identification has been characterized for all taxa present in a study area, and there
are known fixed differences between sympatric species, cheaper PCR or restriction enzyme assays can be developed for species identification (Mills et al. 2000b, Fernandes et al. 2008, Roques et al. 2011). In some populations or species, mtDNA can provide erroneous results due to introgression and nuclear mitochondrial translocations (numts). In these cases using nuclear markers may be necessary to differentiate closely related species (Alves et al. 2008).

**Gender identification**

Determining gender is important for inferring identity, examining social dynamics, and understanding the reproductive potential of a population. In mammals, gender identification is based on determining the presence of the Y chromosome. Primers are available for many species that amplify a small portion (100–400 bp) of male-specific genes on the Y chromosome (Pilgrim et al. 2005, Waits and Paetkau 2005, Janecka et al. 2008b). The results of the PCR amplification can be analyzed on an agarose gel, and presence or absence of the correct male-specific amplicon is used to determine the gender of the sample.

The results must be carefully interpreted because many noninvasive samples are degraded or have low amounts of DNA, which can lead to PCR failure. Therefore, it is important to determine whether lack of amplification of the Y-marker is due to the absence of the Y chromosome (i.e., female), or to PCR failure (i.e., degraded male sample). This can be done by repeating the PCR assay and testing the sample for PCR amplification of another locus on the X or on an autosome (e.g., any of the microsatellite loci used for individual identification). Gender assays based on amelogenin, zinc-finger Y, and sex-determining region Y loci have been developed for carnivores, ungulates, and other taxa (Palsboll et al. 1992, Griffiths and Tiwari 1993, Taberlet et al. 1993, Kurose et al. 2005, Statham et al. 2007, Janecka et al. 2008b, Kim et al. 2009, Rodgers and Janecka 2012).

**Individual identification**

The most common markers used for individual identification are microsatellites in the nuclear genome. Each locus is PCR amplified using one primer labeled with a fluorescent dye and the amplicons are then genotyped (i.e., exact size of alleles determined) with a sequencing instrument (e.g., Applied Biosystems 3730xl DNA Analyzer). There can be differences in allele sizes due to instrumental variation, the type of taq used, the dynamics of the PCR reaction, and interpretation of the raw data. For these reasons, careful steps must be taken to ensure the alleles are sized consistently across samples.
Allele sizes are not directly comparable across studies and laboratories; a subset of the same samples must be run in both labs to calibrate the alleles before genotype data can be combined. Sequencing instruments are very expensive (>US$100,000) and therefore there are few laboratories that can obtain such instruments. Fortunately, PCR amplicons are not restricted under CITES, so in most cases they can be shipped internationally to either collaborators or commercial service labs.

Microsatellite loci vary in the level of variation, PCR amplification failure, allele dropout, and false alleles. These factors affect both the precision and accuracy of the data, primarily when using noninvasive sampling (Selkoe and Toonen 2006). Genotyping errors in noninvasive studies stem from three primary sources: (i) Allele dropout that causes heterozygous individuals to be mistyped as homozygous, (ii) False alleles caused by primers annealing in areas of the genome outside of the target microsatellite locus, and (iii) False alleles resulting from small amounts of contaminating DNA (Paetkau 2003, Bonin et al. 2004, McKelvey and Schwartz 2004a, Pompanon et al. 2005, Miquel et al. 2006, Broquet et al. 2007, Zhan et al. 2010). It is critical that noninvasive studies quantify genotyping error and estimate levels of variation in the study population and then select loci and use protocols that minimize errors to a level that will not significantly affect population estimates (Taberlet et al. 1999, Mills et al. 2000a, Paetkau 2003, McKelvey and Schwartz 2004a, b, Pompanon et al. 2005, Rodgers and Janecka 2012).

The probability that two individuals will share the same alleles at a microsatellite locus and therefore have the same genotype (i.e., the probability of identity \(P_{id}\)) is determined by the frequency of the alleles in their population (Paetkau and Strobeck 1994, Waits et al. 2001). The more loci that are compared between individuals, the smaller the \(P_{id}\) for a composite genotype. If sufficient numbers of microsatellites are examined (usually 3 to 10, depending on the number of alleles) each individual in the population will have a unique genotype and therefore can be individually “tagged” (Mills et al. 2000a) (Figure 3.1). The tradeoff however is that increasing the number of loci will also increase cost, time, and the likelihood of error. The accepted criterion in most wildlife applications is to analyze enough loci to achieve a \(P_{id} < 0.01\) (Mills et al. 2000a, Waits et al. 2001).

**Infectious diseases**

Infectious diseases and parasites play an important role in wildlife populations, from influencing host diversity to altering species composition in ecological communities (Smith et al. 2009, Hedrick 2011). Pathogens can cause massive epidemics; small, endangered populations with reduced diversity are particularly sensitive to such outbreaks and should be closely monitored (Penn et al. 2002,
Acevedo-Whitehouse et al. 2003, Charpentier et al. 2007). Genetic methods can be used to detect the prevalence of pathogens. For example, elephant endotheliotropic herpesvirus (EEHV1) is a major threat in parts of Asia and Africa (Ossent et al. 1990). The presence of this virus can be monitored by testing trunk washes with an EEHV1-specific PCR assay (Stanton et al. 2010). Sequencing the genes from the virus or bacteria that is causing an outbreak can also provide information on the strain, virulence, morbidity, and dispersal through the population.

Population Inferences from Genetic Data

Abundance

Noninvasive genetic surveys provide detailed information on the distribution and abundance of individuals within an area (Figure 3.7). There are now numerous rigorous approaches that can use this data to estimate population size ($N$) (Luikart et al. 2010, Luo et al. 2010, Proctor et al. 2010, Stengelein et al. 2010, Stetz et al. 2010, Tallmon et al. 2010). Emphasis has been placed on complementary development of both field techniques and mark-recapture models. The fundamental approach to $N$ estimation is to “capture”, “mark”, and “recapture” individuals over the course of multiple sampling occasions, and then to analyze the recorded histories of these events across sampling occasions (Chapter 2). Noninvasive genetics relies upon DNA from feces or hair so that individuals are “marked” and “recaptured” using a genetic tag as described above without physically handling them (Figure 3.1). This contrasts with traditional invasive approaches based on physically marking individuals with ear tags or unique collars.

The encounter history of detections for each individual sampled can then be analyzed using an increasing number of methods and free software (Luikart et al. 2010). The most robust of these estimate $N$ from data collected over multiple sampling occasions using mark-recapture models (White and Burnham 1999, Lukacs and Burnham 2005). However, because it is sometimes logistically or financially difficult to have multiple sampling occasions reliable methods are also available for single time periods (Kohn et al. 1999, Valiere 2002, Eggert et al. 2003, Miller et al. 2005, Petit and Valiere 2006).
Figure 3.7. An example of the data that can be obtained from a noninvasive genetic survey. Yellow dots on topographic map represent scats collected along transects. Scat was analyzed using the snow leopard molecular panel for species, sex, and individual identification of Janečka et al. (2008a). The individuals observed at each transect are shown on the map. Note that F11 (female snow leopard 11) was detected on 3 transects. The inset shows the western scat transect. By focusing on ridgelines and saddles, Janečka et al. 2011b was able to detect 15 snow leopards in an 8-day sampling period covering 324 km². Data from joint survey between Texas A&M University (J. Janečka), Mongolian Institute of Biology (B. Munkhtsog), Ibis Mongolia (B. Munkhtsog), and Snow Leopard Conservancy (R. Jackson). Snow leopard inset photo-credit Snow Leopard Conservancy.

Simulation models have shown that the minimum probability of detection should be at least 0.20 for precise N estimates (Boulanger et al. 2004). Pilot studies are needed to identify the sampling design and effort necessary to obtain this level of detection. The most obvious means of increasing the probability of detection is to increase the intensity and duration of sampling occasions (Boulanger et al. 2006). There are many published examples to guide sampling designs for particular taxa (Boulanger et al. 2006). A recent review provides a nice summary of recommended steps to improve non-invasive studies (Marucco et al. 2010). In at least one case, improved study design led to substantially higher grizzly bear abundance estimates than were previously predicted for a remote, mountainous region in the US (Kendall et al. 2009).

The poor quality and quantity of DNA from non-invasive samples discussed above can lead to attributing a sample to the wrong individual, which can affect
abundance estimates (Taberlet et al. 1999, Mills et al. 2000a, Waits and Paetkau 2005). In addition to minimizing errors in the lab there are now methods that incorporate genotype uncertainty into mark-recapture N estimation (Taberlet et al. 1999, Lukacs and Burnham 2005, Waits and Paetkau 2005, Schwartz et al. 2006, Schwartz and Monfort 2008). It is important to recognize that despite the best efforts to reduce contamination and degradation of samples, genotypes frequently can only be obtained from 15% to 90% of non-invasive samples taken from the field (Marucco et al. 2010). Consequently, many samples need to be collected to ensure enough data for abundance analyses.

A single N point estimate is an important goal for a study; however, the trajectory of N provides much more insight into the status of a population. By estimating N over time, we can understand how populations respond to specific stressors, such as habitat removal or road development. Repeated non-invasive sampling provides a way to monitor population trends with minimal disturbance, yielding useful information for conservation and management (Schwartz et al. 1998, Schwartz et al. 2007, Tallmon et al. 2010). In addition, it offers an opportunity to continually improve field and lab protocols.

**Home range, parentage, mating structure**

Non-invasive genetic studies can also provide insights on wildlife behavior and space use (Figure 7). For example, based on the location of sampled individuals, home range distribution can be estimated (Taberlet et al. 1997, Kohn et al. 1999, Taberlet et al. 1999). Similarly, questions on habitat use can be addressed by incorporating geographic, landscape, and habitat information into the analysis (Long et al. 2008). Combining genetic data with radio-telemetry can provide additional insights on relatedness and social structure (Ralls et al. 2001, Tallmon et al. 2002, Widdig et al. 2002, Di Fiore 2003, Janecka et al. 2006).

**Landscape connectivity and population structure**

The applications of molecular markers can be extended by sampling multiple populations. Movement between areas has important conservation ramifications because migration plays a large role in population dynamics and can greatly increase the likelihood of persistence (Sjogren 1991, Schwartz et al. 2002, Vila et al. 2003). The patterns of genetic variation can be used to infer population structure and history; the more isolated populations are the more genetic divergence there will be across loci (Avise 2004). Various methods including assignment tests and genetic-based clustering of samples can indicate contemporary dispersal, identify migrants, and estimate gene flow (movement followed by mating) between populations. Finally, there is an emerging field of landscape genetics that examines
how genotypes are distributed across the landscape and to what extent landscape features hinder or foster connectivity (Manel et al. 2003).

From a practical perspective, the analysis of population structure and connectivity using genetic information requires random, representative samples from each population of interest. Typically, samples of 30 individuals are considered a minimum, although larger numbers (i.e., 50—100 individuals) provide greater statistical power. For rare species, such as snow leopards, this level of sampling may not be feasible. Based on some of the previous studies by the authors, samples of 10 individuals per area can be informative for population structure (Janecka et al. 2008c, Janecka et al. 2011b). Whenever possible, the sampling design used to generate single population estimates of abundance (described above) should be extended to multiple populations to address regional-level questions important for conservation and management of wildlife.

Summary
Genetics provides an important tool for both monitoring wildlife and understanding ecological processes. Noninvasive sampling has opened the horizon for the application of genetics to many wildlife species that previously could not be effectively studied. However, due to poor DNA quality and quantity, analysis must be conducted carefully to ensure errors do not lead to misinterpretation of the data. The data generated from noninvasive genetics can be combined with traditional studies to provide information on distribution and abundance, population trends, home range use, social structure, dispersal, and population connectivity. This knowledge is critical for making effective conservation and management decisions that ensure economic development occurs sustainably, without irrevocable damage to wildlife.

Glossary
Allele: Alternative form of a gene or locus that differs in size and/or composition.

Allele Dropout: The random non-amplification of one of the alleles in a heterozygous sample.

Gene Flow: Genetic exchange between populations as a result of migrants that successfully reproduce.

Genetic Drift: Random change in allele frequencies resulting from chance sampling of gametes. The process contributes to loss of genetic variation (e.g., reduction in the number of alleles) and increased divergence between populations. Its effect is greater when population size is small.
**Genotype:** The characterization of alleles present in a cell or organism. Diploid cells have 2 copies of each nuclear locus.

**Homozygous:** A locus is considered homozygous if an individual has two identical alleles at that locus.

**Heterozygous:** A locus is considered heterozygous if an individual has 2 different alleles at that locus.

**Locus:** A discrete location on a chromosome that is inherited as a unit. A locus may contain a gene; however, it may also contain no genes. Plural: loci.

**Literature Cited**


Appendix 3.1: Snow Leopard Scat Sampling Strategy

This sampling protocol was developed in Mongolia, where snow leopard habitat tends to be fragmented, rather than continuous like the situation in the Himalaya or parts of Tibet and China. The typical minimum home range size is assumed to be about 256 km², based on telemetry studies conducted in South Gobi (ISLT) and Baga Bogd (SLC). Please follow these guidelines in determining where to collect samples.

1) Obtain a 1:200,000 scale topographic map of the proposed survey region. These maps have latitude/longitude and grid lines that produce 4 x 4 km “cells”. These are used as the basis for designated survey blocks.

2) Using the existing lat/long lines, mark 16 x 16 km sized “grids” covering the entire survey region. Each of these “grids” is a square with 4 contiguous 4 x 4 km “cells” on the map. The grids have now become “sampling blocks” (each about 256 km² in size). Number each one consecutively (e.g., 1 – 16 in attached diagram). Each sampling block area corresponds roughly to the size of home ranges of snow leopards in the Gobi. Sample the blocks according to rules described below. If it is possible, sample every block. If you cannot sample every block for logistical reasons, determine the total number that you can feasibly sample during the survey, and then randomly select which blocks you will sample. Remember to avoid sampling only the best sample blocks—this is achieved by first delineating areas of polygons each with a specific habitat suitability class, such as:

- Good or prime habitat (see below)
- Poor habitat (relatively large areas of flat or slightly undulating land, including desert plains, basins with lakes, wide river valleys, expansive forest)
- Other—everything else
- Later (using a GIS or mapping planimeter), you can estimate the proportion of each sampling block that falls within each of these categories

3) On the topographic map (or on tracing paper overlaid onto the map), outline the areas within the sampling blocks that contain suitable snow leopard habitat—broken, steep, rocky or rolling hills and mountains within the landscape. Delineate polygons that contain this type of habitat by following the contour lines.
4) Name (or number) each unique snow leopard habitat polygon (e.g., A—C in attached diagram). These polygons can also be used to guide a presence/absence survey. However, in that case non-snow leopard habitat polygons would also be sampled.

5) In each sampling block, identify specific sites within the snow leopard habitat polygons that can be physically accessed, and have the highest probability of snow leopards' presence and detection. The best sites to select are saddles, outcrops, distinct ridgelines, cliff bases, and distinct drainages that funnel animal movement and/or are used as marking sites. Please refer to the SLIMS manual for more details on sites with high snow leopard activity. Use the following rules to determine the transects that will be sampled:

- In each 16 x 16 km survey block, locate 2 transects that have the greatest concentration of best sites within the snow leopard habitat polygons.
- Transects should be 2–5 km length, depending on ruggedness, access, and number of scats found.
- Transects are a minimum of 3–4 km, but no more than 10–12 km, apart, whether located within the same or in adjacent sampling blocks. Use the closest portions of two transect to estimate the minimum distance between them.
- If possible within each sampling block, have different transects sample different drainages or watersheds, or on different ridgelines.
- Sample both ridgeline and valley (drainage bottom) landforms within a sampling block.
- Collect no more than 36 scats from a particular sampling block.

6) Collected scats should be:

- Near snow leopard sign (i.e., scrape, rock spray, tracks). However, on some transects you may find very little sign. If this is the case continue to collect scats.
- On distinct landscape features such as small outcrops, saddles, bottom of cliff faces, passes between two separate valleys (watersheds).
- Fresh or very fresh. Older scats can be collected provided their surface is relatively intact. If you cannot find fresh scats, you can collect older scats.
Do not collect scats without an outer “shell” of fecal material. These scats look like the outer surface has been washed away, and they consist of only hair and small bone fragments.

7) Record information onto the datasheets. There are two datasheets to be filled out. The first is the TRANSECT DATASHEET. There is one for each transect. The second datasheet is the SCAT DATASHEET: Each individual scat collected will have a record on this sheet.

Numbering Transects: Enter the name and number for each transect. These should be numbered consecutively with initials of primary field person (e.g., BM1 for Bariusha Munkhtsog transect 1). The numbering system should be run consecutively across sites, seasons and years, in order that each transect has a completely unique number. This serves to minimize potential errors arising from duplicate names and transect numbers.

Numbering Scats: Please number all collected scats consecutively along each transect (i.e., BM1-1, BM1-2,...). Start at 1 again when you move to a new transect (ie, BM2-1, BM2-2,...). Be sure you record both the transect name/number and scat number on each collection tube.

Take a photograph of scats before they are sampled (disturbed) with a labeled sampling tube next to it. Record GPS location and other scat / habitat parameters on existing form, then place and store scats in the provided collection tubes following the standard protocol distributed earlier.

After running the transect (or back in camp) make sure to draw it on the map with the sampling blocks. It would be helpful if you were to supplement GPS scat locations with the location of major turning points along the transect, so that we can accurately map the route taken in a GIS for corridor and other spatial analyses.
Appendix 3.2: DNA Sampling Protocol – Scat

Notes

- Numerous scats believed to be snow leopard are often found together on active scrape sites, collect a sample from each intact scat found at each site.
- Do not touch, disturb, or kick, etc. the scat before you sample it.
- Handle samples with new gloves or rock/stick and then dispose of gloves immediately.
- Do not handle the scats with your bare hands.
- Once you collect a sample fill out “Sample Data Sheet”.
- Keep samples cool and dry and in shade.
- Transfer samples to an appropriate lab as soon as possible.

Scat Collection

1) Prepare a new Collection Tube with Silica Desiccant or 96% Ethanol and label it with the Date, Sample ID, and Collector’s Name / Transect number on the side of the tube and with the sample ID on top of the cap.

2) Numbering Transects: Number consecutively with initials of primary field person (e.g., BM1 for Bariusha Munkhtsog transect 1, or BM2 for transect 2). The numbering system should be run consecutively across all sites, seasons and years, in order that each transect has a completely unique number. This serves to minimize potential errors arising from duplicate names and transect numbers.

3) Numbering Scats: Number all collected scats consecutively along each transect (i.e., BM1-1, BM1-2, ...). Start at 1 again when you move to a new transect (i.e., BM2-1, BM2-2, ...). Be sure you record both the transect name/number and scat number on each collection tube.

4) Take the GPS location (decimal degrees, please) and record in Sample Data Form.

5) Fill out the rest of the Sample Data Form.

6) Put on a new pair of gloves. Or using a stone or stick to break up the scat, making sure the part touched by your fingers, never comes into contact with the part touching the scat. Use a new stone/stick for each different scat collected.
7) Break off bits and pieces from the outside part of the scat including pieces not directly in the sun (from the underside sitting on the ground). Collect scat material about the size of a pinkie nail in the tube. Do not fill the rest of the tube with scat. Do not compact scat—the scat should be loose. See Figure A and B on next page for correct amount of scat to store in tube. Close the tube and put it away. If you over fill the tube it will likely not dry and will lead to DNA degradation.

8) Dispose of gloves—keep used gloves away from sample tubes and new clean gloves, in a separate zip-lock bag. ALWAYS HANDLE EACH SCAT WITH NEW GLOVES OR WITH A PIECE OF ROCK OR STICK THAT HAS NOT BEEN USED ON A PREVIOUS SCAT.

9) Do not collect old scats lacking in a surface layer, and/or comprised only of hair, since these contain insufficient DNA for extraction.

10) Fill the form out carefully, ensuring the information you provide is complete, including:

- Species
- Scat diameter (cm), type of segmentation (see Data Code Sheet), sign age
- GPS location and site details, including any landmarks
- Relevant comments (e.g., presence of fresh scrape or rock-scent; human disturbance; scat from camera-trap site)

![Figure 3.2.1. Only a small amount of material from the outside of the scat should be collected. The total should be about the size of a pinkie nail.](image)
Author’s Note: The above Figure 3.2.2 has been corrected from the version in the book. When the figure was made for printing the “No” and X was shifted over to the left and it was not clear which tube the are referring to. The correction more clearly depicts that the tube on the right is not correctly filled.
Appendix 3.3: Field Collection of Biological Samples

DNA Sample Collection

The most desirable samples are tissue:

- Muscle
- Tongue
- Skin (i.e., ear clip)
- Blood
- Hair
- Bone

Scat samples can yield DNA but they are of lower quantity and quality.

It is essential to avoid cross contamination (between samples):

- Wash your cutting instruments and hands (or wear fresh latex gloves) between the handling of samples from different individuals
- Sterilize cutting instrument with a flame

It is essential to properly label each sample and also write your name, date, and a unique ID with something that will not rub off.

Record information:

- Species
- Sex
- Date
- Geographic location (GPS if possible)
- Nearest landmark
- Approximate age (juvenile, subadult, adult)
- The way sample was obtained (road-kill, trapped, shot, etc.)
- Any relevant comments (e.g., parasites observed, morphological abnormalities)

Fill out sample sheet.
As soon as possible deposit the sample into a DNA collection. Many samples are misplaced, lost, or degrade because they are left somewhere (a shelf, in a truck, etc.)

How to sample:

- **Soft tissue (muscle, tongue, lip, ear clip, etc.)** – remove a portion of tissue (about ¼” by ¼”) and cut up into smaller pieces. Place in vial with Longmire’s storage buffer or 96% ethanol. Make sure the tissue is submerged and floats in the tube. The sample should take up no more than 1/3 the volume of buffer. If you put too much tissue in the tube it will not preserve properly. Most sampling errors occur because too much tissue is added to the tube. It take very little tissue (1 mm x 1 mm) to yield enough DNA for analysis.

- **Blood** – place volume of blood equal to the 1/3 volume of Longmire’s storage buffer in vial (may need to remove some buffer from vial, preserve 2–3 ml of blood).

- **Hair** – remove a small tuft of hair (make sure hair bulbs will remain attached) or fragment of bone. Place hairs in coin envelope and tape shut. (store dry, in ziplock with silica desiccant)

- **Bone** – take a small fragment of bone and place in coin envelope and tape shut. (store dry, in ziplock with silica desiccant)

Emergency Method – If you find an animal and do not have sample vials remove some tissue using a flamed cutting instrument, place in a clean bag, and put on ice or freeze as soon as possible until you can preserve it properly (as described above).

**Longmire’s Lysis Buffer, 500ml**

0.1M Tris-HCl, pH 8.0, 0.1M EDTA-Na2•2H2O, pH 8.0, 0.01M NaCl, 0.5% w/vol SDS

*For Dry Reagents*

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCl (or Trizma base)</td>
<td>6.06g</td>
</tr>
<tr>
<td>EDTA-Na2•2H2O</td>
<td>18.6g</td>
</tr>
<tr>
<td>NaCl, 0.29g</td>
<td></td>
</tr>
<tr>
<td>SDS, 2.5g/300ml ddH2O</td>
<td></td>
</tr>
</tbody>
</table>

*For Reagents in Solution*

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1M Tris-HCl, pH 8</td>
<td>0.50ml</td>
</tr>
<tr>
<td>0.5M EDTA, 100ml</td>
<td></td>
</tr>
<tr>
<td>5M NaCl, 1.0ml</td>
<td></td>
</tr>
<tr>
<td>10% SDS, 5ml</td>
<td></td>
</tr>
</tbody>
</table>

Adjust pH to 8.0 and bring to 500ml (or 324ml ddH2O), store at room temp
Appendix 3.4: Transect Datasheet

Transect #: ______  Date: _____  Sampling Block: ________

Field Team: __________________________________________

GPS Location Start: __________  End: _________________

Transect Aspect and Direction: (Aspect and lat-long where the transect changed directions). This is especially critical on transects where very little scat was found.

Compass Aspect (Direction) From Start Location: (e.g., 186°) ______

Intermediate GPS Points & Aspect: ________________________________

________________________________________________________________________

Elevation: ________  Duration: _____ hrs  _____ km (approximate)

NOTE: BE SURE TO ALSO MAP TRANSECT ON THE TOPO MAP WITH SAMPLING BLOCKS

Dominant Topographic Feature (tick dominant one):
___ Ridgeline  ___ Hillside  ___ Cliff Base  ___ River/Drainage Bottom
Other: ______________________(specify)
General comments on topography:

Landform Ruggedness (tick dominant one):
___ Very Broken/Steep  ___ Moderately Broken  ___ Rolling  ___ Flat
General comments on ruggedness: (e.g., terrain changes from very broken to rolling at end of transect)

Primary Vegetation Type (tick dominant one):
___ Barren  ___ Grass  ___ Shrub  ___ Woodland
General comments on habitat: (e.g., grass only on southern slope)

Grazing Status (tick dominant one):
___ Year-Round  ___ Seasonal  ___ Non-Grazing
Type of livestock:

86
General Comments on Grazing: (e.g., very little livestock sign, not much impact)

Other Wildlife Observed and Number:

Other Comments Relevant to Survey: (e.g., found an old ibex kill on transect)
# SCAT SAMPLE DATA FORM

**Collector's Name:**

**Collector's Phone # or Address:**

**Contact:** Dr. Jan E. Janecka, jjanecka@cvm.tamu.edu
Dr. Rod Jackson, rodjackson@mountain.org
R. B. Munkhtsog, istmon@magicnet.mn

**Texas A&M University**
Snow Leopard Conservancy
Irbiis Mongolia/Mongolian Academy of Sciences

<table>
<thead>
<tr>
<th>Date</th>
<th>Sample ID</th>
<th>Transect</th>
<th>Species</th>
<th>Diamtr./Segmnt.</th>
<th>Scat Age</th>
<th>Snow Leopard Sign and Age</th>
<th>Site</th>
<th>Nearest Landmark</th>
<th>GPS Location/Elevation</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Appendix 3.5: Scat Sample Data Sheet**
Data Codes And Definitions For “Scat Data Sheet”:

1) Date = Date the sample was collected. Example: 01 Oct 2010

2) Transect and Sample ID = Provide a unique sample identification. Scats collected and transects conducted are consecutively labeled with the primary field persons initials.
   - Numbering Transects: Enter the name and number for each transect (see attached Transect Datasheet). These should be numbered consecutively with initials of primary field person (e.g., BM1 for Bariusha Munkhtsog transect 1). The numbering system should be run consecutively across sites, seasons and years, in order that each transect has a completely unique number. This serves to minimize potential errors arising from duplicate names and transect numbers.
   - Numbering Scats: Please number all collected scats consecutively along each transect (i.e., BM1-1, BM1-2,...). Start at 1 again when you move to a new transect (i.e., BM2-1, BM2-2,...). Be sure you record both the transect name/number and scat number on each collection tube.

3) Species = Record the name of the species that you believe deposited the scat based on the size, shape, smell, and associated sign. Example: Snow leopard

4) Diamtr./Segmnt. = Diameter/Segmentation = Record the maximum diameter of the scat and the type of segmentation pattern. Example: 2.1 cm/Segmented

5) Scat Age = Record estimate of the relative age of the scat (See the categories below). Example: 1

6) Snow Leopard Sign and Age = Record all sign (See the categories below) that the scat is close to and the age of each sign (See the categories below). Also record the number and age (in parentheses) of each sign. Example: 3 SC (2), 1 SC (0), 1 UR (2). Note if the collected scat was deposited on a scrape or if not, how far away from the nearest scrape which appears to be about the same age as the scat.

7) Site = Description of the geographic feature for the location where the scat is found. Example: Ridge

8) Nearest Landmark = A close well-known geographic location that is relatively close to where the scat was collected. This important so that the
GPS coordinates of the general area where this scat was collected is known. 
Example: Dorji’s gher

9) GPS Location/Elevation = Record GPS location and save in GPS unit (in decimal degrees). Example: 43 04 890, 101 59 758/2300m

10) Comments = Record any additional information that may be useful. This includes parasites in scats, nearby snare-traps, etc. Also include which other scats were in the vicinity. Example: Along river, had horse hair, there was a dead horse near by, next to LD31

**DEFINITIONS**

**Type of Sign:**

<table>
<thead>
<tr>
<th>Type of Sign</th>
<th>Code</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scrape (felid only)</td>
<td>SC</td>
<td>Scrape made by felid</td>
</tr>
<tr>
<td>Scratch (canid only)</td>
<td>SR</td>
<td>Scratch made by canid</td>
</tr>
<tr>
<td>Feces (scat)</td>
<td>FE</td>
<td>Scat or dropping</td>
</tr>
<tr>
<td>Urine</td>
<td>UR</td>
<td>Urination mark</td>
</tr>
<tr>
<td>Scent Spray</td>
<td>SS</td>
<td>Scent mark</td>
</tr>
<tr>
<td>Claw rake</td>
<td>CL</td>
<td>Claw mark on tree or rock left by felid</td>
</tr>
<tr>
<td>Pugmark</td>
<td>PU</td>
<td>Track impression</td>
</tr>
</tbody>
</table>

**Scat Segmentation:**

B = Block-like appearance, with one or more block-like segments with blunt ends and of uniform diameter (typical felid)

T = Tapered appearance, where scat has distinctly tapered tail or tails, often with irregular diameter (typical canid)

**Note that a scat may exhibit both conditions, though one tends to be more dominant. Variations occur due to differences in diet or the condition of food.**

**Age Categories for Sign:**

**Scrape**

<table>
<thead>
<tr>
<th>Age Category</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Old</td>
<td>0</td>
<td>Extensive weathering and disintegration, scrape features poorly defined, often with vegetation growth in the depression and on the pile (age at least 3 to 6 months).</td>
</tr>
<tr>
<td>Old</td>
<td>1</td>
<td>Moderate weathering and disintegration, with the scrape showing a rounded form, occasionally with vegetation in the depression or on the pile (age several months or more).</td>
</tr>
<tr>
<td>Fresh</td>
<td>2</td>
<td>Slight weathering. Scrape has a well-defined form with “sharp” edges, is easily recognizable, and has no new vegetation growing in the scrape depression or pile (age 1 to 4 weeks).</td>
</tr>
<tr>
<td>-------------</td>
<td>---</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Very Fresh</strong></td>
<td>3</td>
<td>Little or no weathering has occurred, so that the scrape has a very sharp and “clean” form, is very easily recognizable, and has no vegetation in its depression or pile. Sand or gravelly material may cover some vegetation, causing it to “bend-down”. Other ephemeral sign such as tracks or urine may be observed, while scats deposited at the same time are obviously still fresh or very fresh (age less than 1 week).</td>
</tr>
</tbody>
</table>

| **Pugmark** |
|-------------|---|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| **Old**     | 0 | Pugmark is very poorly defined, with an obviously “weathered” appearance (more than 2 weeks old). |
| **Fresh**   | 1 | Pugmark has sharply defined edges and shape (several days, but less than one week old). |
| **Very Fresh** | 2 | Pugmark is very fresh, showing fine surface details and having a very sharp edge (made less than 24 hours previously). |

| **Feces** |
|------------|---|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| **Old**    | 0 | Scat with a hard, dull surface and dry interior —some can be mottled and cracked (several weeks to several months of age). |
| **Fresh**  | 1 | Scat is odoriferous and “fresh-looking”, with a glossy sheen inside (more than 2 days but less than 10 days of age). |
| **Very Fresh** | 2 | Scat is still wet outside and moist inside (no older than 2 days). |

| **Scent-Sprayed Rocks** |
|-------------------------|---|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| **None**                | 0 | No detectable odor (more than 3 months old). |
| **Slight**              | 1 | Odor is just detectable. |
| **Moderate**            | 2 | Odor is readily detectable. |
| **Strong**              | 3 | Odor is unmistakable. |
| **Very Strong**         | 4 | Odor is very strong (can be detected from 25 cm or more away; less than several weeks old). |
Wildlife Research Techniques in Rugged Mountainous Asian Landscapes

Edited by
L. Scott Mills
Tshering Tempa
Ellen Cheng

Ugyen Wangchuck Institute for Conservation and Environment