Genomics in Conservation: Case Studies and Bridging the Gap between Data and Application

Brittany A. Garner,[1,2,‡], Brian K. Hand,[1,‡,*] Stephen J. Amish,[1,2] Louis Bernatchez,[3] Jeffrey T. Foster,[4] Kristina M. Miller,[5] Phillip A. Morin,[6] Shawn R. Narum,[7] Stephen J. O'Brien,[8] Gretchen Roffler,[9] William D. Templin,[10] Paul Sunnucks,[11] Jeffrey Strait,[1,2] Kenneth I. Warheit,[12] Todd R. Seamons,[12] John Wenburg,[13] Jeffrey Olsen,[13] and Gordon Luikart[1,2]

[1]Flathead Lake Biological Station, Fish and Wildlife Genomic Group, Division of Biological Sciences, University of Montana, Polson, MT 59860, USA

[2]Wildlife Program, Fish and Wildlife Genomic Group, College of Forestry and Conservation, University of Montana, Missoula, MT 59812, USA

[3]Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC G1V 0A6, Canada

[4]Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, NH 03824, USA

[5]Molecular Genetics Laboratory, Pacific Biological Station, Nanaimo, BC V9T 6N7, Canada

[6]Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 8901 La Jolla Shores Drive, La Jolla, CA 92037, USA

[7]Columbia River Inter-Tribal Fish Commission, Hagerman Fish Culture Experiment Station, 3059-F National Fish Hatchery Road, Hagerman, ID 83332, USA

[8] Theodosius Dobzhansky Center for Genome Bioinformatics, St Petersburg State University, 41 Sredniy Prospect, St Petersburg, Russia

[9]Alaska Department of Fish and Game, Division of Wildlife Conservation, 802 3rd Street, Douglas, AK 99824, USA

[‡]These authors contributed equally.

[10]Gene Conservation Laboratory, Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, AK 99518, USA

[11]School of Biological Sciences, Monash University, Melbourne, VIC 3800 Australia

[12]Washington Department of Fish and Wildlife, Molecular Genetics Laboratory, 600 Capitol Way N., Olympia, WA 98501, USA

[13]Conservation Genetics Laboratory, 1011 East Tudor Road, MS 331, Anchorage, AK 99503, USA

*Correspondence: brian.hand@umontana.edu (Hand, B.K.).

Keywords: conservation genomics; population genomics; conservation biology; next-generation sequencing; conservation practice; genetic monitoring; natural resource management.

We agree with Shafer *et al.* [1] that there is a need for well-documented case studies of the application of genomics in conservation and management as well as increased communication between academics and natural resource managers. However, we challenge Shafer *et al.*'s [1] relatively pessimistic assertion that 'conservation genomics is far from seeing regular application'. Here we illustrate by examples that conservation practitioners utilize more genomic research than is often apparent. In addition, we highlight the work of nonacademic laboratories [government and nongovernmental organizations (NGOs)], some of which are not always well represented in peer-reviewed literature. Finally, we suggest that increased agency–academic collaboration would enhance the application of genomics to real-world conservation and help conserve biodiversity.

There is substantial controversy and confusion surrounding the definition of 'genomics' versus traditional genetic approaches. Here we address this by expanding Shafer *et al.*'s [1] definition to include a broad- and narrow-sense definition to better illuminate the different ways that genomics contributes to conservation practice. We define broad-sense conservation genomics as the use of new genomic techniques and genome-wide information to solve problems in conservation biology (as in Shafer *et al.* [1] and Allendorf *et al.* [2]). Our narrow-sense definition also requires the use of approaches that are conceptually and quantitatively different from traditional genetics that would be impossible using genetic data alone (e.g., detecting genome-wide adaptation, use of transcriptomics, epigenetics, using annotated genomes). This narrow-sense definition includes using hundreds to thousands of mapped or gene-targeted marker loci in combination with recent computational and conceptual approaches such as mapping runs of homozygosity, comparing neutral versus adaptive patterns of population structure or gene flow,

testing for signals of selection to assess adaptation, and testing assumptions of neutrality (e.g., before estimating effective population size or gene flow patterns).

Narrow-sense genomic approaches have been used for diverse conservation applications including identifying conservation units, assessing gene flow, and detecting local adaptation (Table S1 in the supplementary material online). We agree with Shafer *et al.* [1] and others [2] about the general and serious concern of erroneous identification of adaptive loci and their subsequent use (or misuse) in conservation practice. However, we remain cautiously optimistic given the recent efforts to use putatively adaptive loci to inform management practices. For instance, genome-wide scans using diversity array technology (DArTseq) in gimlet trees (*Eucalyptus salubris*) generated 16 122 neutral and putatively adaptive SNP markers used to uncover distinctive molecular lineages signaling adaptation to different environments. These genome-wide scans offered enhanced precision otherwise unavailable with traditional genetics or phenotypic traits alone [3] (Table S1). Such novel insights are important in seed choice for the ecological restoration of gimlet trees, a keystone species in the Great Western Woodlands of Australia, in the wake of wildfires [3].

In many broad-sense studies, next-generation sequencing (NGS) has enabled the discovery of management-informative markers that are subsequently screened in populations of conservation concern. For example, state management agencies in Washington and Idaho, USA used NGS to discover markers of introgression from hatchery broodstock into wild populations of salmonid fishes [4,5]. Other applications of broad-sense conservation genomics are evident (Table S1) and have been enabled by recent NGS and SNP genotyping technologies [6] (http://biorxiv.org/content/early/2015/10/11/028837). These approaches allow genome-wide

discovery and genotyping of highly informative markers, making cost-effective monitoring feasible using relatively small marker sets (e.g., 100–500 markers) [7].

Decreases in costs (e.g., sequencing, library prep, bioinformatics) are sparking the application of NGS to a broader set of conservation questions and taxa where funding is relatively more limited. In addition to the examples above, genomic data are currently applied in conducting parentage analyses in Pacific lampreys (*Lampetra tridentata*) and monitoring for disease in Tasmanian devils (*Sarcophilus harrisii*) [8,9] and fish (Table S1). Power analyses and cost-savings comparisons of using SNPs versus microsatellite markers in conservation genomics would be of great benefit, but such analysis is beyond the scope of this letter. However, using genomic approaches has been shown to provide more statistical power than microsatellites and to cost as low as 1% of traditional Sanger sequencing prices [6,7,10] (Table S1).

We have included multiple case studies from salmonids because these species are of great conservation concern due to their ecological, commercial, and cultural importance in many Northern Pacific Rim river systems. For example, ~30% or more of salmonid populations in the Columbia River Basin (USA–Canada) have been extirpated and many remaining populations are listed as endangered or threatened under the Endangered Species Act (ESA) or the Species at Risk Act in Canada because of, for example, over-harvesting, habitat degradation, pollution, and hydrological dams [11]. Therefore, more money and time is being spent on these species than other taxa due to their multiple conservation concerns (e.g., climate change, hybridization, over-harvesting). There are ~12 nonacademic laboratories (e.g., federal, tribal, NGO, state agencies) using genomic data to work mostly or exclusively on salmonids in the Pacific Northwest of North America. Shafer *et al.* [1] insufficiently acknowledged one of the most significant contributions of genomics to conservation by not fully highlighting the work of these

laboratories, particularly the Alaska Department of Fish and Game (ADFG), a leader in SNP and NGS tool development and application. ADFG genotypes approximately 100 000 fish annually for management using broad-sense conservation genomic approaches, showing that conservation genomics in salmonids is an example of what is possible when adequate resources are devoted to the issue [12] (Table S1).

We highlight recent applications of genomics in real-world management where some are published, but many similar studies are not published or widely disseminated. Some nonacademic laboratories have relatively limited incentive to publish or are delayed due to urgent deadlines reinforced by political, legislative, or legal constraints. For example, some agency laboratories produce reports or declarations used in litigation or the planning of harvest regulations or introductions (e.g., hatchery fish management plans), which can delay scientific publication. Nonacademics could potentially publish more by collaborating with academic groups who have strong incentives to publish (e.g., to 'publish or perish'). Academics could in turn achieve greater conservation impact by working closely with practitioners who can provide benefits such as large sample and data collections, funding and field staff, collection permits, and high-throughput genomics platforms.

While research and publications from some nonacademic laboratories are often underappreciated or delayed, they can help the conservation biology community to understand the extent and feasibility of applying genomics to conservation. We hope by highlighting case studies we will expand discussions and applications of genomic techniques in conservation and encourage the closing of gaps between nonacademic laboratories and academia.

Acknowledgments

The authors thank the following people for providing comments, feedback, and case studies: Margaret Byrne, Todd Cross, Taylor Wilcox, Robb Leary, Sally Aitken, Jon Ballou, Bob Lacy, Eric Peatman, Luciano Beheregaray, Katherine Ralls, Brett Addis, and Kathy Belov. They thank James and Lisa Seeb for extensive contributions to earlier versions of the manuscript. They also thank Fred Allendorf for help in developing the initial idea of a narrow-sense definition and subsequent advice on defining narrow sense (e.g., in Table S1). They especially thank Aaron Shafer, Michael Bruford, and Michael Schwartz for encouraging publication and improving earlier versions the manuscript with a shared vision of enhancing the contribution of genomics in conservation. G.L. and B.K.H. were partially supported by grants from the NSF (DEB 1258203), NASA (NNX14AB84G), and Montana Fish Wildlife and Parks.

References

1 Shafer, A.B.A. *et al.* (2015) Genomics and the challenging translation into conservation practice. *Trends Ecol. Evol.* 30, 78–87

2 Allendorf, F.W. *et al.* (2010) Genomics and the future of conservation genetics. *Nat. Rev. Genet.* 11, 697–709

3 Steane, D.A. *et al.* (2015) Genome-wide scans reveal cryptic population structure in a dry-adapted eucalypt. *Tree Genet. Genomes* 11, 33

4 Warheit, K.I. (2014) *Measuring Reproductive Interaction between Hatchery-Origin and Wild Steelhead (*Oncorhynchus mykiss*) from Northern Puget Sound Populations Potentially Affected by Segregated Hatchery Programs*, Washington Department of Fish and Wildlife

5 Steele, C.A. *et al*. (2013) A validation of parentage-based tagging using hatchery steelhead in the Snake River basin. *Can. J. Fish Aquat. Sci.* 70, 1046–1054

6 Narum, S.R. *et al.* (2013) Genotyping-by-sequencing in ecological and conservation genomics. *Mol. Ecol.* 22, 2841–2847

7 Campbell, N.R. *et al.* (2014) Genotyping-in-thousands by sequencing (GT-seq): a cost effective SNP genotyping method based on custom amplicon sequencing. *Mol. Ecol. Resour.* 15, 855–867

8 Hess, J.E. *et al.* (2015) Use of genotyping by sequencing data to develop a high-throughput and multifunctional SNP panel for conservation applications in Pacific lamprey. *Mol. Ecol. Resour.* 15, 187–202

9 Miller, W. *et al.* (2011) Genetic diversity and population structure of the endangered marsupial *Sarcophilus harrisii* (Tasmanian devil). *Proc. Natl Acad. Sci. U. S. A.* 108, 12348–12353

10 Lemmon, A.R. *et al.* (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst. Biol.* 61, 727–744

11 Gustafson, R. *et al*. (2007) *Pacific Salmon Extinctions: Quantifying Lost and Remaining Diversity. Paper 438*, US Department of Commerce

12 Habicht, C. *et al.* (2012) *Harvest and Harvest Rates of Sockeye Salmon Stocks in Fisheries of the Western Alaska Salmon Stock Identification Program (WASSIP), 2006–2008. Special Publication No. 12-24*, Alaska Department of Fish and Game