Title: Training the next generation of conservation genomicists: ConGen 2018 Workshop

Population genomics training for the next generation of conservation geneticists: ConGen 2018 Workshop

Authors: Amanda Stahlke[1], Donavan Bell[2], Tashi Dhendup[2], Brooke Kern[3], Sam Pannoni[2,4], Zachary Robinson[2], Jeffrey Strait[2], Seth Smith[2,4,5], Brian K. Hand[2,4], Paul A. Hohenlohe[1], Gordon Luikart[2,3,4]

Email Addresses (in order): astahlke@uidaho.edu; donovan.bell@umconnect.umt.edu; tashi.dhendup@umconnect.umt.edu; brooke.kern@umconnect.umt.edu; sam.pannoni@umconnect.umt.edu; zachary.robinson@umconnect.umt.edu; jeffrey.strait@umontana.edu; smithsr.90@gmail.com; brian.hand@flbs.umt.edu; hohenlohe@uidaho.edu; gordon.luikart@umontana.edu.

[1]Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844
[2] Wildlife Biology Program, College of Forestry and Conservation, University of Montana, Missoula, MT 59812
[3] Division of Biological Sciences, College of Humanities and Sciences, University of Montana, Missoula, MT 59812
[4]Flathead Lake Biological Station, Division of Biological Sciences, College of Humanities and Sciences, University of Montana, 59860.
[5]Department of Fisheries & Wildlife, Michigan State University, East Lansing, MI 48824

Abstract:

The increasing availability and complexity of next generation sequencing (NGS) datasets make ongoing training an essential component of conservation and population genetics research. A workshop entitled 'ConGen 2018' was recently held to train researchers in conceptual and practical aspects of NGS data production and analysis for conservation and ecological applications. Sixteen instructors provided helpful lectures, discussions, and hands-on exercises regarding how to plan, produce, and analyze data for many important research questions. Lecture topics ranged from understanding probabilistic (e.g. Bayesian) genotype calling to the detection of local adaptation signatures from genomic, transcriptomic, and epigenomic data. We report on progress in addressing central questions of conservation genomics, advances in NGS data analysis, the potential for genomic tools to assess adaptive capacity, and strategies for training the next generation of conservation genomicists.

Keywords: Population genomic data analysis, conservation genetics pedagogy, effective population size, evolutionary significant units, adaptive capacity

Introduction

Informing conservation efforts is one of the most important and challenging needs of the genomic era (Allendorf 2017; Hunter et al. 2018). To help meet this challenge, sixteen experts from many areas of genomic data analysis met to discuss and teach recent analytical approaches at the 10th International Population Genetics Data Analysis Workshop for Conservation ("ConGen"), held at Flathead Biological Station in September of 2018. The goal of the

23 workshop was to train participants to apply rigorous theory and novel molecular and
24 computational approaches in conservation and population genetics.
25 Since the first ConGen in 2006 (https://cibio.up.pt/congen/index.html), the molecular and
26 computational tools accessible to conservation have grown in number and matured (Andrews and
27 Luikart 2014; Hendricks, Anderson, Antao, Bernatchez, Forester, Garner, Hand, Hohenlohe,
28 Kardos, Koop, et al. 2018; Benestan et al. 2015). ConGen 2018 students originated from >15
29 countries and had a wide range of research questions and career stages including: undergraduate
30 and graduate (Masters and PhD) students, postdoctoral scholars, university faculty, laboratory
31 technicians, and governmental agency scientists. This diversity of origins and perspectives
32 enriched the questions, comments, discussions, and overall learning experience.
33 Historically, ConGen and other conservation genetics courses have focused mainly on
34 approaches (and questions that require) using ~10-15 well-assessed markers (e.g.,
35 microsatellites) such as hybridization, inbreeding, population structure and loss of genetic
36 diversity (Allendorf 2017). Today, the variety of molecular tools, amount of genetic data, and
37 range of computational approaches has greatly expanded. Conservation genomics can be broadly
38 defined as the application of genome-wide markers and new technologies to address problems in
39 conservation. A more narrow-sense definition requires high-density loci to characterize locus- or
40 gene-specific patterns and thus address conceptually novel questions that were intractable using
41 traditional approaches (Allendorf et al. 2010; Garner et al. 2016; Allendorf 2017; Luikart et al.
42 2018).
43 Throughout this genetics-to-genomics transition, many authors, including those of
44 previous ConGen reviews, have reflected on this paradigm shift. They have noted the best
45 practices for filtering, experimental design, analytical approaches, career choices, and the

increasing roles of women (Andrews and Luikart 2014; Benestan et al. 2015; Shaffer et al. 2017; Hendricks et al. 2018; REFs). In this meeting review of ConGen 2018, we focus our reflection on training the next generation of researchers in conservation genomics through the novel components of this years' workshop: progress in understanding key concepts including assessing population differentiation and conservation units, estimation of effective population size, molecular data production and analysis for diverse empirical systems, and prospects for understanding genomic vulnerability.

Progress in Central Concepts

Populations, ESUs, and CUs: How do you identify them using genomics?

Defining biologically meaningful management units within species is challenging (Waples and Gaggiotti 2006; Waples and Lindley 2018; Bradshaw et al. 2018). For the purpose of conservation, the Evolutionary Significant Unit (ESU) is a distinct population or group of populations that can be protected under the U.S. Endangered Species Act (ESA; REF). In Robin Waples' (Northwest Fisheries Science Center) lecture on ESUs, he explained that while there is no single or universal definition of a population, the competing definitions of ESUs emphasize two criteria: 1) substantial reproductive isolation, and 2) an important component of the evolutionary legacy of the species (Waples 1991) (Waples and Gaggiotti 2006). Evolutionary legacy refers to having distinct or different adaptations likely important for species persistence. Molecular genetic data have long been used to assess the isolation criterion for identifying ESUs, but prior to the age of genomics, the evolutionary significance of a population was difficult to determine and was largely inferred by ecological observations.

With genomic data we can now identify loci, alleles, and surrounding regions associated with adaptive differentiation which improves our capacity to define ESUs while taking into account both demographic and selective processes (Funk et al. 2012; Funk et al. 2018). Incorporating adaptive variation into ESU listing raises theoretical and practical challenges (Funk et al. 2018). Mike Miller's ConGen 2018 lecture on an early-migration phenotype in salmonids demonstrated this challenge, wherein previous studies found little evidence for genetic isolation, but locus-specific analysis and simulation modeling provided strong evidence for this phenotype as an important component of the species' evolutionary legacy (Box 1).

Effective Population Size and Effective Number of Breeders (N_e and N_b)

Effective population size (N_e) is one of the most important concepts and parameters in conservation and evolutionary genetics because it influences the rate of loss of genetic variation, the levels of individual inbreeding, and the effectiveness of natural selection and gene flow (Wang, Santiago, and Caballero 2016). Conservation genetics has long employed estimates of effective population size to help assess and monitor the vulnerability of a population to potentially harmful genetic changes as mentioned above.

While genomic data provide greater resolution and ability to estimate N_e in a growing diversity of species and scenarios, these data can also present unique challenges in estimating N_e. In his lecture on N_e, Waples discussed the recent advances in theory and computational analysis which have vastly improved N_e estimation in the genomic era (Waples and Do 2008; Waples, Antao, and Luikart 2014; Hollenbeck, Portnoy, and Gold 2016; Waples, Scribner, et al. 2018; Waples, Grewe, et al. 2018; Zhou et al. 2018). The use of thousands of loci, many of which are likely physically linked, downwardly biases N_e estimates unless physical location (linkage) is

taken into account (Do et al. 2014b; Waples, Larson, and Waples 2016). The recently improved LDNe method implemented in the NeEstimator program (as of version 2.1) improves reliability of confidence intervals and reduces bias in estimating N_e by calculating r^2 on locus pairs, employing positional information from assembled loci or, when available, chromosomes (Do et al. 2014a). Likewise, the improved capability of NeEstimator to handle missing data, which calculates a fixed inverse variance-weighted harmonic mean at each locus (Peel et al. 2013), has been shown to be accurate with up to 50% missing data (Nunziata and Weisrock 2018). Together, these methodological improvements make estimating effective population size more accessible to studies with reduced representation data, with or without a reference genome.

Waples and Andrew Whiteley (University of Montana) highlighted N_b, or the number of effective breeders in a cohort, as a promising parameter for genetic and population management because of its intrinsic relationship to N_e and potential relationship with population abundance or environmental conditions (Kamath et al. 2015; Whiteley et al. 2015). An advantage of estimating N_b, rather than N_e, is that N_b provides frequent (e.g. yearly) information on population status, rather than having to wait to sample between generations which is often required by temporal estimations of N_e (e.g. (Waples and Yokota 2007; Waples, Antao, and Luikart 2014). In his lecture, Whiteley emphasized monitoring population cohorts using a single sample and sib-ship or linkage disequilibrium methods (Kamath et al. 2015; Waples, Scribner, et al. 2018) and demonstrated the nuances of estimating N_b through recent studies of brook trout (*Salvelinus fontinalis*). He cautioned that while estimates of N_b can track abundance in some species (Ferchaud and Hansen 2016), which may supplement or allow demographic-based monitoring, it is unlikely to closely track abundance for species with high variance in reproductive success and limited reproductive habitat. For example, for brook trout that spawn in available habitat patches

N_b estimates had no association with yearly abundance in two populations; however, they provided important information about environmental conditions (Whiteley *et al.* 2015). A comparison among several brook trout populations showed that N_b was the largest at intermediate flow conditions, which is consistent with biological hypotheses (Whiteley *et al.* 2017).

The theory and application of N_b was presented mainly in the context of aquatic organisms. Nonetheless, N_b is easier to estimate than N_e for most taxa (beyond aquatic organisms), requiring only a single sample per generation (REFs). Whiteley's example demonstrated the importance of incorporating detailed biological information in the study design, analysis, and interpretation of effective population size estimates and its relationship to census size (REF). Simulations, such as those conducted by Con Gen 2018 students with EasyPop (Balloux 2001), and those implemented in tools such as AGENE, (REF), AgeStrucNb (Antao et al. in review), NeoGen (REF) and Neff (REF), should be employed to determine an appropriate sampling scheme, implement sensitivity analysis and corroborate empirical results (REF).

Molecular genomic data generation and analysis

Training the next generation of conservation genomicists includes empowering students to evaluate and incorporate a wealth of diverse molecular genetic methods. The first ConGen meetings in 2006-2009 were focused on microsatellites. Since 2010, genomic techniques like restriction-site associated DNA-sequencing (RADseq) have increasingly been the main focus (Andrews and Luikart 2014). Of 33 students at ConGen 2018, 27 had RADseq data, four had exon capture data, and five students had whole-genome sequencing (WGS) data. Several students reported having multiple types of molecular data.

At ConGen 2018, methods both currently applied widely, and those only recently employed in conservation genomics were discussed. Paul Hohenlohe (University of Idaho) reviewed the many variations and utility of RADseq (Andrews et al. 2016), Stefan Prost (Senckenberg Museum) presented a guide to *de novo* genome assembly (Hendricks, Anderson, Antao, Bernatchez, Forester, Garner, Hand, Hohenlohe, Kardos, and Koop 2018; Fuentes−Pardo and Ruzzante 2017), and Rena Schweizer (University of Montana) highlighted the practical and conceptual considerations regarding exon capture (Bi et al. 2012; Schweizer et al. 2016). Here, we highlight advances in RAD-capture, transcriptomics, and epigenomics.

Rapture: A hybrid reduced representation approach

Lectures by Hohenlohe and Seth Smith (University of Montana) demonstrated the utility of Rapture (RAD-capture; (Ali et al. 2016), a reduced representation technique that combines an improved RADseq library preparation protocol (informally referred to as bestRAD) with an in-solution sequence probe capture to enrich sequencing libraries for a subset of RADseq loci (e.g. polymorphic loci, loci near genes, and/or loci with high heterozygosity or high F_{ST}). The major improvements prescribed by the bestRAD protocol are the ability to reduce the proportion of PCR duplicates, efficiency in using smaller starting quantities of DNA, and efficiency in scaling from hundreds to thousands of samples (Ali et al. 2016). We encourage interested readers to see Meek and Larsen (2019) for a detailed review of sequence capture techniques and their utility in conservation. Here we focus on the details each individual researcher must weigh in respect to each individual project: cost, PCR duplication rate, and computational approaches.

Because individual (indexed) samples are pooled early in the bestRAD protocol, the cost of the library preparation kit and capture reaction scales well for large sample sizes. For instance,

159 up to 96 uniquely indexed individual samples are pooled prior to adding sequencing adapters and
160 amplifying the library using a commercially available library. Seth Smith estimated that
161 bestRAD libraries can be generated for <$5.00 per individual after the cost of bestRAD adapters
162 is amortized. The per sample cost for the hybridization capture reaction is ~$0.50 assuming the
163 above multiplexing scheme and a bait panel of up to 20,000 loci. This cost could vary
164 substantially depending on the vendor used for supplies (e.g. the capture array) and does not
165 include labor for the data production which is often the majority of the cost. The cost of
166 sequencing depends on desired coverage. The number of samples that can be multiplexed per
167 sequencing lane is a function of the number of targeted loci, the PCR duplication rate, and the
168 proportion of reads that do not that align to targeted loci. He cautioned that the PCR duplication
169 rate and proportion of off-target reads are expected to vary depending on the proportion of RAD
170 loci targeted for capture and the total number of loci in the original RAD library which can be
171 influenced by sample quality, and PCR duplicate rates may be greater than the typical 20-30%
172 (e.g. Margres et al. 2018).
173 Following sequencing, Rapture data can be analyzed with any method applicable to
174 RAD-type data (Andrews et al. 2016). Among these, Stacks (Catchen et al. 2013) is commonly
175 used for population genomics with RADseq and has been covered at ConGen since 2011. At
176 ConGen 2018, Amanda Stahlke (University of Idaho) taught *de novo* and reference-based locus
177 assembly and genotyping in Stacks version 2.3, which has several major changes from the
178 original implementation (Rochette, Rivera-Colón, and Catchen 2019). Students examined the
179 impacts of removing PCR duplicates and aligning to a reference or not in F-statistics. These
180 choices depend on genetic and financial resources available and the study question, though
181 useful sensitivity frameworks have been described elsewhere (REFs). For example, low-

coverage sequencing can be a cost-effective and powerful approach (Maruki REF),but is also the most sensitive to the effects of PCR duplicates (REF).

As one of the most widely-used software pipelines for genotyping RADseq data and population genomic analysis, the Stacks program (Catchen et al. 2013) has been discussed and used at the ConGen course for several years. Here we highlight some of the recent updates to Stacks (version 2.4 at the time of writing) taught at the 2018 course. These changes are not yet peer-reviewed, but see Rochette et al (2019) and examine the change log on the website for more detail. For users with bestRAD data (Ali et al 2016) the addition of the *--bestrad* flag to process_radtags re-orients paired fastq files such that bestRAD indexes and the remainder of restriction cut-sites are always located at the beginning of the first read, eliminating the requirement of an external script to re-orient the reads prior to input. In Stacks 2, users also have the ability to input paired-end reads and assemble local RAD contigs with data produced by protocols with a randomly sheared end (e.g. Ali et al 2016) or random oligos in ddRAD (double-digest RADseq; Schweyen, Rozenberg, and Leese 2014). Instead of concatenating forward and reverse reads as previously recommended (Rochette and Catchen 2017), paired-end reads are incorporated through the new tsv2bam and gstacks, the new genotyping module, modules, yielding major improvements in memory usage and genotype-calling frameworks (Rochette et al 2019).

Novel genotype-calling algorithms have also been implemented in gstacks, including the diploid Maruki and Lynch (2017) maximum likelihood genotyping model which can incorporate population-level genotype frequencies (the “low-coverage model”) and error-rates with Bayes’ theorem. In gstacks, users may increase *--alpha* to require a greater statistical threshold for calling genotypes, instead of setting a redundant minimum stacks depth flag in the populations

module (-*m* is deprecated). These advances in Stacks hold promise to advance RADseq analysis in conservation genomics by yielding more accurate genotypes, and longer haplotypes (Rochette et al 2019).

Transcriptomics and epigenomics

Transcriptomics and epigenomics, the high-throughput studies of transcribed products and epigenetic modifications of the genome, respectively, can be used to disentangle mechanisms of local adaptation (e.g. plasticity versus Darwinian adaptation) across physiological and temporal scales (Hendricks, Anderson, Antao, Bernatchez, Forester, Garner, Hand, Hohenlohe, Kardos, and Koop 2018; Kelly 2019), though the application of understanding these mechanisms in conservation is still developing Christie et al. 2016; Le Luyer et al. 2017). Recent technological advances in library preparation which better accommodate degraded and low input DNA have made transcriptomic analysis more accessible to systems of conservation concern (Wang, Gerstein, and Snyder 2009; Schuierer et al. 2017). RNAseq has been used to identify the molecular basis for resilience to changing environment in corals (Pratlong et al. 2015; Bay et al. 2017; Barshis et al. 2013) and redband rainbow trout (*Oncorhynchus mykiss gardieri*; (Garvin, Thorgaard, and Narum 2015; Chen et al. 2018).

Still, there are surprisingly few studies that employ these techniques to inform conservation. Perhaps this is due to fewer labs having the capacity to produce and analyze these potentially tissue- and time-specific data, the actual and perceived conflicts in evolutionary paradigms, or the ongoing discussion regarding the role of plasticity in long-term population persistence (Pennisi 2013; Charlesworth, Barton, and Charlesworth 2017; Kelly 2019).

Regardless, transgenerational gene expression and epigenetic changes can underlie an adpative response to environmental change (e.g. corals).

At ConGen 2018, students gained exposure and experience to transcriptomics through an interactive lecture on data production and hands-on analysis of differential gene expression led by Joanna Kelley (Washington State University). Students learned how to functionally annotate variants of interest and perform enrichment analysis with instructor Mackenzie Gavery (University of Washington) and an epigenomic dataset. Here we highlight Gavery's lecture demonstrating the potential utility of epigenomics in conservation with a recent study of DNA methylation of cytosine residues at CpG sites induced by hatchery conditions (Gavery et al. 2018; Box 2).

Promise in Understanding Adaptive Potential and Genomic vulnerability

Genomic methods now allow researchers to determine the genetic basis for variation in fitness, quantify adaptive capacity, and predict potential outcomes for natural populations facing environmental change (Funk et al. 2018). Adaptive potential can be defined as the capacity of species or populations to respond to stressors (e.g. environmental change) by genetically-based changes (Nicotra et al. 2015; Funk et al 2018). Rachael Bay (University of California Davis) and Christen Bossou (Colorado State University) demonstrated the exciting potential for *genomic vulnerability,* which is an estimate of the extent to which allele frequencies of wild populations must change to maintain current genotype-environment associations in the future (Fitzpatrick and Keller 2015; Box 3).

The next generation: Developing theoretical, empirical, and analytical skills

249 Conservation genomics is a multidisciplinary field, requiring practitioners to have a
250 working knowledge of population genetic theory and molecular biology while developing the
251 computational skills necessary to apply novel and conventional analyses to increasingly large
252 datasets. These challenges, raised by (Allendorf, Hohenlohe, and Luikart), (Shafer et al. 2015),
253 and (Garner et al. 2016) were discussed by students and instructors alike at ConGen 2018.
254 Conservation genomicists often need to navigate social (e.g. legal), ecological, and molecular
255 dimensions, sometimes in the most challenging of field conditions (Groom, Meffe, and Carroll
256 2006).
257 Researchers must also be able to effectively communicate with stakeholders, including
258 agency managers, NGOs, policy makers, and the public (Hand et al. 2018). The diversity of
259 topics covered during lectures, discussions, and hands-on activities during ConGen 2018
260 demonstrates the importance of taking a holistic approach when tackling questions in
261 conservation genomics. One recommendation from managers at ConGen to help conservation
262 geneticists ensure their data is used for conservation management was to design a study with a
263 manager who has plans in place (e.g. including permits, policy, etc.) to use the genetic data once
264 it is available to make management decisions (Pers. Comm. M. Boyer). This recommendation is
265 an important consideration for future discussion in conservation and genetics workshops where
266 open forums and group conversations can be organized. Other big-group discussion topics
267 ranged from the best programming languages for population genomics (e.g., R, and shell
268 scripting), to career choices.
269 Theory in population genetics has a long and rich history, and yet, is still developing with
270 effective population size concepts and empirical estimation methods among the most important
271 areas (e.g., Waples et al. 2014; (Ceballos et al. 2018). The importance of theory, and specifically

effective population size, is exemplified by the following quotes: “Nothing in evolution makes sense except in light of population genetics,” (Lynch and Walsh 2007) and “Nothing in population genetics makes sense except in light of effective population size,” which Robin Waples at ConGen 2018 said was a quote from Fred Allendorf (University of Montana). For example, when testing for genotype-phenotype associations, knowing the effective population size is helpful because N_e influences the extent of linkage disequilibrium along chromosomes, which in turn determines the density of markers and molecular methods needed to conduct a powerful genome-wide scan (e.g. (Kardos et al. 2016).

The increasing diversity and complexity of analysis also requires that code be well annotated and highly reproducible. A number of instructors shared version-controlled worksheets and R code via Github including Racheal Bay, Eric Anderson (Southwest Fisheries Science Center), Joanna Kelley, and Brenna Forester (Colorado State University). Kelley, for example, provided instruction and materials for transcriptome assembly and quantifying differential gene expression (https://github.com/jokelley/congen-2018). Also of discussion was the increasing availability of R packages to efficiently analyze and visualize NGS datasets and results and the importance this has in increasing reproducibility and reliability, and lowering the barrier on bioinformatics and data analysis in general (Paradis et al. 2017).

Summary and Conclusions

In conclusion, major conceptual advances discussed at ConGen 2018 include estimating the effective population size per year or generation (e.g. N_b with age structure, using thousands of loci) and using adaptive genetic information to identify conservation units. New approaches have emerged for cheaper genome-wide data production (e.g. Rapture) and data analysis (e.g.

295 major updates in Stacks). Emphasis in recent years at ConGen including the use of tools
296 becoming more cost-effective and available to conservation genomics including rapture,
297 transcriptomics, epigenomics, genome-wide and reference-genome-based work. The purpose of
298 ConGen remains to introduce recent novel techniques and approaches to a wide range (globally
299 and career path) of students. Recent work by ConGen instructors and other researchers include
300 large multifaceted data sets (e.g. see Transcriptomics and Epigenomics; and Forester et al. 2018).
301 A researcher now often has multiple data types that may include everything from *de novo*
302 genome assemblies to RADseq to differential gene expression among populations – and more.
303 While the amount of genomic data production grows exponentially, the continuing challenge for
304 genomicists remains in obtaining a solid foundation in population genetics theory, data filtering,
305 and computational analysis. Through training and experiences such as those available at
306 workshops like ConGen 2018, the modern conservation and population genomicist will be able
307 to examine a wide range of central questions, evaluate the appropriate tools for data production
308 and analysis, and integrate across different data types from RADseq to whole genome
309 resequencing, RNAseq, and more. As the field continues to evolve, we hope this review of
310 ConGen 2018 will help serve as a benchmark and starting point for information and references
311 for readers from multiple disciplines world-wide.
312
313 Acknowledgements. This work was supported by National Science Foundation grant DEB-
314 1655809 to P.A.H. and National Science Foundation grant DEB-1639014, DOB and NASA
315 grant NNX14AB84G to B.K.H. and G. L. The authors thank the many instructors and students of
316 ConGen 2018 for invaluable discussion throughout ConGen 2018; as well as the three
317 anonymous reviewers whose insight improved this manuscript. We especially thank R. Waples

318 who is the only instructor to have lectured at all ten ConGen courses (plus helped organize them)
319 since its inception in 2006; he also was awarded the prestigious 2018 Molecular Ecology Prize
320 while at ConGen 2018.

References

Ali, O. A., S. M. O'Rourke, S. J. Amish, M. H. Meek, G. Luikart, C. Jeffres, and M. R. Miller. 2016. 'RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping', *Genetics*, 202: 389-400.

Allendorf, Fred W. 2017. 'Genetics and the conservation of natural populations: allozymes to genomes', *Molecular ecology*, 26: 420-30.

Allendorf, Fred W., Paul A. Hohenlohe, and Gordon Luikart. 2010. 'Genomics and the future of conservation genetics', *Nature reviews genetics*, 11: 697.

Allendorf, Frederick William. 1977. 'GENETIC VARIABILITY IN A SPECIES POSSESSING EXTENSIVE GENE DUPLICATION: GENETIC INTERPRETATION OF DUPLICATE LOCI AND EXAMINATION OF GENETIC VARIATION IN POPULATIONS OF RAINBOW TROUT'.

Andrews, Kimberly R., Jeffrey M. Good, Michael R. Miller, Gordon Luikart, and Paul A. Hohenlohe. 2016. 'Harnessing the power of RADseq for ecological and evolutionary genomics', *Nature Reviews Genetics*, 17: 81.

Andrews, Kimberly R., and Gordon Luikart. 2014. 'Recent novel approaches for population genomics data analysis', *Molecular ecology*, 23: 1661-67.

Arciniega, Martha, Anthony J. Clemento, Michael R. Miller, Matt Peterson, John Carlos Garza, and Devon E. Pearse. 2016. 'Parallel evolution of the summer steelhead ecotype in multiple populations from Oregon and Northern California', *Conservation genetics*, 17: 165-75.

Barshis, Daniel J., Jason T. Ladner, Thomas A. Oliver, François O. Seneca, Nikki Traylor-Knowles, and Stephen R. Palumbi. 2013. 'Genomic basis for coral resilience to climate change', *Proceedings of the National Academy of Sciences*, 110: 1387-92.

Bay, R. A., N. Rose, R. Barrett, L. Bernatchez, C. K. Ghalambor, J. R. Lasky, R. B. Brem, S. R. Palumbi, and P. Ralph. 2017. 'Predicting Responses to Contemporary Environmental Change Using Evolutionary Response Architectures', *Am Nat*, 189: 463-73.

Bay, Rachael A., Ryan J. Harrigan, Vinh Le Underwood, H. Lisle Gibbs, Thomas B. Smith, and Kristen Ruegg. 2018. 'Genomic signals of selection predict climate-driven population declines in a migratory bird', *Science*, 359: 83-86.

Benestan, Laura, Thierry Gosselin, Charles Perrier, Bernard Sainte-Marie, Rémy Rochette, and Louis Bernatchez. 2015. 'RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the A merican lobster (H omarus americanus)', *Molecular ecology*, 24: 3299-315.

Bi, Ke, Dan Vanderpool, Sonal Singhal, Tyler Linderoth, Craig Moritz, and Jeffrey M. Good. 2012. 'Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales', *BMC genomics*, 13: 403.

Catchen, Julian, Paul A. Hohenlohe, Susan Bassham, Angel Amores, and William A. Cresko. 2013. 'Stacks: an analysis tool set for population genomics', *Molecular ecology*, 22: 3124-40.

Ceballos, Francisco C., Peter K. Joshi, David W. Clark, Michèle Ramsay, and James F. Wilson. 2018. 'Runs of homozygosity: windows into population history and trait architecture', *Nature Reviews Genetics*, 19: 220.

Charlesworth, Deborah, Nicholas H. Barton, and Brian Charlesworth. 2017. 'The sources of adaptive variation', *Proceedings of the Royal Society B: Biological Sciences*, 284: 20162864.

Chen, Zhongqi, Anthony P. Farrell, Amanda Matala, Nicholas Hoffman, and Shawn R. Narum. 2018. 'Physiological and genomic signatures of evolutionary thermal adaptation in redband trout from extreme climates', *Evolutionary applications*, 11: 1686-99.

Chilcote, Mark W., Bruce A. Crawford, and Steven A. Leider. 1980. 'A genetic comparison of sympatric populations of summer and winter steelheads', *Transactions of the American Fisheries Society*, 109: 203-06.

Christie, M. R., M. L. Marine, S. E. Fox, R. A. French, and M. S. Blouin. 2016. 'A single generation of domestication heritably alters the expression of hundreds of genes', *Nat Commun*, 7: 10676.

Do, C., R. S. Waples, D. Peel, G. M. Macbeth, B. J. Tillett, and J. R. Ovenden. 2014a. 'NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data', *Mol Ecol Resour*, 14: 209-14.

Do, Chi, Robin S. Waples, David Peel, G. M. Macbeth, Bree J. Tillett, and Jennifer R. Ovenden. 2014b. 'NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data', *Molecular ecology resources*, 14: 209-14.

Ellis, Nick, Stephen J. Smith, and C. Roland Pitcher. 2012. 'Gradient forests: calculating importance gradients on physical predictors', *Ecology*, 93: 156-68.

Ferchaud, Anne-Laure, and Michael M. Hansen. 2016. 'The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: three-spine sticklebacks in divergent environments', *Molecular ecology*, 25: 238-59.

Fitzpatrick, M. C., and S. R. Keller. 2015. 'Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation', *Ecol Lett*, 18: 1-16.

Forester, Brenna R., Erin L. Landguth, Brian K. Hand, and Niko Balkenhol. 2018. 'Landscape Genomics for Wildlife Research'.

Fuentes-Pardo, Angela P., and Daniel E. Ruzzante. 2017. 'Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations', *Molecular ecology*, 26: 5369-406.

Funk, W. C., Brenna R. Forester, Sarah J. Converse, Catherine Darst, and Steve Morey. 2018. 'Improving conservation policy with genomics: a guide to integrating adaptive potential into US Endangered Species Act decisions for conservation practitioners and geneticists', *Conservation Genetics*: 1-20.

Funk, W. C., J. K. McKay, P. A. Hohenlohe, and F. W. Allendorf. 2012. 'Harnessing genomics for delineating conservation units', *Trends Ecol Evol*, 27: 489-96.

Garner, B. A., B. K. Hand, S. J. Amish, L. Bernatchez, J. T. Foster, K. M. Miller, P. A. Morin, S. R. Narum, S. J. O'Brien, G. Roffler, W. D. Templin, P. Sunnucks, J. Strait, K. I. Warheit, T. R. Seamons, J. Wenburg, J. Olsen, and G. Luikart. 2016. 'Genomics in Conservation: Case Studies and Bridging the Gap between Data and Application', *Trends Ecol Evol*, 31: 81-83.

Garvin, Michael R., Gary H. Thorgaard, and Shawn R. Narum. 2015. 'Differential expression of genes that control respiration contribute to thermal adaptation in redband trout (Oncorhynchus mykiss gairdneri)', *Genome biology and evolution*, 7: 1404-14.

Gavery, Mackenzie R., Krista M. Nichols, Giles W. Goetz, Mollie A. Middleton, and Penny Swanson. 2018. 'Characterization of Genetic and Epigenetic Variation in Sperm and Red Blood Cells from Adult Hatchery and Natural-Origin Steelhead, Oncorhynchus mykiss', *G3: Genes, Genomes, Genetics*, 8: 3723-36.

Groom, Martha J., Gary K. Meffe, and Carl Ronald Carroll. 2006. *Principles of conservation biology* (Sinauer Associates Sunderland).

Hand, Brian K., Courtney G. Flint, Chris A. Frissell, Clint C. Muhlfeld, Shawn P. Devlin, Brian P. Kennedy, Robert L. Crabtree, W. Arthur McKee, Gordon Luikart, and Jack A. Stanford. 2018. 'A social–ecological perspective for riverscape management in the Columbia River Basin', *Frontiers in Ecology and the Environment*, 16.

Hendricks, Sarah, Eric C. Anderson, Tiago Antao, Louis Bernatchez, Brenna R. Forester, Brittany Garner, Brian K. Hand, Paul A. Hohenlohe, Martin Kardos, and Ben Koop. 2018. 'Recent advances in conservation and population genomics data analysis', *Evolutionary Applications*, 11: 1197-211.

Hendricks, Sarah, Eric C. Anderson, Tiago Antao, Louis Bernatchez, Brenna R. Forester, Brittany Garner, Brian K. Hand, Paul A. Hohenlohe, Martin Kardos, Ben Koop, Arun Sethuraman, Robin S. Waples, and Gordon Luikart. 2018. 'Recent advances in conservation and population genomics data analysis', *Evolutionary Applications*, 11: 1197-211.

Hess, Jon E., Joseph S. Zendt, Amanda R. Matala, and Shawn R. Narum. 2016. 'Genetic basis of adult migration timing in anadromous steelhead discovered through multivariate association testing', *Proc. R. Soc. B*, 283: 20153064.

Hollenbeck, C. M., D. S. Portnoy, and J. R. Gold. 2016. 'A method for detecting recent changes in contemporary effective population size from linkage disequilibrium at linked and unlinked loci', *Heredity*, 117: 207.

Hunter, Margaret E., Sean M. Hoban, Michael W. Bruford, Gernot Segelbacher, and Louis Bernatchez. 2018. 'Next-generation conservation genetics and biodiversity monitoring', *Evolutionary applications*, 11: 1029-34.

Kamath, Pauline L., Mark A. Haroldson, Gordon Luikart, David Paetkau, Craig Whitman, and Frank T. Van Manen. 2015. 'Multiple estimates of effective population size for monitoring a long-lived vertebrate: an application to Y ellowstone grizzly bears', *Molecular ecology*, 24: 5507-21.

Kardos, M., and A. B. A. Shafer. 2018. 'The Peril of Gene-Targeted Conservation', *Trends Ecol Evol*, 33: 827-39.

Kardos, Marty, Arild Husby, S. Eryn McFarlane, Anna Qvarnström, and Hans Ellegren. 2016. 'Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in natural populations', *Molecular Ecology Resources*, 16: 727-41.

Kelly, Morgan. 2019. 'Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes', *Philosophical Transactions of the Royal Society B*, 374: 20180176.

Langin, Katie. 2018. "A different animal." In.: American Association for the Advancement of Science.

Le Luyer, Jérémy, Martin Laporte, Terry D. Beacham, Karia H. Kaukinen, Ruth E. Withler, Jong S. Leong, Eric B. Rondeau, Ben F. Koop, and Louis Bernatchez. 2017. 'Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon', *Proceedings of the National Academy of Sciences*, 114: 12964-69.

Lynch, Michael, and Bruce Walsh. 2007. *The origins of genome architecture* (Sinauer Associates Sunderland (MA)).

Margres, Mark J., Menna E. Jones, Brendan Epstein, Douglas H. Kerlin, Sebastien Comte, Samantha Fox, Alexandra K. Fraik, Sarah A. Hendricks, Stewart Huxtable, and Shelly Lachish. 2018. 'Large-effect loci affect survival in Tasmanian devils (Sarcophilus harrisii) infected with a transmissible cancer', *Molecular ecology*, 27: 4189-99.

Maruki, Takahiro, and Michael Lynch. 2017. 'Genotype calling from population-genomic sequencing data', *G3: Genes, Genomes, Genetics*: g3-117.

Meissner, Alexander, Andreas Gnirke, George W. Bell, Bernard Ramsahoye, Eric S. Lander, and Rudolf Jaenisch. 2005. 'Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis', *Nucleic acids research*, 33: 5868-77.

Micheletti, Steven J., Jon E. Hess, Joseph S. Zendt, and Shawn R. Narum. 2018. 'Selection at a genomic region of major effect is responsible for evolution of complex life histories in anadromous steelhead', *BMC evolutionary biology*, 18: 140.

Narum, Shawn R., Alex Di Genova, Steven J. Micheletti, and Alejandro Maass. 2018. 'Genomic variation underlying complex life-history traits revealed by genome sequencing in Chinook salmon', *Proceedings of the Royal Society B: Biological Sciences*, 285: 20180935.

Nicotra, Adrienne B., Erik A. Beever, Amanda L. Robertson, Gretchen E. Hofmann, and John O'Leary. 2015. 'Assessing the components of adaptive capacity to improve conservation and management efforts under global change', *Conservation Biology*, 29: 1268-78.

Nunziata, S. O., and D. W. Weisrock. 2018. 'Estimation of contemporary effective population size and population declines using RAD sequence data', *Heredity (Edinb)*, 120: 196-207.

Paradis, Emmanuel, Thierry Gosselin, Jerome Goudet, Thibaut Jombart, and Klaus Schliep. 2017. 'Linking genomics and population genetics with R', *Molecular ecology resources*, 17: 54-66.

Peel, David, Robin S. Waples, G. M. Macbeth, Chi Do, and Jennifer R. Ovenden. 2013. 'Accounting for missing data in the estimation of contemporary genetic effective population size (Ne)', *Molecular Ecology Resources*, 13: 243-53.

Pennisi, Elizabeth. 2013. "Evolution heresy? Epigenetics underlies heritable plant traits." In.: American Association for the Advancement of Science.

Pratlong, Marine, Anne Haguenauer, Olivier Chabrol, Christophe Klopp, Pierre Pontarotti, and Didier Aurelle. 2015. 'The red coral (Corallium rubrum) transcriptome: a new resource for population genetics and local adaptation studies', *Molecular ecology resources*, 15: 1205-15.

Prince, Daniel J., Sean M. O'Rourke, Tasha Q. Thompson, Omar A. Ali, Hannah S. Lyman, Ismail K. Saglam, Thomas J. Hotaling, Adrian P. Spidle, and Michael R. Miller. 2017. 'The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation', *Science advances*, 3: e1603198.

493 Rochette, Nicolas C., and Julian M. Catchen. 2017. 'Deriving genotypes from RAD-seq short-read
494 data using Stacks', *nature protocols*, 12: 2640.
495 Rochette, Nicolas C., Angel G. Rivera-Colón, and Julian M. Catchen. 2019. 'Stacks 2: Analytical
496 Methods for Paired-end Sequencing Improve RADseq-based Population Genomics',
497 *bioRxiv*: 615385.
498 Schuierer, Sven, Walter Carbone, Judith Knehr, Virginie Petitjean, Anita Fernandez, Marc Sultan,
499 and Guglielmo Roma. 2017. 'A comprehensive assessment of RNA-seq protocols for
500 degraded and low-quantity samples', *BMC genomics*, 18: 442.
501 Schweizer, Rena M., Jacqueline Robinson, Ryan Harrigan, Pedro Silva, Marco Galverni, Marco
502 Musiani, Richard E. Green, John Novembre, and Robert K. Wayne. 2016. 'Targeted
503 capture and resequencing of 1040 genes reveal environmentally driven functional
504 variation in grey wolves', *Molecular ecology*, 25: 357-79.
505 Schweyen, Hannah, Andrey Rozenberg, and Florian Leese. 2014. 'Detection and removal of PCR
506 duplicates in population genomic ddRAD studies by addition of a degenerate base
507 region (DBR) in sequencing adapters', *The Biological Bulletin*, 227: 146-60.
508 Shafer, A. B., J. B. Wolf, P. C. Alves, L. Bergstrom, M. W. Bruford, I. Brannstrom, G. Colling, L.
509 Dalen, L. De Meester, R. Ekblom, K. D. Fawcett, S. Fior, M. Hajibabaei, J. A. Hill, A. R.
510 Hoezel, J. Hoglund, E. L. Jensen, J. Krause, T. N. Kristensen, M. Krutzen, J. K. McKay, A. J.
511 Norman, R. Ogden, E. M. Osterling, N. J. Ouborg, J. Piccolo, D. Popovic, C. R. Primmer, F.
512 A. Reed, M. Roumet, J. Salmona, T. Schenekar, M. K. Schwartz, G. Segelbacher, H. Senn,
513 J. Thaulow, M. Valtonen, A. Veale, P. Vergeer, N. Vijay, C. Vila, M. Weissensteiner, L.
514 Wennerstrom, C. W. Wheat, and P. Zielinski. 2015. 'Genomics and the challenging
515 translation into conservation practice', *Trends Ecol Evol*, 30: 78-87.
516 Thompson, Tasha Q., M. Renee Bellinger, Sean M. O'Rourke, Daniel J. Prince, Alexander E.
517 Stevenson, Antonia T. Rodrigues, Matthew R. Sloat, Camilla F. Speller, Dongya Y. Yang,
518 and Virginia L. Butler. 2019. 'Anthropogenic habitat alteration leads to rapid loss of
519 adaptive variation and restoration potential in wild salmon populations', *Proceedings of*
520 *the National Academy of Sciences*, 116: 177-86.
521 Thompson, Tasha Q., Renee M. Bellinger, Sean M. O'Rourke, Daniel J. Prince, Alexander E.
522 Stevenson, Antonia T. Rodrigues, Matthew R. Sloat, Camilla F. Speller, Dongya Y. Yang,
523 and Virginia L. Butler. 2018. 'Anthropogenic habitat alteration leads to rapid loss of
524 adaptive variation and restoration potential in wild salmon populations', *bioRxiv*:
525 310714.
526 Wang, J., E. Santiago, and A. Caballero. 2016. 'Prediction and estimation of effective population
527 size', *Heredity (Edinb)*, 117: 193-206.
528 Wang, Zhong, Mark Gerstein, and Michael Snyder. 2009. 'RNA-Seq: a revolutionary tool for
529 transcriptomics', *Nature reviews genetics*, 10: 57.
530 Waples, R. K., W. A. Larson, and R. S. Waples. 2016. 'Estimating contemporary effective
531 population size in non-model species using linkage disequilibrium across thousands of
532 loci', *Heredity (Edinb)*, 117: 233-40.
533 Waples, Robin S. 1991. 'Pacific salmon, Oncorhynchus spp., and the definition of" species"
534 under the Endangered Species Act', *Marine Fisheries Review*, 53: 11-22.
535 Waples, Robin S., Tiago Antao, and Gordon Luikart. 2014. 'Effects of overlapping generations on
536 linkage disequilibrium estimates of effective population size', *Genetics*: genetics-114.

Waples, Robin S., and C. H. I. Do. 2008. 'LDNE: a program for estimating effective population size from data on linkage disequilibrium', *Molecular ecology resources*, 8: 753-56.
Waples, Robin S., and Oscar Gaggiotti. 2006. 'INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity', *Molecular ecology*, 15: 1419-39.
Waples, Robin S., Peter M. Grewe, Mark W. Bravington, Richard Hillary, and Pierre Feutry. 2018. 'Robust estimates of a high Ne/N ratio in a top marine predator, southern bluefin tuna', *Science advances*, 4: eaar7759.
Waples, Robin S., and Steven T. Lindley. 2018. 'Genomics and conservation units: The genetic basis of adult migration timing in Pacific salmonids', *Evolutionary Applications*, 11: 1518-26.
Waples, Robin S., Kim Scribner, Jennifer Moore, Hope Draheim, Dwayne Etter, and Mark Boersen. 2018. 'Accounting for age structure and spatial structure in eco-evolutionary analyses of a large, mobile vertebrate', *Journal of Heredity*, 1: 15.
Waples, Robin S., David J. Teel, James M. Myers, and Anne R. Marshall. 2004. 'Life-history divergence in Chinook salmon: historic contingency and parallel evolution', *Evolution*, 58: 386-403.
Waples, Robin S., and Masashi Yokota. 2007. 'Temporal estimates of effective population size in species with overlapping generations', *Genetics*, 175: 219-33.
Zhou, Ying, Xiaowen Tian, Brian L. Browning, and Sharon R. Browning. 2018. 'POPdemog: visualizing population demographic history from simulation scripts', *Bioinformatics*, 1: 2.

560 **Figures**
561 Figure 1. Adaptive locus and alleles in Chinook salmon

Box 1. How will an adaptive locus influence listing of distinct salmonid populations under the Endangered Species Act (ESA) of the United States?

Chinook salmon (*Oncorhynchus tshawytscha*) and steelhead (*O. mykiss*) have distinct spring (premature) and fall (mature, normal) migratory phenotypes (called runs) in several river basins across western United States. The spring run phenotype differs substantially in behavior and physiology , but has declined in abundance throughout the ranges of both species. Spring-run phenotypes have ecological, economic, and cultural importance, and are valuable to commerce and ecosystems for their greater fat content (REF). They also have had long histories with local tribes, including documented ritualistic management (REF). Due to reliance on cool, clean water in the summer, spring-run salmonids are particularly vulnerable to anthropogenic effects and have dramatically declined (Thompson et al 2018).

Low genetic divergence (e.g. $F_{ST} < 0.03$) between premature and mature migrants within local rivers was found by multiple studies (Allendorf 1977; Chilcote, Crawford, and Leider 1980; Waples et al. 2004; Arciniega et al. 2016). Based on these findings, premature migrant forms did not meet the first criterion for ESU status, sufficient reproductive isolation (Waples and Lindley 2018). However, recent genomic studies by Prince et al (2017) and Thompson (2018) have identified a single locus that has a major effect on the migration phenotype and highlighted the potential for the loss of allelic variation at this locus to have significant ecological consequences, leading to legal action (Hess et al. 2016; Prince et al. 2017; Micheletti et al. 2018; Thompson et al. 2018; Narum et al. 2018; Thompson et al. 2019) NMFS 2018). Prince et al. (2017) conducted a genome-wide association study that identified a single genetic locus (*GREB1L*) associated with premature migration (Fig. 1). Further phylogenetic analyses

584 suggested that the *GREB1L* alleles determining the premature migrant phenotype arose only once
585 in each species, and subsequently spread through dispersal and positive selection.
586 Thompson et al. (2018) further examined selection against the premature migrant
587 phenotype of Chinook salmon in the Rogue River in Oregon after the construction of a dam.
588 They estimated the strength of selection needed to explain the change in allele frequencies at
589 *GREB1L* under multiple dominance scenarios and predicted allele frequencies in future
590 populations. Results suggested that the premature migration allele is likely codominant with
591 respect to fitness and may be lost from the population if the current selection pressure continues
592 (Fig 1b).
593 Together, these findings suggest that the premature migration phenotype (and allele) is
594 vulnerable to loss and unlikely to reappear for a long time if lost from a population. Populations
595 where *GREB1L* early-migration alleles are prevalent may deserve special legal protection. Based
596 on these results, the Karuk Tribe submitted a petition to list the Klamath premature Chinook
597 under the ESA (Langin 2018). In February 2018, National Oceanic and Atmospheric
598 Administration (NOAA) Fisheries announced a finding of substantial scientific evidence
599 indicating the creation and listing of a new ESU as threatened or endangered may be warranted.
600 At the time of writing, the National Marine Fisheries status review of the Upper Klamath and
601 Trinity River Chinook salmon was still pending. The decision on whether to list Klamath
602 premature Chinook could have wide-reaching implications for conservation (Waples and Lindley
603 2018).
604 These studies and the resultant legal action have recharged debate over whether, when,
605 and how species should be managed for single genes (Kardos and Shafer 2018). A concern is
606 that as genomics continues to make it easier to find adaptive genetic variation, management units

607 could be over-split as more and more important loci and alleles are identified. As this case study
608 in Pacific salmon shows, iterative and focused genomic studies have the power to identify crucial
609 adaptive variation and to inform long-standing debates.

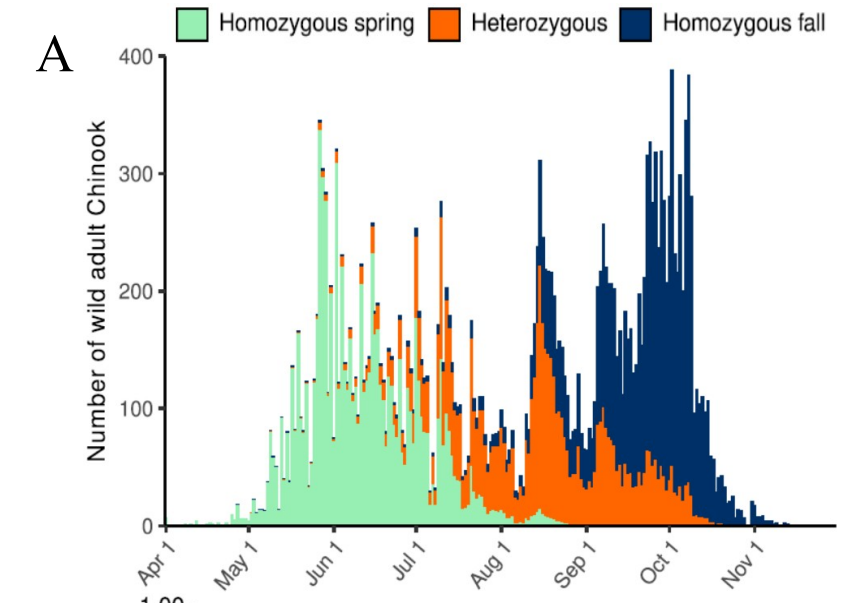
Box 2. How will changes in DNA methylation influence adaptation to artificial environments in hatchery fish?

A common goal of captive breeding programs is to support declining wild populations (i.e. genetic rescue); however, there is concern that rearing in artificial conditions may inadvertently reduce fitness. In conservation salmonid hatcheries, there is mounting evidence that tank-rearing conditions can induce developmental plasticity and impact life-history traits. To examine the role of epigenetic changes in hatchery-reared steelhead trout (Oncorhynchus mykiss), Gavery et al. (2019) raised steelhead in an artificial stream and small simulated hatchery tank for two years, well past germ cell differentiation, then sampled individuals and performed reduced representation bisulfite sequencing (Meissner et al. 2005) to determine methylation patterns. After accounting for familial relationships influencing methylation patterns, they were able to discern up-methylated and down-methylated gene differences between their two conditions (artificial stream vs tank). While family relatedness had the largest effect, environmental differences also caused significant changes in the methylation pattern. If these epigenetic changes occur at an early stage in development in response to environmental pressures, they may not only affect the organism's growth, but will continue to persist well past the time when those environmental pressures are no longer present. This has implications for conservation of salmonids and other species if environmentally induced epigenetic shifts are transmitted to offspring and grand offspring. For example, if hatchery-adaptive epigenetic changes are transmitted to wild fish, the fitness of wild fish could decline (Christie et al. 2016; Le Luyer et al. 2017). There is substantial evidence of maladaptive introgression in wild populations (REF), though more work must be conducted to determine if epigenetic changes can

632 persist, be transmitted across multiple generations, and spread within and among natural
633 populations (Charlesworth et al. 2017).
634

Box 3. How will genomic vulnerability of yellow warblers influence their evolutionary response to climate change?

In their workshop lecture, Bay and Bossou invited students to assess genomic vulnerability of the yellow warbler (*Setophaga petechia*), a migratory songbird distributed across much of North America (Fig. 2; Bay et al. 2018). First, students identified the environmental variables that best explained variation at a subset of genome-wide SNPs using gradient forest analysis, a regression tree-based machine learning approach (Ellis, Smith, and Pitcher 2012). Then, genomic vulnerability was calculated as the difference between current versus predicted gradient forest-transformed climate variables. A significant negative association was found between genomic vulnerability and current population trends, suggesting that populations with high genomic vulnerability may have already been impacted (Bay et al. 2018). This approach provides a useful starting point to incorporate evolution into models that predict the effects of climate change on biodiversity. Important future extensions of the model could include incorporating additional evolutionary components, such as gene flow and population sizes. Predictive modeling, such as the strategy taught by Bay and Bossou, will become increasingly useful for conservation as it incorporates both local adaptation and projected environmental conditions.


A
Homozygous spring
Heterozygous
Homozygous fall
Number of wild adult Chinook
400
300
200
100
0
Apr 1
May 1
Jun 1
Jul 1
Aug 1
Sep 1
Oct 1
Nov 1


C

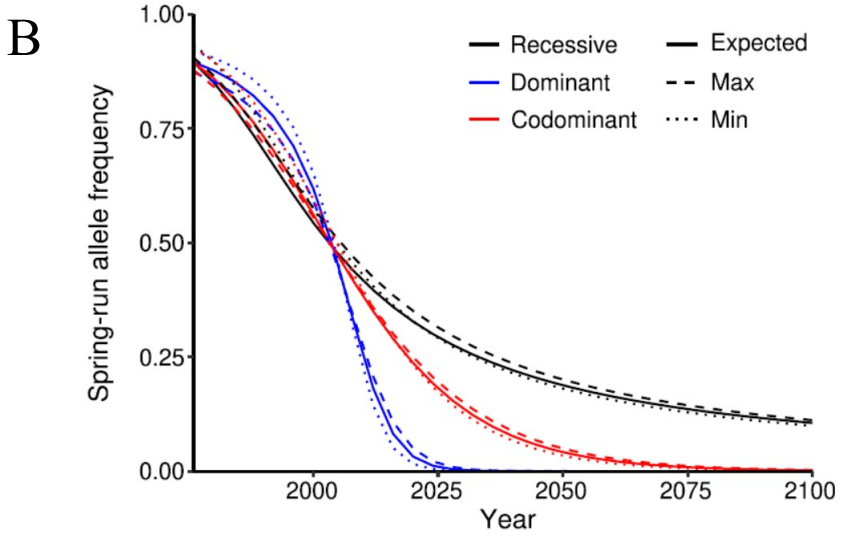

B
Recessive
Dominant
Codominant
Expected
Max
Min
Spring-run allele frequency
1.00
0.75
0.50
0.25
0.00
2000
2025
2050
2075
2100
Year

Figure 1. (A) Stacked bar graph representing the number of wild adult Chinook salmon passing Gold Ray Fish Counting Station on the Rogue river in 2004; colors represent estimated proportion of each GREB1L locus genotype. (B) Selection modeling in Rogue Chinook. Curves representing the decline (or loss) of the spring-run allele frequency over time under a recessive, dominant, or codominant scenario. Spring-run alleles are thought to be codominant and predicted to be lost by ~2075 (red curve). The modeling assumes random mating and no genetic drift. (C) Image of a Chinook salmon. Figure modified from Thompson et al. 2019.

Figure 2. The wide breeding range of the yellow warbler (*Setophaga petechia*), pictured here, and recent population declines in some regions motivated the hands-on tutorial of Bay and Bossou. Photo by Daniel Karp, reproduced with permission.

670

671

672

673 Figure 3. Empirical examples provided by instructors at ConGen 2018 across a broad range of
674 data types, questions, and taxa. A) RAD-Capture and GWAS in characterizing the genetic
675 architecture of disease-related traits in Tasmanian devils (*Sarcophilus harisii*; Margres et al
676 2018), B) targeted-capture, demographic modeling, and linkage-disequilibrium analysis in
677 understanding the evolutionary history of color polymorphism of the grey wolf (*Canis lupus*;
678 Schweizer et al 2018), and C) RADseq and analysis of population structure in identifying range
679 expansion and hybridization of the tamarisk beetle (*Diorhabda* spp.), a recently introduced
680 biocontrol agent. Photos by A) Menna Jones, B) Marco Musiani, and C) Ed Kosmicki;
681 respectively, reproduced with permission.