Genetic Re-assessment of Population Subdivision in Yellowstone National Park Bison

Sam Stroupe[1], Chris Geremia[2], Rick L. Wallen[2], P. J. White[2], James N. Derr[1]

[1]Department of Veterinary Pathobiology, Texas A&M University College of Veterinary Medicine and Biomedical Science, College Station, Texas, USA

[2]Yellowstone Center for Resources, Yellowstone National Park, Mammoth Hot Springs, Wyoming, USA

*Corresponding Author: James N. Derr j-derr@tamu.edu

Keywords: North American bison, *Bison bison*, Yellowstone National Park, population structure, wildlife management

Abstract

Yellowstone National Park is home to the only plains bison population that has continually existed as wildlife, on the same landscape, through the population bottleneck of the late 19th century. Nevertheless, by the early 1900s, only 23 wild bison were known to have survived poaching. Salvation efforts included the addition of 18 females from Montana and 3 bulls from Texas to augment this population. A century later, nuclear microsatellite-based population level assessment revealed two genetically distinct bison sub-populations. However, in 2016 an analysis of mitochondrial haplotypes showed the two founding lineages were distributed throughout the park. This study is designed to delineate any current sub-structure in the Yellowstone bison population by strategically sampling the two major summer breeding herds and the two major winter ranges. Population level metrics were derived using the same microsatellite loci as the original study along with a newly developed set of highly informative bison specific Single Nucleotide Polymorphisms (SNPs). Our analyses reveal that the modern bison in Yellowstone National Park currently consist of one interbreeding metapopulation. Our findings have important conservation and management implications on efforts to restore this large and historically important bison population.

Introduction

Yellowstone National Park is home to arguably the most iconic bison population. Since its establishment in 1872 as the first National Park in the United States and the world, significant and unwavering efforts have been dedicated to the preservation of the bison population residing there. Not only has this population played a part in bison becoming a national icon, but they are also a true testament to the resilience, hope, and strength bison symbolize. Yellowstone bison are the largest free-roaming bison population in the United States, and the only plains bison to have continually existed in the wild and on same landscape since prehistoric times.

The history of this population is a tale of near tragedy turned conservation success story. During the population bottleneck of the late 19th century, the Greater Yellowstone Area bison population narrowly avoided extinction, with an estimated 23 individuals remaining in 1902 (American Bison Society, 1908). To preserve this population, additional bison were brought in from private ranches. Eighteen females from the Pablo-Allard herd in Montana, three males from the Goodnight herd in Texas, and four calves from the indigenous herd were used to establish a secondary "introduced" population (Coder, 1975; M. Meagher, 1973). The introduced herd was moved to the Lamar Valley in 1907 and closely day herded or corralled through at least 1915 (M. Meagher, 1973). Meanwhile the indigenous herd was isolated, wintering in the Pelican Valley and summering in the high elevation grasslands of the Upper Lamar River.

The subsequent recovery of Yellowstone bison is well documented (M. Meagher, 1973; White et al., 2015). Intermingling of the native and introduced herds increased after 1921 when managers fenced the introduced herd out of the Lamar Valley and moved them to areas used by the indigenous herd during summer. Managers also preferentially removed adult males from the introduced herd to increase the indigenous lineage. The annual report of Superintendent Toll in 1929 indicated "[t]here seems to be a gradual intermingling of the wild [Pelican] and tame [Lamar] herd. It has reached a point where it is difficult to distinguish the buffalo of the wild herd from those of the Lamar Valley herd" (Toll, 1929). Therefore, it is likely the bison formed a single "Northern Herd" by the 1930s summering together and separating into two wintering units called the "Lamar" and "Pelican" bison (Meagher, 1973). In 1936, managers relocated 71 bison from the Lamar bison to the Firehole and Hayden Valleys (Skinner & Alcorn, 1942). The animals formed the "Mary Mountain Bison" or "Central Herd." Population reductions and subsequent recovery likely kept the Northern and Central Herds separated through the 1970s. Movements between the herds were believed to increase through the 1980s when Northern Herd animals wintering in the Pelican Valley began moving to the Hayden Valley and integrating into the Central Herd (Meagher, 1993, 1998). By the 1990s, the animals from the Central Herd began moving to wintering areas of the Northern Herd (Meagher, 1989, 1993, 1998). Today, all Yellowstone bison roam relatively freely within Yellowstone National Park, with limitations on their distribution into areas outside the park boundaries. During the past decade, the population has fluctuated between 4000 and 6000 animals. Modern GPS technology confirms that bison congregate into two main herds (Northern, Central) during the breeding season, some animals from each herd share wintering areas, and some individuals switch breeding areas over time.

In a study by Halbert et al. (2012a), evaluating 661 Yellowstone bison sampled between 1997 and 2003, Yellowstone bison appeared to split into two genetically distinct subpopulations

defined by microsatellite genotype diversity and allelic distributions. At that time, a concern raised by Halbert et al. (2012a) involved unequal culling across these two bison populations and the impact this could have on long-term genetic diversity. Subsequently, White and Wallen (2012) suggested "managers should be promoting the conservation of wildness and natural selection to retain adaptive capabilities, rather than preconceived notions of "natural" genetic or populations structures that were likely created or exacerbated by human actions". As this discussion continued, Halbert et. al. (2012b) pointed out that since all modern bison populations are due to anthropogenic activities, it may be best to error on the side of caution and focus Yellowstone bison management to preserve genetic diversity in its, at the time, current state (Halbert et al., 2012b). While no definitive conclusion was reached at the time, it was agreed that there should be continued monitoring of genetic variation in Yellowstone bison and use the best available scientific data as a foundation for future bison management.

Following the study by Halbert et al. (2012a), a mitochondrial genome study was conducted to analyze haplotype diversity among Yellowstone bison (Forgacs et al., 2016). In this study 25 mitochondrial genomes of Yellowstone bison were evaluated and compared to an additional 20 bison from diverse populations. This study revealed two distinct mitochondrial haplotype clades found within Yellowstone and the overall population of bison. This evidence was consistent with historical documentation of a second maternal lineage brought in from the Pablo-Allard herd in the early 1900s (Coder, 1975; Meagher, 1973). Unlike the previous study (Halbert et al., 2012a), there was no evidence, based on mtDNA haplotypes, to support population subdivision among the Yellowstone bison. In fact, this study was able to identify that bison of both mitochondrial haplotype clades, and therefore maternal lineages, were present in both present day breeding areas (Forgacs et al., 2016).

However, both studies had limitations to their scope. Namely neither study included Yellowstone bison sampled from both breeding populations during the same breeding season. In the study conducted by Halbert et al. (2012a), although a portion of the samples were collected within the park boundaries, over 90% of the samples were collected during the winter migration when bison were leaving the park. Considering the northern and central groups have an overlapping winter range in the Gardiner Basin, it was not possible to confidently assign home ranges without associated individual movement data such as radio telemetry (Halbert et al., 2012a; Meagher, 1989). Therefore, no definitive conclusions could be drawn as to whether the identified genetic population substructure extended to the breeding groups. In the work of Forgacs et al. (2016), the mitochondrial DNA only portrays a portion of the genetic story. While able to assign bison to maternal lineages and compare life histories, it failed to characterize levels of admixture within these individuals.

This study was designed to delineate any current sub-structure in the Yellowstone bison population by strategically sampling the two major summer breeding populations, where one would expect to see genetic differentiation in ongoing population subdivision, and the two major winter ranges, where migration patterns of both breeding herds can overlap. Metrics used to evaluate population dynamics were determined using a set of 24 microsatellite loci, previously characterized in Yellowstone bison, as well as a newly developed set of highly informative bison specific Single Nucleotide Polymorphisms (SNPs) (Stroupe et al. submitted). Identifying current

structure of the Yellowstone bison population is important to understanding the history of this iconic population, as well as developing strategies for its future conservation.

Methods

Biological Material

Tissue biopsy samples were collected from 282 Yellowstone National Park bison. These samples were collected during the summer of 2019 and winter of 2021, in two major ranges within the park, Central and Northern (Figure 1). The summer collection included 154 samples, 62 from the Central breeding area and 92 from the Northern breeding area. The winter collection included 128 samples, 44 from the Central wintering area and 84 from the Northern wintering area. Within each geographic areas bison groups were opportunistically located during sampling periods with up to 30% of group members randomly selected for sampling. Tissue samples were collected using a Type P Pneu-Dart Biopsy RDD. Tissue samples were extracted from the dart using forceps cleaned in ethanol. Tissue samples were suspended in a 10% ethanol solution and refrigerated prior to lab analysis. Biopsy tissue samples were collected according to the NPS IACUC IMR_YELL_White_Ungulates_2022.A3.

Tissue samples were cut into smaller portions in an isolation hood and divided for microsatellite or SNP genotyping. Of the 282 samples, eight were identified as duplicates with SNP data using Sequoia v2.5.3 (Huisman, 2017) and confirmed with microsatellite data. Duplicates were removed from final data analysis.

Microsatellites

DNA was extracted from a portion of each biopsy tissue sample using the Gentra Puregene kit (Qiagen) according to manufacture protocols. A set of 24 microsatellite markers were used in this study including the core set of loci used for parentage determination (Halbert, 2003; Schnabel et al., 2000). Samples were genotyped according to established lab protocol (Halbert, 2003) using an ABI 3730 Genetic Analyzer and STRand software (Toonen & Hughes, 2001; https://vgl.ucdavis.edu/STRand). Three samples had a call rate below 90% and were therefore removed from further analysis resulting in a final microsatellite dataset of 271 samples and 24 microsatellite markers. Of those, 148 samples were collected during the summer breeding season.

Cervus v3.0.7 (Kalinowski et al., 2007) was used for initial overview of population genetic diversity and formatting data set for downstream analysis. A series of principal component analysis (PCA) were used to establish patterns in the data using collection location and time of year to group samples. Ade4 v1.7-22 (Dray & Dufour, 2007) was used to calculate eigenvalues and eigenvectors on the final dataset and plotted using ggplot (Wickham, 2016) in R v4.1.2 (R Core Team, 2022). Observed heterozygosity (H_O), expected heterozygosity within populations (H_S), allelic richness (A_R), F_{ST}, and F_{IS} with 95% confidence intervals were calculated with the Hierfstat R package (Goudet, 2005). F_{ST} was calculated using the Weir & Cockerham (Weir & Cockerham, 1984) equation. Values calculated per individual or per loci were averaged across each population. These analyses were carried out in R v4.1.2 (R Core Team, 2022).

Population parameters were evaluated according to the following methods. Evidence of population structure was assessed using STRUCTURE (Pritchard et al., 2000). The dataset was evaluated across 20 iterations at each *K* from 1 to 6. Notable changes in default program setting include a burn-in period of 40,000 replicates and 80,000 Markov chain Monte Carlo replicates. Structure Harvester was used to compile and format the resulting data (Earl & vonHoldt, 2012). Results were visualized with all iterations at each K aligned and merged using the R package pophelper v2.3.1 (Francis, 2017).

Single Nucleotide Polymorphism (SNPs)

Portions of the biopsy tissue samples were sent directly to NeoGen Canada for processing, DNA extraction, and genotyping. Samples were genotyped according to standard Illumnia Infinium HD Ultra Assay protocol guideline on the GGP Equine-Bison chip (NeoGen). Initial data quality control was done by NeoGen using Illumina's GenomeStudio with a call cutoff of 0.95. Details regarding the development of the SNP panel used in this study are described in Stroupe & Derr (submitted).

The Illumnia SNP Chip final report data was converted to plink lgen format using the script "illumina_to_lgen.R", originally written by Ryan Schubert (github@RyanSchu), with adjustments specific to our dataset. In addition, a plink fam file was created from the sample list. The data was then converted into VCF format using Plink v1.9 (Purcell et al., 2007). The nuclear and mitochondrial SNPs were then separated into different data sets for downstream analysis.

The nuclear SNP dataset was further thinned to remove monomorphic SNPs and SNPs with a genotyping call rate below 90% among the Yellowstone samples, using VCFtools v0.1.16 (Danecek et al., 2011). Two samples had a genotyping call rate below 90% and were therefore removed resulting in a final SNP dataset of 272 samples and 725 SNPs. Of those, 151 samples were collected during the summer breeding season.

A series of principal component analysis (PCA) were used to establish patterns in the data using collection location and time of year to group samples. Plink v1.9 (Chang et al., 2015; Purcell et al., 2007) was used to calculate eigenvalues and eigenvectors on the final dataset and plotted using ggplot (Wickham, 2016) in R v4.1.2 (R Core Team, 2022). Observed heterozygosity (H_O), expected heterozygosity within populations (H_S), allelic richness (A_R), F_{ST}, and F_{IS} with 95% confidence intervals were calculated with the Hierfstat R package (Goudet, 2005). F_{ST} was calculated using the Weir & Cockerham (1984) equation. Values calculated per individual or per loci were averaged across each population. These analyses were carried out in R v4.1.2 (R Core Team, 2022).

fastSTRUCTURE 1.0 (Raj et al., 2014) was then used to determine population structure among the samples with values of K=1 to K=6 and visualized using the R package pophelper v2.3.1 (Francis, 2017). fastSTRUCTURE chooseK.py was used to test for the most likely number of subpopulations.

Mitochondrial haplotype assignments were made according to Stroupe & Derr (submitted). Briefly, mitochondrial SNPs originally identified by Forgacs et al. (Forgacs et al., 2016) were

used to assign interspecific haplotypes (bison or domestic cattle) and intraspecific clade haplotypes (Clade I or Clade II).

Results

Biopsy tissue samples were collected from 282 bison from Yellowstone National Park during the summer breeding season of 2019 and the winter of 2021. Collection location was used to group bison into either the Central or Northern Herd according to previous studies, migration patterns, population history, and observations (Geremia et al., 2014; Halbert et al., 2012a; Meagher, 1973, 1989) (Figure 1). Collection time of year was used to distinguish samples during summer breeding season and winter migration. Of these samples, eight were identified as the same bison and removed from final analyses. All duplicate bison were sampled on the same range; however, two bison were sampled during both the summer and winter on the same range. In those cases, the sample collected during the summer breeding season was kept and the winter sample was removed from the dataset. Additionally, two samples from the SNP data and three samples from the microsatellite data were removed due to a genotyping call rate below 90%. The summer breeding season samples included 151 bison in the SNP dataset and 148 bison in microsatellite dataset.

Samples were genotyped at 24 microsatellites (Halbert, 2003; Schnabel et al., 2000), 798 autosomal SNPs, and 13 mitochondrial SNPs (Stroupe & Derr submitted). The autosomal SNP data was further filtered to remove monomorphic variants and those with a genotyping call rate below 90% resulting in a final dataset of 725 autosomal SNPs. All microsatellites and mitochondrial SNPs passed the 90% genotyping threshold.

Evidence of population structure and differentiation were evaluated using principal component analysis (PCA), admixture analysis, and calculations of F_{ST} (Figure 2, Figure 3, Table 1). In the PCA of only samples collected during the summer breeding season, there is not a definitive separation between the breeding herds (Figure 2). In the PCA based on 725 SNPs, the central herd seems to only contain a subset of the genetic diversity found within Yellowstone while the northern herd has a wider distribution (Figure 2a). In the PCA based on 24 microsatellites, there is more overlap between the two breeding herds sample distribution (Figure 2b). The inclusion of samples collected during the winter migration, did not substantially change the distribution of samples in the PCA for either herd (Supplemental Figure 1). In fact, there are minimal differences between the PCAs of summer samples only and all samples besides the density of sample points. Additionally, fastStructure, with 725 SNPs, and STRUCTURE, with 24 microsatellites, admixture analyses did not separate the geographically defined herds into genetically distinct subpopulations (Figure 3, Supplemental Figure 2). In the fastStructure analysis ran from K=1 to K=6 for only the samples collected during the breeding season, both the model complexity that maximizes marginal likelihood and model components used to explain the data structure were revealed to equal one, meaning the best fit for this data is one population (Figure 3a). This was the same when including winter samples (Supplemental Figure 2a). In the microsatellite-based STRUCTURE analysis, there was again no distinction between the two summer breeding herds (Figure 3b). Evaluation of multiple runs revealed the best fit as K=3 according to the delta K (Evanno et al., 2005) meaning there is evidence for three genetically defined clusters. However, this did not extend to geographically defined breeding herds. In fact,

all samples seem to have an even representation of all clusters at all instances of K. Inclusion of samples collected during the winter migration were consistent with the summer only analysis in admixture of individuals, however the best fit was K=2 instead (Supplemental Figure 2b).

Estimates of F_{ST} were used to measure genetic differentiation among each collection group using both the SNP and microsatellite datasets (Table 1). The central herd summer and winter samples had the lowest measures of differentiation in both the SNP and microsatellite comparisons, F_{ST} =-0.007 and -0.0043 respectively. In the SNP comparison the highest level of differentiation was between the central winter and northern summer samples (F_{ST} = 0.0050) which was slightly higher than the difference between the central and northern summer groups (F_{ST} = 0.0049). However, in the microsatellite comparison the highest differentiation was between the central and northern summer samples (F_{ST} = 0.0044) and the second highest was between the central summer and northern winter samples (F_{ST} = 0.0031).

To characterize the genetic diversity of each collection group (central herd- summer, central herd- winter, northern herd- summer, and northern herd- winter) estimates of mean observed heterozygosity (H_O), mean gene diversity within populations (H_S), mean F_{IS}, and mean allelic richness (A_R) were calculated across each group for 725 autosomal SNPs and 24 microsatellites (Supplemental Table 1). Estimates of genetic diversity had a larger range when calculated with the microsatellite dataset compared the SNP dataset. In the SNP dataset, the central herd had on average higher observed heterozygosity than the northern herd. Both central groups were above the overall mean (H_O= 0.4239), while both northern groups were below the overall mean. Though in the microsatellite dataset, the central summer was above the overall mean (H_O= 0.5906) and the other groups were below the overall mean with the central winter group estimate the lowest. Mean F_{IS} was not significantly different between each group and all 95% confidence level range estimates overlapped with each other and the overall mean 95% confidence interval in both datasets.

All Yellowstone bison had bison derived mitochondrial haplotypes based on ten interspecific mtDNA SNPs (Forgacs et al., 2016; Ward et al., 1999). No samples had evidence of domestic cattle mitochondrial DNA. Additionally, three SNPs distinguished the major intraspecific mitochondrial clades found in bison (Forgacs et al., 2016). All 272 samples in the final SNP dataset were assigned a mitochondrial haplotype based on the criteria outlined in Stroupe & Derr (submitted). Clade II had the highest frequency of 65% in overall assignment (Figure 4). Within each group, Clade II was also the dominant mitochondrial haplotype. However, the central herd winter samples had the highest frequency of Clade II followed by central summer, northern winter, and northern summer with frequencies of 78%, 70%, 65%, and 57% respectively.

Discussion

No evidence supported the hypothesis that the bison in Yellowstone National Park are currently comprised of genetically distinct and independently breeding subpopulations. The presented genetic analyses using both microsatellite and SNP markers did not reveal substantial differentiation between bison sampled in the northern and central ranges during the summer breeding season. Moreover, analyses showed clear support for considering the bison in

Yellowstone as one interbreeding metapopulation. However, there is undeniable evidence of multiple genetic lineages contributing to the current genetic diversity.

The sets of microsatellites and SNPs used in this study have both revealed genetic distinctions between populations in previous studies (Halbert & Derr, 2008; Stroupe & Derr submitted). Therefore, it is unlikely the lack of population subdivision is due to a lack of sensitivity in the selected genetic markers. Both previous studies revealed an observational difference in multiple measures of genetic diversity between Yellowstone National Park and other federal bison populations. The level of genetic differentiation in the Yellowstone breeding herds is comparable to bison populations from the same source that have been separated less than 8 years (F_{ST} = 0.0049 vs 0.0030) (Stroupe & Derr submitted). Furthermore, the genetic differentiation between the Yellowstone National Park breeding herds is lower than between closely related federal bison populations (Stroupe & Derr submitted).

While the results of the SNP and microsatellite analyses both revealed the lack of subdivision among Yellowstone bison, there were some observed differences. PCA and admixture analysis using SNPs, seemed to reveal a higher sensitivity of subtle differences. PCA of SNPs exposed the central herd as a subset of the whole while the microsatellite PCA showed no observable differences in the herds. Admixture analysis of the SNPs showed variable levels of admixture among samples while the microsatellite data was congruent across samples. Differences could be due to SNPs being more effective in identifying signals of population division compared to microsatellites which has been observed in other wildlife population studies (Zimmerman et al., 2020).

Due to anthropogenic movements of bison since the population bottleneck in the late 19th century, all modern bison populations are derived from multiple historic lineages, and all have evidence of domestic cattle introgression (Coder, 1975; Stroupe et al., 2022). Therefore, when engaging in tasks such as gauging population divergence and migration rates, a comprehensive understanding of a population's historical trajectory is imperative for accurate interpretation of findings.

During the population bottleneck in the late 19th century bison in the Greater Yellowstone Area consisted of a dwindling population of indigenous bison that had persisted since prehistoric times. However, when the population reached an estimated low of 23 individuals in 1902, genetically distinct animals were brought in from Montana (Pablo-Allard herd) and Texas (Goodnight herd) to form an introduced population in the northern range (Meagher, 1973). Therefore, the Yellowstone bison population, would be better described as initially two subpopulations of genetically distinct lineages that have become a metapopulation through population growth, range expansion, natural selection, and migration instead of a historic divergence then convergence of a single source population (Meagher 1989, Meagher 2002, White and Wallen 2012).

Although previous studies provided evidence of genetic subdivision in Yellowstone bison, there is no evidence this persists in the modern population (Halbert et al., 2012a). No conclusions at the time could be drawn as to whether the identified genetic population substructure extended to the breeding groups due to limited location data during the summer breeding season.

Additionally, two decades (1997-2003 versus 2019-2021) separate when the population was sampled and showed subdivision compared to the presented study which did not. It is not clear if the observed differences represent a change in population dynamics over time, management actions, sampling strategy and study design, or a combination of these factors.

Our estimates of genetic differentiation between the central and northern breeding herds were much lower (Microsatellite F_{ST} = 0.0044, SNP F_{ST} = 0.0049) compared to previous estimates (F_{ST} = 0.0321) (Halbert et al., 2012a). This difference could represent a change in population structure over time or could be due to grouping samples by genetic cluster rather than geographic distribution or a combination thereof. Moreover, previous estimated migration rate of 2 per generation is likely underestimated based on the assumptions used, separation of samples based on genetic clustering rather than location, and lack of samples collected during the breeding season.

Throughout the years, there have been many changes in conditions and management that could alter the behavior and movement patterns of bison (White et al., 2015). Yellowstone National Park is one of the only places where bison population size and behavior freely respond to environmental changes such as predators, resource limitations, and climate. Furthermore, population numbers also fluctuate due to culling bison during the winter at the parks boundaries to reduce numbers and the risk of brucellosis transmission to domestic cattle (Geremia, 2022; White et al., 2015). In some years, this management practice has resulted in the removal of over 1,000 animals from this population, with disproportionate removals from the central and northern herds (Geremia, 2022; White et al., 2011).

In addition, we were able to evaluate maternal lineages and mitochondrial haplotype distribution within the Yellowstone bison population. Similar to the findings of Forgacs et al. (Forgacs et al., 2016), mitochondrial haplotype clades I and II were distributed across the central and northern ranges in both the summer (breeding) and winter. These two mitochondrial haplotype clades likely reflect the mixed lineage history of indigenous and introduced animals that comprise the current Yellowstone National Park bison herd. Overall mitochondrial haplotypes from Clade II were found in higher frequency that Clade I haplotypes (65% for Clade II and 35% for Clade I) and Clade II was also found in higher frequency in both the Central and Northern populations in summer and winter collections (Figure 4)

This reevaluation of the Yellowstone bison population to delineate the current population structure was able to improve upon previous studies by evaluating the two major summer breeding populations with previously used methods and new SNP-based technologies. The SNP panel has proven effective in population differentiation among closely related bison populations and estimating measures of genetic diversity (Stroupe & Derr submitted). The SNP-based evaluation of this important population provides a point-in-time observation of the Yellowstone bison. With the increased efficiency of SNP-based platforms due to large number of loci, consistency, easier automation (Anderson & Garza, 2006), lower mutation rates (Amorim & Pereira, 2005), reduced influence of inbreeding (Fernández et al., 2013), and ease of data sharing (Forcina & Leonard, 2020) there is more opportunity for utilization in long-term management. Moreover, samples collected during the breeding season are imperative for establishing the current structure since that is when genetic exchange between populations occurs and migration

patterns can overlap in the winter ranges. Though the addition of samples collected in the winter in our analyses showed similar results.

Here we present evidence, developed from multiple analyses, that indicate bison within Yellowstone National Park represent a single interbreeding population. Even though there are multiple breeding units and clear evidence of historical bison lineages, it appears substantial gene flow is occurring throughout the herd. Thus, there is no way to confidently assign individuals outside of the summer breeding season to their respective breeding unit without tracking individual movement histories. The bison at Yellowstone National Park are an important biological resource that is essential to the long-term conservation of this species. Continued genetic monitoring is imperative to track genetic diversity indices of both nuclear and mitochondrial DNA in order to maintain stewardship of this iconic bison resource.

Funding

This work was supported by the Department of Interior, National Park Service grant M1802772. SS support was provided by a Throlson American Bison Foundation Scholarship, The Houston Safari Club, Jim Womack Endowed Fund in Animal Genomics, and the School of Veterinary Medicine and Biomedical Sciences, Texas A&M University.

Acknowledgements

Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing.

References

American Bison Society. (1908). *Annual Report of the American Bison Society 1905-1907*.

Amorim, A., & Pereira, L. (2005). Pros and cons in the use of SNPs in forensic kinship investigation: A comparative analysis with STRs. *Forensic Science International*, *150*(1), 17–21. https://doi.org/10.1016/j.forsciint.2004.06.018

Anderson, E. C., & Garza, J. C. (2006). The power of single-nucleotide polymorphisms for large-scale parentage inference. *Genetics*, *172*(4), 2567–2582. https://doi.org/10.1534/genetics.105.048074

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 1–16. https://doi.org/10.1186/s13742-015-0047-8

Coder, G. D. (1975). The national movement to preserve the American buffalo in the United States and Canada between 1880 and 1920. In *ProQuest Dissertations and Theses*. http://proxy.library.tamu.edu/login?url=https://www.proquest.com/dissertations-theses/national-movement-preserve-american-buffalo/docview/302745670/se-2?accountid=7082

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Group, 1000 Genomes Project Analysis. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dray, S., & Dufour, A.-B. (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*, 1–20.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fernández, M. E., Goszczynski, D. E., Lirón, J. P., Villegas-Castagnasso, E. E., Carino, M. H., Ripoli, M. v., Rogberg-Muñoz, A., Posik, D. M., Peral-García, P., & Giovambattista, G. (2013). Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. *Genetics and Molecular Biology*, *36*(2), 185–191. https://doi.org/10.1590/S1415-47572013000200008

Forcina, G., & Leonard, J. A. (2020). *Tools for Monitoring Genetic Diversity in Mammals: Past, Present, and Future BT - Conservation Genetics in Mammals: Integrative Research Using Novel Approaches* (J. Ortega & J. E. Maldonado, Eds.; pp. 13–27). Springer International Publishing. https://doi.org/10.1007/978-3-030-33334-8_2

Forgacs, D., Wallen, R. L., Dobson, L. K., & Derr, J. N. (2016). Mitochondrial Genome Analysis Reveals Historical Lineages in Yellowstone Bison. *PLOS ONE*, *11*(11), e0166081. https://doi.org/10.1371/journal.pone.0166081

Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*(1), 27–32.

Geremia, C. (2022). *Status Report on the Yellowstone Bison Population to the Superintendent Bison Program*.

Geremia, C., White, P. J., Hoeting, J. A., Wallen, R. L., Watson, F. G. R., Blanton, D., & Hobbs, N. T. (2014). Integrating population- and individual-level information in a movement model of Yellowstone bison. *Ecological Applications*, *24*(2), 346–362. https://doi.org/https://doi.org/10.1890/13-0137.1

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*(1), 184–186.

Halbert, N. D. (2003). *The utilization of genetic markers to resolve modern management issues in historic bison populations: implications for species conservation*. Texas A&M University.

Halbert, N. D., & Derr, J. N. (2008). Patterns of genetic variation in US federal bison herds. *Molecular Ecology*, *17*(23), 4963–4977. https://doi.org/10.1111/j.1365-294X.2008.03973.x

Halbert, N. D., Gogan, P. J. P., Hedrick, P. W., Wahl, J. M., & Derr, J. N. (2012a). Genetic Population Substructure in Bison at Yellowstone National Park. *Journal of Heredity*, *103*(3), 360–370. https://doi.org/10.1093/jhered/esr140

Halbert, N. D., Gogan, P. J. P., Hedrick, P. W., Wahl, J. M., & Derr, J. N. (2012b). Yellowstone Bison Genetics: Let Us Move Forward. *Journal of Heredity*, *103*(5), 754–755. https://doi.org/10.1093/jhered/ess051

Huisman, J. (2017). Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources*, *17*(5), 1009–1024.

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, *16*(5), 1099–1106.

Meagher, M. M. (1973). *The bison of yellowstone national park* (Issue 1). US Government Printing Office.

Meagher, M. M. (1989). Range Expansion by Bison of Yellowstone National Park. *Journal of Mammalogy*, *70*(3), 670–675. https://doi.org/10.2307/1381449

Meagher, M. M. (1993). *Winter recreation-induced changes in bison numbers and distribution in Yellowstone National Park*. Mary Meagher.

Meagher, M. M. (1998). Recent changes in Yellowstone bison numbers and distribution. *International Symposium on Bison Ecology and Management (L. Irby and J. Knight, Eds.). Montana State University Press, Bozeman, Montana*, 107–112.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/https://doi.org/10.1086/519795

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Schnabel, R. D., Ward, T. J., & Derr, J. N. (2000). Validation of 15 microsatellites for parentage testing in North American bison, Bison bison and domestic cattle. *Animal Genetics*, *31*(6), 360–366. https://doi.org/10.1046/j.1365-2052.2000.00685.x

Skinner, C. K., & Alcorn, W. B. (1942). History of the bison in Yellowstone National Park. *Project Report to the Chief Ranger. Mammoth Hot Springs, Yellowstone National Park, Wyoming (Updated in 1952 by WL Evans)*.

Stroupe, S., Forgacs, D., Harris, A., Derr, J. N., & Davis, B. W. (2022). Genomic evaluation of hybridization in historic and modern North American Bison (Bison bison). *Scientific Reports*, *12*(1), 1–11. https://doi.org/10.1038/s41598-022-09828-z

Toll, R. W. (1929). *Annual Report for Yellowstone National Park*.

Toonen, R. J., & Hughes, S. (2001). Increased throughput for fragment analysis on an ABI Prism®377 automated sequencer using a membrane comb and STRand software. *Biotechniques*, *31*(6), 1320–1325.

Ward, T. J., Bielawski, J. P., Davis, S. K., Templeton, J. W., & Derr, J. N. (1999). Identification of domestic cattle hybrids in wild cattle and bison species: a general approach using mtDNA markers and the parametric bootstrap. *Animal Conservation*, *2*(1), 51–57. https://doi.org/https://doi.org/10.1111/j.1469-1795.1999.tb00048.x

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358–1370.

White, P. J., & Wallen, R. L. (2012). Yellowstone Bison—Should We Preserve Artificial Population Substructure or Rely on Ecological Processes? *Journal of Heredity*, *103*(5), 751–753. https://doi.org/10.1093/jhered/ess050

White, P. J., Wallen, R. L., Geremia, C., Treanor, J. J., & Blanton, D. W. (2011). Management of Yellowstone bison and brucellosis transmission risk--Implications for conservation and restoration. *Biological Conservation*, *144*(5), 1322–1334.

White, P. J., Wallen, R. L., & Hallac, D. E. (2015). *Yellowstone Bison: Conserving an American Icon in Modern Society*. Yellowstone Association.

Wickham, H. (2016). Programming with ggplot2. In *ggplot2: Elegant Graphics for Data Analysis* (pp. 241–253). Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4_12

Zimmerman, S. J., Aldridge, C. L., & Oyler-McCance, S. J. (2020). An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC Genomics*, *21*, 1–16.