Frequently-Asked Questions About Wild Sheep Genetics and Genomics

CLINTON W. EPPS, *Department of Fisheries and Wildlife, Oregon State University, Nash Hall Room 104, Corvallis, OR 97330, USA, Clinton.Epps@oregonstate.edu*

LIZABETH BOWEN*, *Western Ecological Research Center, 1 Shields Avenue, Davis, CA 95616, USA*

MICHAEL R. BUCHALSKI, *California Department of Fish and Wildlife, Wildlife Investigations Lab, 1701 Nimbus Road, Suite D, Rancho Cordova, CA 95670, USA*

FRANCIS CASSIRER, *Idaho Department of Fish and Game, 3316 16th St., Lewiston, ID 83501, USA*

DAVID COLTMAN, *Department of Biological Sciences, CW 405, Biological Sciences Building, University of Alberta, Edmonton, AB T6G 2E9, Canada*

WARREN C. CONWAY, *Department of Natural Resources Management, Texas Tech University, Lubbock, TX 79409, USA*

MIKE COX, *Nevada Department of Wildlife, 1100 Valley Road, Reno, Nevada 89512, USA*

RICH HARRIS, *Washington Department of Fish and Wildlife, Box 43141, Olympia, WA 98504-3200, USA*

JOSHUA P. JAHNER, *Department of Biology, University of Nevada, Reno, NV 89557, USA*

MARJORIE MATOCQ, *Department of Natural Resources and Environmental Science; Program in Ecology, Evolution and Conservation Biology, University of Nevada, Reno, NV 89557, USA*

THOMAS L. PARCHMAN, *Department of Biology; Program in Ecology, Evolution, and Conservation Biology, University of Nevada, Reno, NV 89557, USA*

CALEB D. PHILLIPS, *Department of Biological Sciences and Museum, Texas Tech University, Lubbock, TX 79409, USA*

HELEN SCHWANTJE, *Wildlife Veterinarian, Wildlife and Habitat Branch, Ministry of Forests, Lands, Natural Resource Operations and Rural Development, 2080 Labieux Road, Nanaimo, BC V9T 6J9, Canada*

ZIJIAN SIM, *Department of Biological Sciences, CW 405, Biological Sciences Building, University of Alberta, Edmonton, AB T6G 2E9, Canada*

LISETTE WAITS, *Department of Fish and Wildlife Science, University of Idaho, Moscow ID 83844-1136, USA*

JOHN D. WEHAUSEN, *Sierra Nevada Bighorn Sheep Foundation, P.O Box 1183, Bishop, CA 93515, USA*

WILD SHEEP GENOMICS WORKING GROUP

*Authors alphabetical after Epps

ABSTRACT: Research focusing on population genetic variation has informed our understanding of North American wild sheep conservation and management for decades. New techniques, however, are emerging at an increasing rate, including genomics approaches that investigate a larger portion of the genome than was previously possible. Research has addressed questions at many scales, from assessing relationships among individuals within single populations, to broad-scale patterns of genetic variation across multiple states or provinces, to deep evolutionary relationships among species. Meanwhile, wildlife managers seek answers for management-relevant questions old and new that could be informed by appropriate analyses using traditional population genetics tools and newer genomic methodologies. Improving communication among researchers and wildlife managers responsible for wild sheep populations or other taxa across multiple jurisdictions, as well as clarifying what questions are best addressed by different genetic or genomic approaches, could facilitate collaboration and improve research. To that end, the informal Wild Sheep Genomic Working Group was established in 2018, including members from both research and management backgrounds. A "frequently-asked questions" (FAQ) document was established to facilitate conversation; questions were collated from group members and presented as part of a special session on genetics and genomics at the Northern Wild Sheep and Goat Council Symposium held in Whitefish, Montana, during May 2018. Here, we present an edited version of those questions with responses provided by the group, including a glossary of technical terms. We address methodological choices and ways to improve collaboration, provide examples of how population genetics research has informed management, and discuss genetic diversity, subspecies management, and genetics as a tool to understand disease. We note that the projects discussed and researchers contributing herein are not a comprehensive list of current genetic research on wild sheep, but we view this as a first step towards improving collaborative research to inform conservation and management of wild sheep and goats in North America.

Biennial Symposium of the Northern Wild Sheep and Goat Council 21:48-75; 2018

KEY WORDS Disease; DNA; genetics; genomics; local adaptation; microsatellites; SNPs; sequence; taxonomy.

For the last three decades, conservation and management of wild sheep in North America has drawn upon a rich body of research using population genetic approaches. Wild sheep have been attractive subjects for genetic research because of their evolutionary history, extensive history of population translocation and management, high value to the public as a game animal and iconic symbol of desert or alpine habitats, and because their small founder or current population sizes raise concerns about the maintenance of genetic diversity. Genetic studies have better informed wild sheep conservation and management on numerous topics, including **phylogeography** (Ramey II 1995, Luikart and Allendorf 1996, Wilder et al. 2014, Malaney et al. 2015, Buchalski et al. 2016, Sim et al. 2016), population structure and connectivity (Gutierrez-Espeleta et al. 2000, Buchalski et al. 2015, Epps et al. 2018), trans-boundary movement (Flesch et al. 2010, Buchalski et

al. 2015), horn growth (Fitzsimmons et al. 1995, Miller et al. 2018), population growth (Johnson et al. 2011), reintroduction effects (Ramey II et al. 2000, Whittaker et al. 2004, Hedrick and Wehausen 2014, Malaney et al. 2015, Gille et al. *In press,* Jahner et al. *In press*), inbreeding depression and genetic rescue (Hedrick et al. 2001, Hogg et al. 2006, Miller et al. 2012, Olson et al. 2012), and most recently local adaptation (Kardos et al. 2015, Roffler et al. 2016) to name but a few. Thus, genetic research is likely to remain an integral component of wild sheep conservation and management for the foreseeable future.

Over the last decade, tools for genetic analysis have expanded rapidly, including approaches sometimes referred to as *genomics* because they address variation across a larger proportion of a research subject's **genome**. The distinction between genetics and genomics research, however, is not always clear, and may be best viewed as a continuum of research from small portions of the genome (e.g. 10-20 **microsatellite loci** or partial sequence of **mitochondrial DNA** [mtDNA] cytochrome b or control region) all the way to comparing nearly complete genome sequences of many individuals. Technological advancements for high-throughput DNA sequencing over the last decade have been transformative (Mardis 2013), resulting in extremely rapid change in data type and volume. These technologies can sequence millions of genomic fragments ("reads") for multiple individuals all at once, versus targeted sequencing of individual **loci**, or can sequence many targeted loci at once as in high-throughput microsatellite genotyping (De Barba et al. 2017). This means that in addition to traditional single-locus approaches, geneticists can now sequence tens or hundreds of thousands of loci simultaneously and at a very small cost. Further, these loci (e.g., **single nucleotide polymorphisms**, or **SNP**s) can occur throughout the genome including gene coding regions. Thus, it is now possible for geneticists to simultaneously assess genomic variation resulting from both demographic processes and natural selection.

With the rapid advancement of techniques and increasing sophistication of research questions, however, has come increased need for strengthening communication between wildlife managers and researchers, as well as improving collaborations across state, provincial, and national borders. The Western Association of Fish and Wildlife Agencies (WAFWA) Wild Sheep Working Group began to address that need in 2018 by convening an informal Wild Sheep Genomic Working Group, including members from both research and management backgrounds. The group has focused to date on North American wild sheep, including bighorn sheep (*Ovis canadensis*) and the thinhorn sheep (*Ovis dalli*), but the approach could be extended to mountain goats (*Oreamnos americanus*) given similar evolutionary, ecological, conservation, and life histories.

After a suggestion by Zijian Sim, the group decided to collect a list of genetics questions and allow group members to respond individually, providing a diversity of viewpoints, with the intent that those questions and responses could be collated and published. Questions fell into two general categories. One category included those questions that originated with geneticists and were directed to wildlife managers, asking what data gaps currently limit informed management and how genetics research can be of benefit. The other category included questions from

wildlife managers directed to geneticists regarding the various approaches to genetic research, and the appropriateness of the various methods for assisting with wild sheep conservation and management. Questions were provided by Mike Cox of Nevada Dept. of Wildlife (NDOW), based in part on a January 2018 survey of Wild Sheep Working Group representatives regarding critical data gaps for wild sheep genetics and genomics, with significant input from Rich Harris of Washington Dept. of Fish and Wildlife (WDFW) and authors of this manuscript. Questions were submitted in April-May 2018, edited, and then were presented for discussion at a genetics and genomics symposium at the Northern Wild Sheep and Goat Council meeting held in Whitefish Montana during May 2018.

Here, we present those questions and the responses gathered during and after the symposium. The questions and responses are organized into topics, including 1) methodological choices and ways to improve collaboration, 2) examples of how genetics research has informed management, 3) causes and consequences of genetic diversity, 4) management of taxonomic units below the species level, and 5) genetics as a tool to understand disease. Questions are numbered sequentially across categories and italicized; answers are numbered and attributed to the source. We note, however, that as participation was self-selected, not all projects and researchers currently working on wild sheep genetics are represented. We have included a glossary of technical terms (Appendix 1); terms included in the glossary are indicated in bold-face type at first usage.

QUESTIONS AND ANSWERS

Questions and needs relating to methods and collaborations

1) *What are the various types of genotype/loci datasets each group is collecting- can we assess the possibility for comparisons in the future?*

Answer I (Epps): The Oregon State University (Epps) research group has collected **neutral microsatellite** (13-14 loci) and **adaptive-linked microsatellite** data (3-8 loci) for bighorn sheep across most of the range of desert bighorn sheep (*O. c. nelsoni*) in California (excluding Peninsular populations), desert bighorn in Utah and the Grand Canyon of Arizona, and are generating neutral and adaptive-linked microsatellite data for 'California' (*O. c. canadensis*, formerly *O. c. californiana*) bighorn populations in Oregon. These data can be compared with other microsatellite datasets using overlapping sets of loci by running reference samples in multiple labs to allow alignment (e.g., Gille et al. *In press*).We recommend that researchers extract a large quantity of DNA from several individuals to create such reference samples.

Previously, we have merged our datasets with data produced by John Wehausen, Kathy Longshore, and Jef Jaeger (southern Nevada) as part of Tyler Creech's dissertation work (Epps et al. 2016, Creech et al. 2017), Mike Buchalski, Walter Boyce, and other's data (Buchalski et al. 2015) for Peninsular bighorn work (in preparation), and with Daphne Gille, Mike Buchalski, Walter Boyce, Holly Ernest, and others for Arizona-wide analyses (Gille et al. *In press*). Although

the number of overlapping **loci** was small in the last case, we were able to align datasets to identify important patterns of **genetic structure** across Arizona. We hope to share data between Nevada/Oregon populations of 'California' bighorn with Marjorie Matocq at University of Nevada and potentially people working in Idaho. Microsatellite data seem to do a good job of addressing questions of bighorn genetic structure, **gene flow**, connectivity, and population history (e.g., Epps et al. 2018), although we are still determining what additional resolution **Single Nucleotide Polymorphisms (SNPs)** might bring.

In addition, the Epps Lab recently used a method of **restriction site-associated DNA sequencing** (RADseq, Wang et al. 2012, see http://people.oregonstate.edu/~meyere/docs/2bRAD_25Aug2016.pdf) to generate SNP genotypes for ~10,000 loci from Sierra Nevada bighorn sheep, Rocky Mountain bighorn sheep from Alberta and British Columbia, and desert bighorn sheep from throughout the majority of that subspecies' natural range for a **landscape genomics** study investigating local adaptation to climate, in collaboration with CDFW (Buchalski), University of Nevada (Matocq), and other partners. These SNPs provide a signal of genetic population structure consistent with previous analyses of microsatellite data (Buchalski et al. 2016) and at least one SNP locus shows strong evidence of selection (adaptive-linked), suggesting local adaptation among remnant herds of the Great Basin (Buchalski et al. *In prep*).

Answer II (Jahner): The University of Nevada group (Matocq, Parchman, Jahner) has collaborated with NDOW to generate genetic data from populations of California, desert, and Rocky Mountain bighorn sheep across Nevada. Our initial effort entailed generating microsatellite data (16 loci) from 347 individuals across 55 herds representing all three "subspecies" in Nevada (Malaney et al. (2015). These data showed marked genetic subdivision between 'California', desert, and Rocky Mountain bighorn in Nevada as well as clear signatures of hybridization where these lineages come into spatial proximity. More recently, we have generated genotyping-by-sequencing data for several hundred individuals (analogous to a type of restriction site-associated DNA sequencing (Peterson et al. 2012)), using the protocol found in Parchman et al. (2012), and have a new paper describing the genetic consequences of Nevadan desert bighorn sheep translocations over the past 50 years (Jahner et al. *In press*). This is a reduced representation method, where subsets of genomes are sampled and sequenced based on the distribution of restriction enzyme cut sites. These approaches allow direct sequencing and SNP detection in tens of or hundreds of thousands of genomic regions and can be executed with large numbers of individuals in a time and cost-effective manner to rapidly generate population genomic level data (Andrews et al. 2016). Our dataset of more than 17,000 SNPs was able to detect very fine-scale genetic differentiation among geographically proximate, remnant herds in southern

Nevada, a signal consistent with microsatellite data, as well as recovering signatures of **admixture** in translocated desert herds that reflect the history of translocation. Our ongoing sequencing efforts are focused on establishing a genome-wide view of **hybridization** among subspecies to complement our microsatellite-based studies, and to investigate whether some remnant, central Nevadan herds are relicts of a Great Basin lineage of desert bighorn sheep (we will be collaborating closely with Mike Buchalski and CDFW for this second question). The sequencing data that we generate should be fairly comparable with studies using similar library preparation methods, but will have little comparability to studies using other marker types. We have interest in trying the **Ovine SNP chip** (array) that other labs have successfully used (e.g., Miller et al. 2018) as a mechanism for generating genotypic data that could be readily transferable among any groups using this resource. Because we now have parallel microsatellite and SNP datasets for multiple individuals and herds, we can compare these measures of genetic variation to help bridge interpretation between these different data types (see next question).

Answer III (Sim & Coltman): The Coltman research group (University of Alberta) has two datasets that may be of interest:

i) A ~10,000 SNP dataset typed on 55 individuals using the OvineHD SNP array. This dataset formed the basis of Sim et al. (2016), which investigated the **phylogeographical** history of thinhorn sheep (*Ovis dalli*). That paper helped resolve the relationship between Dall's, Stone's and Fannin sheep. Data from this paper are publicly available on the Dryad database: https://datadryad.org/resource/doi:10.5061/dryad.c87rm. Data from SNP arrays are comparable with other datasets generated from the same array.

ii) A dataset of 153 neutral SNPs typed on ~2500 individuals covering the entire thinhorn sheep range. This dataset was used to study the range-wide population genetic structure of thinhorn sheep. The paper describing the results are in review so the data likely will not be available until the paper is published. In the meantime, we are happy to share the primer sequences should other groups be interested in genotyping their own samples and/or combining data. Our SNP discovery panel was chosen to ensure even representation from all jurisdictions so ascertainment bias should be minimal.

Answer IV (Waits): The Waits lab (University of Idaho) has multiple datasets that may be of interest, including:

i) A dataset of neutral and adaptive-linked microsatellite loci for ~450 sheep in Idaho including some historic specimens, plus a mitochondrial DNA (mtDNA; control region**)** dataset for ~200 sheep.

ii) A dataset of neutral and adaptive-linked microsatellite loci

for 300 sheep in Washington and the Lostine and Lookout populations of Oregon, as well as putative source populations in the Fraser/Williams Lake area of British Columbia (n=112).

iii) A SNP dataset generated by RADseq for the Lostine population (n=100, 142-523 loci) in Oregon (Andrews et al. 2018).

Answer V (Wehausen): In 1998, with Rob Ramey, we developed the first population genetic data for Sierra Nevada bighorn sheep. We used 13 microsatellite loci to assess genetic population structure of the surviving populations, initially using DNA extracted from fecal samples and tissue from skulls collected in the field, and sampled all populations of desert bighorn sheep immediately east of the southern Sierra Nevada to Death Valley to provide geographical context. We quickly learned that Sierra bighorn had a strong signature of a recent bottleneck and notably lower genetic diversity than other sampled populations; this finding was added to an early draft of the recovery plan for Sierra bighorn (USFWS 2007). The desert bighorn sheep data were later augmented with additional microsatellite loci and combined with additional data from the Death Valley region run in the Epps lab (Epps et al. 2016). In combination with other efforts, this completed the genetic sampling of essentially all bighorn sheep populations in California, allowing comparisons of genetic diversity metrics at neutral microsatellite loci.

The genetic research on Sierra bighorn began by using the San Diego Zoo genetics lab, courtesy of Ollie Ryder, then transitioned to a private lab Ramey and Wehausen set up in Nederland CO. That was followed by the creation of another independent lab in Bishop, CA through the Sierra Nevada Bighorn Sheep Foundation, where all subsequent work has been performed by Wehausen. One project run in this lab investigated the question of potential fitness influences of the low genetic diversity of Sierra bighorn and found a statistically significant heterozygosity-fitness relationship in Sierra bighorn involving the probability of females having lambs in summer (Johnson et al. 2011). More than 50 microsatellite loci are now used, with more to be added in order to be able to determine parentage of all lambs in one small population genotyped from fecal samples for 20 years. Another graduate project is further examining the robustness of the findings of Johnson et al. (2011) using considerably more loci and data.

A large number of samples from Sierra bighorn have now been genotyped and resulting genetic data have been used as a basis for management actions to influence genetic diversity in reintroduced populations and existing populations through augmentation (genetic rescue), and to track genetic diversity of all populations to evaluate success of genetic management efforts. These data also have been used to investigate gene flow between some populations.

Another project from this lab has characterized genetic population structure and gene flow in southern Nevada immediately south and north of Las Vegas, in collaboration with Jef

Jaeger at UNLV and Kathy Longshore of USGS (see Epps answer above). Finally, a long-term project is using mtDNA to investigate the evolutionary history of North American wild sheep, in collaboration with R. Ramey and C. Epps. This began with the sequencing of control region for a large number of samples to compliment what Epps et al. (2010) and Boyce et al. (1999) had produced. After finding that the control region could not adequately address this question, this effort shifted to conserved protein coding genes. Additional sequence data and samples have been added periodically. This currently includes 4 genes (ND5, ND6, ND2, and COIII) for a total of 3,336 bp.

2) *What past and present methods have been used to describe wild sheep genetics: microsatellites, SNPs, others? What is the most appropriate sample material for each? What level of detail and types of questions that are best answered by each method? Are results from each method comparable?*

Answer I (Sim): Combining SNP datasets can be fairly straightforward in some cases, unlike for microsatellites where a bit of calibrating/standardizing work needs to be done (although as the Epps collaborations have shown, it is not impossible).

Answer II (Parchman): This is especially the case if SNPs are called (i.e., determining what base pair is present at each variable site) using a high throughput SNP chip, such as the OvineHD SNP array. Genotyping-by-sequencing (RADseq) data can be easily combined if they are generated with the same library preparation method and if the raw sequencing data are available.

Answer III (Epps): I have started writing a review paper to explore this topic in detail. Different types of markers are typically best suited for different questions and may not be directly comparable. Wild sheep genetics have been addressed by many methods, including **Restriction Fragment Length Polymorphisms** (**RFLPs)** at mitochondrial or **Major Histocompatibility Complex (MHC)** genes (e.g., Ramey II 1995, Boyce et al. 1997), neutral microsatellite markers (e.g., Coltman et al. 2002, Whittaker et al. 2004, Epps et al. 2005, Miller et al. 2012), adaptive-linked microsatellite markers (e.g., Luikart et al. 2008a, Plowright et al. 2017), mitochondrial **DNA sequences** (e.g., Boyce et al. 1999, Buchalski et al. 2015), SNPs (e.g., Miller et al. 2014, Sim et al. 2016), and other approaches. In recent publications, microsatellite loci, mtDNA and nuclear sequence data, and SNPs seem to be most commonly employed. Microsatellites are still commonly used to estimate genetic structure and diversity (e.g., Epps et al. 2018), and identify individuals from non-invasive (e.g., feces or hair) samples as for population estimation or other purposes. Microsatellite loci are well-characterized by many labs, usually can be aligned with existing datasets by re-running individual samples previously genotyped elsewhere and are relatively easy to amplify from non-invasive samples. Mitochondrial and other DNA sequence data may be best suited for phylogeographic or phylogenetic analyses or direct assessment of variation at genes of interest. Short

sequences (e.g., mitochondrial DNA sequences of up to ~600 bp or so) can be easily amplified from non-invasive samples, but "next generation" sequencing approaches usually require larger quantities of DNA with little contamination, as can be obtained from blood or tissue.

Modern high-throughput sequencing or genotyping approaches are capable of generating many more loci (usually SNPs) across the genome, and therefore allow finer-scaled or more accurate estimates of genetic structure, inbreeding, diversity, and **heterozygosity** that will be more representative of the rest of the genome than estimates from a small number of microsatellite loci (Miller et al. 2014). If sufficient marker density is achieved, SNPs generated from such approaches can be used to identify variable markers showing evidence of selection (e.g., landscape genomics) or association with different traits (e.g., a **genome-wide association study [GWAS]**). While microsatellites may also show evidence of selection if closely associated with genes of interest ("adaptive-linked") (e.g., Luikart et al. 2008a, Plowright et al. 2017), the greatly-increased number of loci in most SNP studies increases the chance of detecting associations, although strong genetic drift and isolation common to wild sheep and goat populations still make such studies challenging. Some methods of generating SNP datasets (e.g., RADseq) may not work well with non-invasive samples, degraded DNA, or small amounts of DNA, and thus are best suited to large amounts of clean DNA obtained from well-preserved blood or tissue. SNP datasets can be difficult to align or compare among research groups if library preparation methods differ (but see Parchman's comment above). However, SNP assays can be designed to work with other types of material and such assays can be quite comparable across research groups (Carroll et al. 2018). Investment in the development of an assay (e.g., a SNP chip) that can consistently recover a large, shared set of target loci from a range of tissues or non-invasive samples will be important for moving forward in a coordinated way.

3) *What are standards and options for long-term storage of wild sheep genetic material?*

   Answer 1 (Sim): Here are some options with positive (+) and negative (-) attributes indicated:

   i) Horn core drillings (can be stored in paper envelopes)
      (+) Highly stable over long periods of time (decades)
      (+) Can be collected from animals whose tissue have putrefied
      (+) Can be stored at room temperature
      (+) Very space efficient
      (-) DNA yield is low – sufficient for genotyping/sequencing methods that include an amplification step (e.g., microsatellites) but may not be suitable for next generation library preparation methods like RADseq. Matocq notes, however, that by designing an assay that includes a DNA enrichment step, it may be

feasible to work with such material.

(-) Processing of horn cores for DNA extraction is messy – higher risk of contamination

ii) Frozen tissue

(+) High DNA and RNA yield provided tissue is fresh and preserved soon after collection

(+) Relatively stable if stored at -80° C, although storage in liquid nitrogen (-196° C) reduces degradation of RNA and DNA over time

(+) Provides insight into gene expression (functional genomics), not just gene sequence data

(-) Susceptible to equipment failure, power outages, flooding, etc., requires access to freezers

iii) Tissue in ethanol

(+) High DNA yield provided tissue is fresh and preserved soon after collection

(+) Can be stored at room temperature

(+) Stable over long periods of time (years)

(-) Ethanol is highly flammable – Safety rules of your institution may require special permits or special conditions for storage

(-) Stability is highly dependent on the seal quality of storage container – containers with gasketed caps are required for long-term storage. Ethanol in containers with snap on caps (e.g., 1.5 mL centrifuge tubes) will eventually evaporate.

(-) Tissue with high moisture content can dilute the ethanol, making it less effective.

Answer II (Waits): Blood is a valuable source of DNA not mentioned above, particularly because many samples have been collected for disease work. It should be frozen or mixed with lysis buffer for long term storage. When mixed 1:4 with lysis buffer, it can be stored at room temperature for months.

Answer III (Epps): We and many other groups have made extensive use of DNA from blood and feces (e.g., Wehausen et al. 2004, Luikart et al. 2008b, Luikart et al. 2011, Driscoll et al. 2015, Epps et al. 2018), and some use of RNA from blood or tissue:

i) DNA from feces can be archived for decades in Tris-EDTA buffer at -80C; I have amplified microsatellite markers from 15-year-old samples collected and stored in this manner. Feces appear to be a stable source of DNA for amplification (PCR)-based analyses for years if pellets are kept dry and dark and in a controlled environment, but Wehausen got no DNA from 15-year-old samples stored in a very dry environment. There is apparent degradation over time, so it is best to extract and store the DNA as soon as possible (within a few years) or freeze the samples. Currently, however, we cannot easily use this type of sample for some of the SNP-

type analyses (e.g., RADseq), although we could presumably do targeted assessments of specific SNPs thought to be associated with important genes, or as part of a large assay panel (see above). As others have shown, the relative ease of obtaining fecal samples greatly augments our sampling options and can easily be adapted to citizen science projects, as long as they are carefully overseen.

ii) Marjorie Matocq notes that another benefit of collecting fecal samples is that we can use high-throughput sequencing methods to gain insight into diet (Pompanon et al. 2012) and the gut microbial communities maintained by these animals (Kohl 2017). We are learning more and more about the role of the microbiome in nutrient acquisition and immunity (Alberdi et al. 2016), and Marjorie suggests that our group should advocate including inventory of diet and microbiome in our screens of population parameters. Understanding how diet and gut microbiomes vary spatially and temporally should become an important element of how we inventory these populations, especially in relation to translocations and changing food availability and plant chemistry that is anticipated with climate change.

iii) DNA from blood has been a mainstay of work using samples collected during live capture of bighorn sheep. Whole blood in EDTA, frozen, can yield usable DNA for sequencing or microsatellite analyses, but we have had trouble getting sufficient yield for genomics approaches from this type of sample. We have had the best results from spinning down blood tubes shortly after capture, pipetting off the "buffy coat" (white blood cells), and freezing that.

iv) RNA can be isolated from blood or tissue but requires special buffers and handling. RNA is used for studies of gene transcription and expression.

v) Hair is often collected at captures, can serve as a source of good-quality DNA, and can also be used in isotope or hormone analysis. Hair has been used for "next generation" analyses (Russello et al. 2015), but DNA quantity is low. Thus, we do not usually recommend relying on hair for DNA work on captured animals, given that blood or tissue samples are easy to obtain in most cases.

vi) Captures and necropsies also present opportunities for detecting DNA of pathogens. Collection and storage

methods vary depending on disease and material available but may include swabs of nasal or pharyngeal cavities on live animals (respiratory pathogens) or sections of lung, bone marrow, or other tissue.

vii) Consider splitting important samples (blood, extracted DNA, etc.) into multiple subsamples that are stored in different locations, as freezer failures and loss of material do occur periodically. As Texas Tech University has opened a new long-term storage facility for bighorn sheep DNA, this would be an excellent repository for backup samples and could facilitate collaborations or follow-up studies.

viii) Finally, make sure associated data are or can be linked to samples: age, sex, locality, source stock or population history, reproductive status, data on disease at time of capture, etc.

Answer IV (Conway, Phillips): The Natural Sciences Research Laboratory (NSRL), housed within the Museum of Texas Tech University, has installed Liquid Nitrogen (LN) freezer storage specifically for bighorn sheep samples (tissue; blood; fecal; swabs; etc.) via funds from the Wild Sheep Foundation and The Texas Bighorn Society. This storage will provide the in-perpetuity archival of a variety of bighorn sheep samples supporting the range-wide Disease Management Venture, and the genomics-based research on bighorn sheep in Texas and range-wide. The NSRL is a research repository with the mission to archive biological samples and their associated data for scientists throughout the world; requests for samples and associated data can be made by national and international researchers. The Genetic Resources Collection (GRC) of the NSRL to date curates approximately 375,000 tissue samples from >100,000 individual specimens distributed worldwide. Currently, the bighorn sheep LN freezer contains samples from ~200 desert bighorn sheep in Texas from the last three years of disease sampling captures. Efforts to obtain samples from Idaho have recently been completed – although curation and database development for those samples has just been initiated. Curation of additional samples from previous bighorn sheep work in Texas is also underway, through collaborations with Texas Parks and Wildlife Department biologists and staff and private landowners.

4) *What are the pros and cons of using the domestic sheep genome versus efforts to improve the wild sheep genome: should all jurisdictions help contribute to this effort, and which genomics lab is best suited to take it on if worthwhile?*

Answer I (Jahner): In our genotyping-by-sequencing (RADseq) dataset ('California', desert, and Rocky Mountain bighorn sheep populations in Nevada), we were able to align a higher percentage of bighorn sheep reads (relatively short DNA sequences generated by "next-generation" sequencing platforms) to the domestic sheep genome than to the Rocky Mountain bighorn sheep genome.

Recent improvements to sequencing platforms that generate longer reads (e.g., PacBio) or enable improved scaffolding (basically, arranging sequences and gaps of known length, e.g. Chicago libraries, Hi-C) have made highly contiguous genome assemblies readily attainable even when no reference genome exists (e.g., Bredeson et al. 2016, Putnam et al. 2016). The latter methods mean that near chromosomal level assemblies can be generated *de novo* for most non-model organisms at a fraction of the cost and time that would be have been required just a few years ago. Improved genome assemblies for wild sheep, including for each recognized taxonomic lineage, will be important resources for improving inference from reduced-representation resources (e.g. higher density SNP chips) and eventually enabling whole-genome resequencing for understanding the genetics of adaptation.

Answer II (Sim): A draft genome of a Rocky Mountain bighorn sheep is available via Josh Miller, formerly of the Coltman Lab (Miller et al. 2015). As it stands now, almost all the genetic resources including microsatellites, SNPs, and genome have originated from or are heavily dependent on domestic sheep, so I am personally in favor developing some resources directly from wild sheep. Jahner et al.'s (*In press*) observation that their dataset aligned better to the domestic sheep genome than the Rocky Mountain bighorn sheep genome may be related to the fact that the Rocky Mountain bighorn sheep genome was assembled via alignment to the domestic sheep genome, so everything ends up looking like the domestic sheep. This perhaps further speaks to the need for more genetic resources generated directly using wild sheep.

How has genetics informed wild sheep management?

5) *How successful have agencies been in implementing management decisions based on genetics research?*

Answer I (Cox): Obviously Sim's work on redefining Stone/Dall Sheep distribution (Sim et al. 2016) has been accepted by British Columbia, Yukon, and NWT. I think other jurisdictions would make the right decisions based on peer-reviewed sound science. In 2000, I was successful in using the Wehausen and Ramey II (2000) study that showed the entire Great Basin was a single bighorn subspecies of desert bighorn to help dispell the old myth, assumed by my agency and past biologists, that there were 3 subspecies of bighorn in Nevada (desert, 'California', Rocky Mountain). We greatly encouraged broader distribution of desert bighorn sheep translocations after that time; unfortunately, many introductions of 'California' and Rocky Mountain bighorn had already occurred.

Answer II (Epps): The Sierra Nevada Bighorn Sheep program (California Dept. of Fish and Wildlife) has used John Wehausen's estimates of genetic diversity at population and individual levels to guide translocation and augmentation strategies. Connectivity analyses by Creech et al. (2014) that were based on landscape genetic analyses of bighorn sheep in the Mojave (Epps et al. 2007) have been used by the National Park Service and Bureau of Land Management to guide decision making around renewable energy sites.

Oregon Department of Fish and Wildlife used genetic studies to explore the consequences of multiple founder effects and conduct experimental translocations (Whittaker et al. 2004, Olson et al. 2012).

Answer III (Waits): Frances Cassirer and Hollie Miyasaki from Idaho Department of Game and Fish (IDGF) have been using results from genetic analyses by the Waits research group to inform a variety of management decisions. For instance, the genetic connecivity analysis has been used to inform land management decisions by the Forest Service, and genetic structure has been used by IDGF to evaluate how bighorn sheep were grouped into population management units.

Questions about genetic diversity

6) *For wild sheep managers forced to manage relatively small and isolated herds, are there, or can there be constructed, guidelines that outline what is adequate genetic diversity before herd performance is compromised?*

Answer I (Cox): I think that those of us who are managers would love to have some simple guidelines and metrics we can collect, measure, and take action on to maintain a certain level of genetic diversity.

Answer II (Epps): I think there is much to be done. The SNP-type datasets should be helpful as they cover more of the genome, but diversity estimates vary depending on marker choice. SNP diversity correlates but is not directly comparable with estimates from microsatellites (e.g., Miller et al. 2014), and may be sensitive to choice of SNPs. We need measures of herd and individual performance to link to the genetic data, and need to recognize that many other non-genetic factors will affect those variables as well, which will make it challenging but not impossible to establish relationships between genetic diversity and performance. Sim (below) argues that estimates of genetic diversity from 8-20 microsatellites may give little resolution with respect to herd performance: I agree that many more loci would do a lot to increase power, but also believe we have not yet had many chances to confront estimates of genetic diversity of any type with herd performance metrics over large datasets. Luckily, I think we are now poised to do some of that.

Answer III (Sim): Unfortunately, there is no one magic number that genetics can give that says, for instance, "if heterozygosity > 0.6145 then performance will not be compromised". What you want to know is the extent to which genetic diversity is correlated with some measure or proxy of herd performance or fitness (fecundity, horn size, body weight, etc.). What exactly herd performance means will ultimately, I think, be up to the manager and likely influenced by other ecological factors. Without the corresponding data for herd performance, a single measure of diversity is probably not useful unless the values are extremely low [for reference, Epps notes that microsatellite-based estimates of expected heterozygosity <0.5 characterize isolated populations established from small numbers of founders (e.g., Hedrick and Wehausen 2014), which are sometimes associated with poor population performance (e.g., Hogg et al. 2006)]. As well, the

traditional 8-20 microsatellite approach to genotyping likely cannot measure genome wide diversity precisely enough to be useful for comparisons of genetic diversity and herd performance (although there is discussion on this topic, see above). Hopefully, with the price of sequencing coming down, assays using tens/hundreds of thousands of SNPs will provide enough genomic coverage to give precise enough measures of diversity for this kind of comparison. Beyond the cost of sequencing, the post-sequencing processing of large SNP datasets need to be worked out before widespread adoption since most current methods are not very user friendly and require very steep learning curves. The field is moving fast though so there is great hope on this front.

Answer IV (Waits): I agree and would like to emphasize Sim's point above that the key to understanding these relationships is to have large datasets of phenotypic/morphological/fitness trait data from individuals linked to the genetic data. This is a key area where managers and geneticists can work together.

7) *Given that we have little choice now but to manage most wild sheep populations as islands (perhaps linked by artificial migration), and most are much smaller in size than historically the case, what is the current best wisdom regarding levels of diversity, inbreeding, drift that can be tolerated before we do a disservice to our populations, and how do we figure it out? Perhaps these animals are adapted to lower levels of heterozygosity than we would initially be comfortable with…but we have also seen tantalizing hints that small, stagnant populations are sometimes jump-started by infusions of new genetic material, suggesting that heterozygosity or specific* ***alleles*** *matter. Given that practitioners lack the resources to conduct genuine studies on any but a fraction of our populations, how do we learn more and make better decisions?*

Answer I (Cox): As with the Disease Management Venture (an ongoing effort by agencies to share data on bighorn sheep performance and disease), we need to share data on genetic diversity levels and simple herd performance metrics and compare/contrast west-wide. This would allow us to see the spectrum of values and identify "breakpoints" on which we can agree, for example, identifying minimum values or triggers that would require management action, or provide managers with some confidence that herds are doing just fine without genetic "rescues" or intervention.

Answer II (Jahner): It is worth noting that translocations have been successfully used in the past to elevate genetic diversity in bighorn sheep populations (Hogg et al. 2006, Miller et al. 2012, Olson et al. 2012), but we still have much to learn about 1) what constitutes "low" genetic diversity in bighorn sheep, and 2) how levels of genetic diversity actually affect important population demographic parameters.

Answer III (Epps): I believe that we can learn lessons from some of the truly isolated herds, such as the Sespe population in southwestern California, in which we see some physical abnormalities. In more natural arrays of populations connected by occasional

dispersal, as long as there is occasional gene flow among populations, I doubt that genetic diversity will decline to point where we would see obvious inbreeding effects. The biggest management question probably occurs in systems like Oregon, where decades of sequential founder effects have apparently created populations with lower genetic diversity, albeit interlinked in most cases by occasional gene flow. Is it worth bringing in individuals from different source populations, as was done experimentally in Oregon (Olson et al. 2012;2013)? Careful studies of individual and population performance, with and without challenges such as disease, coupled with better characterization of genetic diversity (e.g., using many more markers than is usual in previous microsatellite-based studies), are needed and in some cases are underway.

8) *Should managers be placing more emphasis in maintaining unique remnant herds that may be isolated with low genetic diversity but are performing well? How do you measure, quantify, or place a value on the importance of these remnant herds maintaining their uniqueness?*

Answer I (Cox): We had some of these unique herds detected from the UNR Genomics Lab (Matocq, Jahner, Parchman) recently. We certainly want to maintain these remnant herds and their genetic integrity and even use them as source stock to expand this important historic remnant in more areas in Nevada. So yes, westwide, such herds need to be identified and enhanced-- as long as geneticists feel that their low diversity does not compromise their performance and is outweighed by their unique traits that may be better suited to the local environmental conditions.

Answer II (Jahner): This is a complicated problem. Translocations have the potential to quickly eliminate any signatures of local adaptation within a population through gene flow, which would result in the loss of ecologically and evolutionarily important variation. However, genetically isolated or unique populations are not necessarily locally adapted, and perhaps could benefit from augmentations. Any metric that attempts to prioritize some herds over others for maintaining their uniqueness should ideally consider genetic, morphological, and ecological information, but this further complicates this endeavor.

Answer III (Harris): How do we distinguish between uniqueness owing to adaptive evolution versus genetic drift? Conserving every uniquely-identifiable genetic variant can increase extinction risk due to demographic and genetic processes: take a look at Weeks et al. (2016).

Answer IV (Epps): Harris' concern above identifies a real risk in my opinion-statistical significance does not imply biological significance. We found the distinction between drift and local adaptation to be a challenging problem in our recent landscape genomic project looking for adaptive genetic variation associated with climate variation across the range of desert bighorn sheep (Buchalski et al. *in prep*). Our solution was to exert stringent criteria (e.g., markers identified by multiple methods) to avoid false positives due to isolation

and drift, but this field of study is still in its infancy.

Questions about subspecies management

9) *How do we manage 'California' vs. Rocky Mountain bighorn populations (formerly considered subspecies)? Can we describe genetics of original remnant 'California' bighorn herds in BC, compare them to the translocated 'California' bighorn herds, and provide guidelines to managers on maintaining or mixing the 2 lineages and the pros and cons?*

Answer I (Cox): Helen Schwantje is providing samples to Marjorie Matocq to get genetics of the remnant BC 'California' bighorn herds to compare to the rest of our introduced herds in lower 48. We just need a few geneticists to share values from all the introduced herds and develop guidelines through series of brainstorm sessions with subset of geneticists and managers.

Answer II (Schwantje): Additional work may be required to target sample the "original herds".

Answer III (Epps): I am definitely interested in working on this question as well, and have worked with ODFW and Ph.D. student Rob Spaan and ongoing research to sample many of the 'California' herds in Oregon. We are characterizing neutral genetic markers at this stage for other questions (we are in process of genotyping hundreds of samples at 15-20 microsatellite loci), but some of these samples would be suitable for SNPs etc. We plan to share data with Matocq's group and are interested in comparing data with Waits, Miyasaki, and others working in Idaho. John Wehausen, Mike Buchalski, Rob Ramey, and I have been working on some phylogeography/phylogenetics of bighorn sheep; this work could be expanded.

Answer IV (Jahner): I am currently working on generating a next-generation sequencing dataset to evaluate the degree of differentiation between populations of Rocky Mountain and 'California' bighorn sheep in Nevada. However, the only way to truly answer to this question is to do a broad, range-wide genetic study of populations across the entire range of the two putative subspecies, preferably using remnant populations that have not been heavily influenced by translocations. Additionally, a phylogenetic study including all of the named varieties of wild sheep in North America could provide complementary insights the history of differentiation. Indeed, as we pointed out in Malaney et al. (2015), the only way to address this question is by sampling across an west-to-east transect in the Canadian portion of the distribution. Our high level of differentiation within Nevada between 'California' and Rocky Mountain bighorn may be the result of genetically distinct groups in the native/northern range, which could manifest on the landscape as a fairly sharp transition between distinct genetic groups. Alternatively, Nevada's sheep could simply be the result of having sampled at the western end ('California' source herds) and at the eastern end (Rocky Mountain source herds) of an otherwise broad pattern of isolation by distance.

Answer V (Waits): We have also genotyped ~100 individuals from British Columbia herds with 15 microsatellite loci and have a mixture of 'California'

and Rocky Mountain genetic groups, so I think it would be useful to discuss results with other labs.

10) *For jurisdictions that have known hybridization occurring among wild sheep subspecies, what information should we be collecting to describe the consequences, and how should managers treat new or potential occurrences of hybridization?*

Answer I (Cox): Nevada and Arizona both have desert/Rocky Mountain hybrids. We are collecting muscle tissue from every ram harvested in our Nevada desert/Rocky herd. I think we need to accept certain level of hybridization-- but yes, we also need to collect data to monitor its progression and any adaptations we document. We have had discussions in the past on how best to deal with hybrid herds (promote or eliminate) and no agreement was made; each situation and set of circumstances is different.

Answer II (Waits): Based on our current results, British Columbia has California/Rocky Mountain bighorn hybrids.

Answer III (Schwantje): Some of those hybrid animals may have been translocated into US herds, again emphasizing the need for targeted sampling in British Columbia.

Answer IV (Matocq): This is a point that will require close communication with the hunting community. I have presented our hybridization results directly to members of Nevada Bighorns Unlimited, and they were enthralled with the biology and natural history that these genetic patterns suggest. Understanding the genomic and phenotypic consequences of hybridization should be an important goal for our group, in addition to non-hybrid genotype-phenotype relationships in this system. Sportsmen/women seem to be as eager as we are to learn about these relationships.

11) *Are managers on the right track by managing as separate taxa the various 'types' of bighorn sheep we currently recognize? We seem well past the days of depending on traditional taxonomy…do quantifiable differences at the molecular level signify adaptive differences that we had best not lose, or alternatively, signify drift (or worse yet, loss of alleles associated with small population size) that we would best attempt to counteract? Geneticists may not all agree about the relative risks of these two. So, what do we know now, and what do we need to learn to better figure it out?*

Answer I (Epps): We are still working this out, but I feel we are in position to give good guidance in many areas of bighorn sheep range. Our recent SNP-based landscape genomics study (Buchalski et al. *in prep*) and other analyses (e.g., Buchalski et al. 2016) suggest long-term separation between some regional populations of bighorn sheep, including those now considered a single subspecies (e.g., "Peninsular" and "Mojave" populations of desert bighorn). In absence of other information, I think applying the precautionary principle to future translocations is wise- use the nearest source available, and try not to mix lineages- but check with the latest research about what "lineages" are supported!

12) *How do we consider the above in relation to our questions about the*

importance of genetic diversity and managing for same? In other words, when might we consider mixing individuals from more distantly related groups in cases where we think genetic diversity is extremely low and locally-appropriate stock for augmentation may not be available?

Answer I (Cox): This concern by managers is becoming prevalent westwide. We do need help in deciding what best alternatives we have for mixing if diversity is low but when we do not have source animals that are well adapted to a particular mountain range and its climate, topography, forage base, and other conditions.

Answer II (Epps): I recommend being very cautious before considering augmentation from distantly-related populations to increase genetic diversity. Here are some considerations for this type of situation:

- Is the population performing well? If so, I would not be quick to intervene just because genetic diversity is low.
- Is the population performing poorly due to documented problems with disease, predation, or some other external cause? Genetic factors could play a role, as for disease, but probably are not the fundemental problem and thus unlikely to be a quick fix.
- I would be particularly reluctant to intervene in a native (or "remnant") population not originating from a translocation. Bringing in animals from other ecosystems raises the possibility of outbreeding depression (i.e., disrupting suites of genes acting in concert due to local adaptation).
- More obvious cases for intervention would be populations that are artificially isolated by translocation history or enclosures, and when phyical anomalies are observed.
- Systems of reintroduced populations that have suffered from sequential founder effects are probably the best place to learn from experimentation, as was done in Oregon (Olson et al. 2012).

Questions about disease

13) Who is doing what to investigate wild sheep genetics that control or influence immune response to virulent pathogens or similarly measuring animal resilience, resistance, or stress indicators? Should jurisdictions be collecting and contributing samples to this research?

Answer I (Bowen): USGS (Bowen) and NDOW have been measuring molecular immune response, and response to environmental stressors. Additional samples from jurisdictions with distinct questions would be excellent. Note: we are at the very beginning of this line of research (*i.e.*, it is not a crystal ball).

Answer II (Waits): We recently published a paper that addressed this to some degree (Plowright et al. 2017) and plan to follow up with a GWAS study using our SNP data for the Lostine herd of Rocky Mountain bighorn sheep in Oregon.

Answer III (Epps): With collaborators at Oregon State University, California Department of Fish and Wildlife, and the National Park Service, as part of the study of respiratory disease in desert

bighorn in the Mojave National Preserve, we have collected data on *Mycoplasma ovipneumoniae* infection, genetic diversity at neutral and adaptive-linked microsatellite markers, and measures of immune phenotype to examine correlations among these measures. We are collecting similar data in 'California' bighorn populations in Oregon. Analyses are ongoing but we hope to submit a manuscript on the Mojave system soon. Recently we published a paper investigating differences in immune phenotype in Peninsular and Mojave bighorn sheep (Dugovich et al. 2017), in which we found higher levels of natural antibody in blood of bighorn sheep from the Peninsular Ranges compared to the Mojave. We also observed markedly greater ability of plasma from Peninsular bighorn sheep to kill *E. coli* bacteria *in vitro* compared to plasma from bighorn sheep in the Mojave. Those differences suggest that important geographical variation in immune response can exist.

14) *What do we currently know about the genetic basis for (or genetic correlations with) susceptibility to, and response if infected by, pathogens leading to chronic pneumonia? If we do not know as much as we would like (which I suspect is true), what practical actions can we take to learn more?*

Answer I (Cox): We need to bring together a subset of geneticists and managers that have collected a great deal data already on pathogen profiles and herd responses and see if there is a logical next horizon to explore. This may include digging deeper into the animal genetics for susceptibility and recovery and maybe even setting up an experiment of sorts to challenge animals, if we have candidate herds that have unique genetics or traits and past responses.

Answer II (Cassirer): I think it would be important to bring some immunological expertise to the discussion to identify specific genes or loci that we should be looking at for this particular disease. As far as that goes, I think we also need to know more about pneumonia, why bighorn sheep are so susceptible and what parts of the immune system are (likely) responsible for that.

Answer III (Epps): We are seeing some interesting correlations between genetic measures, immune phenotype, and disease in the Mojave system, but this is work in progress and needs to go through peer review.

Answer IV (Bowen): Our group is working with transcription of immunological genes. We have a long way to go and need input from all of you.

CONCLUSIONS

This exercise was intended to improve communication and facilitate collaboration among managers and researchers. Some concensus has appeared regarding important research directions: these include resolving taxonomy among lineages of wild sheep (particularly 'California' versus Rocky Mountain bighorn), research to inform managing isolated, remnant, or genetically depauperate populations, and the need to clarify the utility of specific data types to answer different questions, for instance via parallel analyses of microsatellites and SNPs, using SNPs on non-invasive samples, or evaluating ways to merge SNP datasets generated from different labs. The discussion of ongoing

research projects has already led to the realization that there are opportunities for better data sharing and collaboration.

Beyond the questions addressed above, the Wild Sheep Genomics Working group has considered procedural questions about the way forward. In particular: what efforts or standards need to be done or met to promote collaboration and comparing of wild sheep genomics west-wide to help answer many of the questions we have in common? Is there a need to develop a broad west-wide plan of data sharing for different genetics questions, and are there any geneticists willing to take on (with funding) the meta-analyses of west-wide datasets vs. each jurisdiction doing their own work? Finally, can we create a simple inventory list of current wild sheep genetics research in order to identify opportunities for collaboration or sharing of resources?

In the discussions associated with preparing this document, interest has been expressed that we should seek funding to draft a data sharing plan to encourage range (west)-wide analyses. An introduction to this idea was presented to the Wild Sheep Working Group at the Western Association of Fish and Wildlife Agencies (WAFWA) meeting in Eugene, OR in July 2018 (Epps, Cox). Matocq proposed that investing in the generation of a high quality bighorn genome would be an important next step. That, combined with the reduced representation (SNP) datasets already being generated range-wide will provide the needed information to generate a bighorn-specific SNP array that captures the specific loci needed to address the breadth of questions that interest this group. She further proposed that we seek funds to pay for 1-2 Postdoctoral Scholars that rotate among the key labs involved in the effort so that approaches and analyses are necessarily shared and consistent across labs. Finally, establishing a database of ongoing research projects was proposed to facilitate future collaborations. We expect that the Wild Sheep Genomic Working Group will continue to seek opportunities for discussing and developing these approaches.

ACKNOWLEDGEMENTS

We would like to acknowledge input from all the participants in the Wild Sheep Genomics Working Group, particularly those who contributed to the January 2018 survey or participated in the initial conference call, as well as members of the Wild Sheep Working Group of the Western Association of Fish and Wildlife Agencies.

LITERATURE CITED

Alberdi, A., O. Aizpurua, K. Bohmann, M. L. Zepeda-Mendoza, and M. T. P. Gilbert. 2016. Do vertebrate gut metagenomes confer rapid ecological adaptation? Trends in Ecology & Evolution 31:689-699.

Andrews, K. R., J. R. Adams, E. F. Cassirer, R. K. Plowright, C. Gardner, M. Dwire, P. A. Hohenlohe, and L. P. Waits. 2018. A bioinformatic pipeline for identifying informative SNP panels for parentage assignment from RADseq data. Molecular Ecology Resources Online Early:1-19.

Andrews, K. R., J. M. Good, M. R. Miller, G. Luikart, and P. A. Hohenlohe. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics 17:81-92.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, Massachusetts, USA.

Boyce, W. M., P. W. Hedrick, N. E. MuggliCockett, S. Kalinowski, M. C. T. Penedo, and R. R. Ramey. 1997. Genetic variation of major histocompatibility complex and micrcosatellite loci: A comparison in bighorn sheep. Genetics 145:421-433.

Boyce, W. M., R. R. Ramey, T. C. Rodwell, E. S. Rubin, and R. S. Singer. 1999. Population subdivision among desert bighorn sheep (*Ovis canadensis*) ewes revealed by mitochondrial DNA analysis. Molecular Ecology 8:99-106.

Bredeson, J. V., J. B. Lyons, S. E. Prochnik, G. A. Wu, C. M. Ha, E. Edsinger-Gonzales, J. Grimwood, J. Schmutz, I. Y. Rabbi, C. Egesi, P. Nauluvula, V. Lebot, J. Ndunguru, G. Mkamilo, R. S. Bart, T. L. Setter, R. M. Gleadow, P. Kulakow, M. E. Ferguson, S. Rounsley, and D. S. Rokhsar. 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. Nature Biotechnology 34:562-570.

Buchalski, M. R., A. Y. Navarro, W. M. Boyce, T. W. Vickers, M. W. Tobler, L. A. Nordstrom, J. A. Garcia, D. A. Gille, M. C. T. Penedo, O. A. Ryder, and H. B. Ernest. 2015. Genetic population structure of Peninsular bighorn sheep (*Ovis canadensis nelsoni*) indicates substantial gene flow across US-Mexico border. Biological Conservation 184:218-228.

Buchalski, M. R., B. N. Sacks, D. A. Gille, M. C. T. Penedo, H. B. Ernest, S. A. Morrison, and W. M. Boyce. 2016. Phylogeographic and population genetic structure of bighorn sheep (*Ovis canadensis*) in North American deserts. Journal of Mammalogy 97:823-838.

Carroll, E. L., M. W. Bruford, J. A. DeWoody, G. Leroy, A. Strand, L. Waits, and J. Wang. 2018. Genetic and genomic monitoring with minimally invasive sampling methods. Evolutionary Applications 11:1094-1119.

Coltman, D. W., M. Festa-Bianchet, J. T. Jorgenson, and C. Strobeck. 2002. Age-dependent sexual selection in bighorn rams. Proceedings of the Royal Society B-Biological Sciences 269:165-172.

Creech, T. G., C. W. Epps, E. L. Landguth, J. D. Wehausen, R. S. Crowhurst, B. Holton, and R. J. Monello. 2017. Simulating the spread of selection-driven genotypes using landscape resistance models for desert bighorn sheep. Plos One 12.

Creech, T. G., C. W. Epps, R. J. Monello, and J. D. Wehausen. 2014. Using network theory to prioritize management in a desert bighorn sheep metapopulation. Landscape Ecology 29:605-619.

De Barba, M., C. Miquel, S. Lobreaux, P. Y. Quenette, J. E. Swenson, and P. Taberlet. 2017. High-throughput microsatellite genotyping in ecology: improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. Molecular Ecology Resources 17:492-507.

Driscoll, C. C., J. G. Driscoll, C. Hazekamp, J. B. Mitton, and J. D. Wehausen. 2015. A tale of two markers: Population genetics of Colorado Rocky Mountain bighorn sheep estimated from microsatellite and

mitochondrial data. Journal of Wildlife Management 79:819-831.

Dugovich, B. S., M. J. Peel, A. L. Palmer, R. A. Zielke, A. E. Sikora, B. R. Beechler, A. E. Jolles, C. W. Epps, and B. P. Dolan. 2017. Detection of bacterial-reactive natural IgM antibodies in desert bighorn sheep populations. Plos One 12:e0180415.

Epps, C. W., R. S. Crowhurst, and B. S. Nickerson. 2018. Assessing changes in functional connectivity in a desert bighorn sheep metapopulation after two generations. Molecular Ecology 27:2334-2346.

Epps, C. W., P. J. Palsboll, J. D. Wehausen, G. K. Roderick, R. R. Ramey, and D. R. McCullough. 2005. Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. Ecology Letters 8:1029-1038.

Epps, C. W., J. D. Wehausen, V. C. Bleich, S. G. Torres, and J. S. Brashares. 2007. Optimizing dispersal and corridor models using landscape genetics. Journal of Applied Ecology 44:714-724.

Epps, C. W., J. D. Wehausen, P. J. Palsboll, and D. R. McCullough. 2010. Using genetic tools to track desert bighorn sheep colonizations. Journal of Wildlife Management 74:522-531.

Epps, C. W., J. D. Wehausen, W. B. Sloan, S. Holt, T. G. Creech, R. S. Crowhurst, J. R. Jaeger, K. M. Longshore, and R. J. Monello. 2016. Fifty years after Welles and Welles: distribution and genetic structure of desert bighorn sheep in Death Valley National Park. Pages 70-91 *in* G. White, editor. First Death Valley Natural History Conference Proceedings. Death Valley Natural History Association, Death Valley, California.

Fitzsimmons, N. N., S. W. Buskirk, and M. H. Smith. 1995. Population history, genetic variability, and horn growth in bighorn sheep. Conservation Biology 9:314-323.

Flesch, A. D., C. W. Epps, J. W. Cain III, M. Clark, P. R. Krausman, and J. R. Morgart. 2010. Potential effects of the United States-Mexico border fence on wildlife. Conservation Biology 24:171-181.

Gille, D. A., M. R. Buchalski, D. Conrad, E. S. Rubin, A. Munig, B. Wakeling, C. W. Epps, T. G. Creech, R. S. Crowhurst, B. Holton, R. J. Monello, W. M. Boyce, C. Penedo, and H. B. Ernest. *In press*. Genetic outcomes of translocation of bighorn sheep in Arizona. Journal of Wildlife Management.

Gutierrez-Espeleta, G. A., S. T. Kalinowski, W. M. Boyce, and P. W. Hedrick. 2000. Genetic variation and population structure in desert bighorn sheep: implications for conservation. Conservation Genetics 1:3-15.

Hedrick, P. W., G. A. Gutierrez-Espeleta, and R. N. Lee. 2001. Founder effect in an island population of bighorn sheep. Molecular Ecology 10:851-857.

Hedrick, P. W., and J. D. Wehausen. 2014. Desert bighorn sheep: changes in genetic variation over time and the impact of merging populations. Journal of Fish and Wildlife Management 5:3-13.

Hogg, J. T., S. H. Forbes, B. M. Steele, and G. Luikart. 2006. Genetic rescue of an insular population of large mammals. Proceedings of the Royal

Society B-Biological Sciences 273:1491-1500.

Jahner, J. P., M. D. Matocq, J. L. Malaney, M. Cox, P. Wolff, M. Gritts, and T. L. Parchman. *In press*. The genetic legacy of 50 years of desert bighorn sheep translocations. Evolutionary Applications.

Johnson, H. E., L. S. Mills, J. D. Wehausen, T. R. Stephenson, and G. Luikart. 2011. Translating effects of inbreeding depression on component vital rates to overall population growth in endangered bighorn sheep. Conservation Biology 25:1240-1249.

Kardos, M., G. Luikart, R. Bunch, S. Dewey, W. Edwards, S. McWilliam, J. Stephenson, F. W. Allendorf, J. T. Hogg, and J. Kijas. 2015. Whole-genome resequencing uncovers molecular signatures of natural and sexual selection in wild bighorn sheep. Molecular Ecology 24:5616-5632.

Kohl, K. D. 2017. An introductory "how-to" guide for incorporating microbiome research into integrative and comparative biology. Integrative and Comparative Biology 57:674-681.

Luikart, G., and F. W. Allendorf. 1996. Mitochondrial-DNA variation and genetic-population structure in rocky mountain bighorn sheep (*Ovis canadensis canadensis*). Journal of Mammalogy 77:109-123.

Luikart, G., S. J. Amish, J. Winnie, A. Beja-Pereira, R. Godinho, F. W. Allendorf, and R. B. Harris. 2011. High connectivity among argali sheep from Afghanistan and adjacent countries: Inferences from neutral and candidate gene microsatellites. Conservation Genetics 12:921-931.

Luikart, G., K. Pilgrim, J. Visty, V. O. Ezenwa, and M. K. Schwartz. 2008a. Candidate gene microsatellite variation is associated with parasitism in wild bighorn sheep. Biology Letters 4:228-231.

Luikart, G., S. Zundel, D. Rioux, C. Miquel, K. A. Keating, J. T. Hogg, B. Steele, K. Foresman, and P. Taberlet. 2008b. Low genotyping error rates and noninvasive sampling in bighorn sheep. Journal of Wildlife Management 72:299-304.

Malaney, J. L., C. R. Feldman, M. Cox, P. Wolff, J. D. Wehausen, and M. D. Matocq. 2015. Translocated to the fringe: genetic and niche variation in bighorn sheep of the Great Basin and northern Mojave deserts. Diversity and Distributions 21:1063-1074.

Mardis, E. R. 2013. Next-generation sequencing platforms. Pages 287-303 *in* R. G. Cooks, andJ. E. Pemberton, editors. Annual Review of Analytical Chemistry, Vol 6. Annual Reviews, Palo Alto.

Miller, J. M., M. Festa-Bianchet, and D. W. Coltman. 2018. Genomic analysis of morphometric traits in bighorn sheep using the Ovine Infinium (R) HD SNP BeadChip. Peerj 6:20.

Miller, J. M., R. M. Malenfant, P. David, C. S. Davis, J. Poissant, J. T. Hogg, M. Festa-Bianchet, and D. W. Coltman. 2014. Estimating genome-wide heterozygosity: effects of demographic history and marker type. Heredity 112:240-247.

Miller, J. M., S. S. Moore, P. Stothard, X. Liao, and D. W. Coltman. 2015. Harnessing cross-species alignment to discover SNPs and generate a draft genome sequence of a bighorn

sheep (*Ovis canadensis*). BMC Genomics 16:397.

Miller, J. M., J. Poissant, J. T. Hogg, and D. W. Coltman. 2012. Genomic consequences of genetic rescue in an insular population of bighorn sheep (*Ovis canadensis*). Molecular Ecology 21:1583-1596.

Miller, J. M., J. Poissant, J. W. Kijas, D. W. Coltman, and C. Int Sheep Genomics. 2011. A genome-wide set of SNPs detects population substructure and long range linkage disequilibrium in wild sheep. Molecular Ecology Resources 11:314-322.

Olson, Z. H., D. G. Whittaker, and O. E. Rhodes. 2012. Evaluation of experimental genetic management in reintroduced bighorn sheep. Ecology and Evolution 2:429-443.

_____. 2013. Translocation history and genetic diversity in reintroduced bighorn sheep. Journal of Wildlife Management 77:1553-1563.

Parchman, T. L., Z. Gompert, J. Mudge, F. D. Schilkey, C. W. Benkman, and C. A. Buerkle. 2012. Genome-wide association genetics of an adaptive trait in lodgepole pine. Molecular Ecology 21:2991-3005.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyipng in model and non-model species. Plos One 7:11.

Plowright, R. K., K. R. Manlove, T. E. Besser, D. J. Paez, K. R. Andrews, P. E. Matthews, L. P. Waits, P. J. Hudson, and E. F. Cassirer. 2017. Age-specific infectious period shapes dynamics of pneumonia in bighorn sheep. Ecology Letters 20:1325-1336.

Pompanon, F., B. E. Deagle, W. O. C. Symondson, D. S. Brown, S. N. Jarman, and P. Taberlet. 2012. Who is eating what: diet assessment using next generation sequencing. Molecular Ecology 21:1931-1950.

Putnam, N. H., B. L. O'Connell, J. C. Stites, B. J. Rice, M. Blanchette, R. Calef, C. J. Troll, A. Fields, P. D. Hartley, C. W. Sugnet, D. Haussler, D. S. Rokhsar, and R. E. Green. 2016. Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. Genome Research 26:342-350.

Ramey II, R. R. 1995. Mitochondrial DNA variation, population structure, and evolution of mountain sheep in the southwestern United States and Mexico. Molecular Ecology 4:429-439.

Ramey II, R. R., G. Luikart, and F. J. Singer. 2000. Genetic bottlenecks resulting from restoration efforts: The case of bighorn sheep in Badlands National Park. Restoration Ecology 8:85-90.

Roffler, G. H., S. J. Amish, S. Smith, T. Cosart, M. Kardos, M. K. Schwartz, and G. Luikart. 2016. SNP discovery in candidate adaptive genes using exon capture in a free-ranging alpine ungulate. Molecular Ecology Resources 16:1147-1164.

Russello, M. A., M. D. Waterhouse, P. D. Etter, and E. A. Johnson. 2015. From promise to practice: pairing non-invasive sampling with genomics in conservation. Peerj 3:18.

Sim, Z., J. C. Hall, B. Jex, T. M. Hegel, and D. W. Coltman. 2016. Genome-wide set of SNPs reveals evidence for two glacial refugia and admixture from

postglacial recolonization in an alpine ungulate. Molecular Ecology 25:3696-3705.

U.S. Fish and Wildlife Service. 2007. Recovery plan for the Sierra Nevada bighorn sheep., U. S. Fish and Wildlife Service, Sacramento, California.

Waits, L., and C. W. Epps. 2015. Population genetics and wildlife habitat. Pages 63-83 *in* M. L. Morrison, andH. A. Mathewson, editors. Wildlife habitat conservation: concepts, challenges, and solutions. Johns Hopkins University Press, Baltimore, USA.

Wang, S., E. Meyer, J. K. McKay, and M. V. Matz. 2012. 2b-RAD: a simple and flexible method for genome-wide genotyping. Nature Methods 9:808-+.

Weeks, A. R., J. Stoklosa, and A. A. Hoffmann. 2016. Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: the case of Australian mammals. Frontiers in Zoology 13:9.

Wehausen, J. D., and R. R. Ramey II. 2000. Cranial morphometric and evolutionary relationships in the northern range of *Ovis canadensis*. Journal of Mammalogy 81:145-161.

Wehausen, J. D., R. R. Ramey II, and C. W. Epps. 2004. Experiments in DNA extraction and PCR amplification from bighorn sheep feces: the importance of DNA extraction method. Journal of Heredity 95:503-509.

Whittaker, D. G., S. D. Ostermann, and W. M. Boyce. 2004. Genetic variability of reintroduced California bighorn sheep in Oregon. Journal of Wildlife Management 68:850-859.

Wilder, B. T., J. L. Betancourt, C. W. Epps, R. S. Crowhurst, J. I. Mead, and E. Ezcurra. 2014. Local extinction and unintentional rewilding of bighorn sheep (*Ovis canadensis*) on a desert island. Plos One 9:9.

APENDIX 1: GLOSSARY OF TERMS USED IN GENETIC AND GENOMIC RESEARCH ON WILD SHEEP. Definitions denoted by a superscript 1 were adapted from Waits and Epps (2015).

Admixture: interbreeding of two or more previously distinct genetic lineages.

Allele: a variant at a gene or locus.

DNA sequence: a representation of the actual sequence of nucleotides in a molecule of DNA, rather than an estimate of fragment number or size (as in microsatellites or RFLPs).

Heterozygosity: the degree of genetic variation in individuals or populations, as measured at variable sites (loci or genes). To be comparable, heterozygosity estimates should incorporate the same or at least similar types of loci.

[1]**Gene flow:** the movement of alleles between populations, also referred to as "migration" in the context of population genetics.

[1]**Genetic structure:** spatial variation in the frequency or identity of alleles

Genome: the complete DNA sequence of an organism.

Genome-wide association study (GWAS): A research approach that assesses variation at a large number of variable markers, e.g., SNPs, across the genomes of a large number of individuals, and then tests for association of alleles with phenotypes or traits in the sampled individuals. This approach can help identify genes or particular alleles associated with traits such as disease risk.

[1]**Hybridization:** interbreeding of individuals from genetically distinct populations, subspecies, or species.

Landscape genomics: a research approach that attempts to identify signals of selection on variable genetic markers while controlling for genetic differences resulting from isolation, drift, and phylogeographic history. The approach can help identify genes or variants associated with different local adaptations, but has methodological challenges.

Locus (plural, loci): a variable site in the genome, i.e., one that differs among individuals included in a particular analysis.

Major histocompatibility complex (MHC): a group of genes that codes for proteins involved with acquired immunity; those proteins help the immune system identify foreign material. More genetic variation at these genes, in theory, increases the chance that a specific threat can be recognized and neutralized by the immune system.

Microsatellite: sometimes referred to as short tandem repeats (STRs) or simple sequence repeats (SSRs), this type of marker includes repeating sequences of two to six base pairs. Mutations resulting in different numbers of repeated elements at a particular locus are common, making these markers suitable for analyses where individual-level variation is useful.

- **Adaptive-linked:** a locus where the frequency of alleles may be affected by selection on a nearby gene.

-[1]**Neutral:** a locus where the frequency of alleles is not affected by selection.

[1]**Mitochondrial DNA (mtDNA):** a circular DNA molecule found in the mitochondria of cells; mtDNA is haploid (only one copy) and generally is inherited only from the mother.

Ovine SNP chip (array): an assay designed for assessing SNP variation in domestic sheep which has been applied to wild sheep, albeit only a subset of loci are variable (Miller et al. 2011). A newer version of the array provides more variable loci (Miller et al. 2018).

Phylogeography: the study of the spatial arrangement of genealogical lineages, especially within and among conspecific populations and closely related species (Avise 2000).

Restriction fragment length polymorphism (RFLP): a type of analysis in which DNA is cut using enzymes that recognize particular short sequences and the size of the resulting fragments is visualized; variation in sequence among individuals leads to different patterns of cutting and thus different sized fragments. Rarely used anymore.

Restriction site-associated DNA sequencing (RADseq): reduced representation sequencing method that sequences a subsample of genomic regions guided by the genomic location and frequency of restriction enzymes. Individual DNA samples are labeled by attaching DNA barcodes, allowing large number of individuals to be pooled and simultaneously sequenced on individual lanes of the Illumina platform. These methods have revolutionized population genetics and made population genomic scale data readily available for virtually any organism. Genotyping By Sequencing (GBS) is another name often used for the same type of method (see Andrews et al. 2016 for a thorough review).

Single Nucleotide Polymorphism (SNP): variation at a single base pair of DNA. Methods such as RADseq or SNP arrays can identify or assess variation at thousands of variable sites across the genome. While the information content of a single SNP locus is lower than that of a single microsatellite locus, SNP loci can be efficiently assessed in much larger numbers for some types of DNA samples (Dugovich et al. 2017).