Chapter 02 Bibliography

Matthew D. Buehler

SNPs vs Microsatellites

Shawna J Zimmerman et al. "An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern". In: *BMC genomics* 21 (2020), pp. 1–16, SNP/Micro Comparison Grouse

Demographic Parameters

Population Connectivity

Desislava Petkova et al. "Visualizing spatial population structure with estimated effective migration surfaces". In: *Nature genetics* 48.1 (2016), pp. 94–100, EEMS Description

Steven M Mussmann et al. "BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data". In: *Methods in Ecology and Evolution* 10.10 (2019), pp. 1808–1813, BA3-SNP

Gregory A Wilson and Bruce Rannala. "Bayesian inference of recent migration rates using multilocus genotypes". In: *Genetics* 163.3 (2003), pp. 1177–1191, BayesAss3

Conservation Genomics

Soraia Barbosa et al. "Integrative approaches to guide conservation decisions: using genomics to define conservation units and functional corridors". In: *Molecular Ecology* 27.17 (2018), pp. 3452–3465, Delimiting Conservation Units

Population Genetics Bibliography

Barbosa et al.: Integrative approaches to guide conservation decisions: using genomics to define conservation units and functional corridors **barbosa2018IntegrativeConservationGenomics**

Soraia Barbosa, Frederico Mestre, Thomas A White, Joana Paupério, Paulo C Alves, and Jeremy B Searle. "Integrative approaches to guide conservation decisions: using genomics to define conservation units and functional corridors". In: *Molecular Ecology* 27.17 (2018), pp. 3452–3465.

Annotations: **USE THIS PAPER TO SET UP THE CONSERVASTION UNITS FOR EIS**. This paper deliniates and proposes a unit schema for defining populations for conservation. First is the Conservation Unit (CU) or the Evolutionary Significant Unit. This designation lumps populations based on their genetic composistion and should be seperated, ideally, by phylogeographic history (e.g., the GA and FL EIS populations). These conservation units would be delimited using all loci in a genomic dataset.ESUs can be subdivided further into management units and adaptive units. Management units are the short term currency of conservation, clustering populations based on their neutral genetic diversity. This delimits populations that are genetically unique from one another. Adaptive units are identified using potential adaptive loci. I could use outlier methods to identify putatively adaptive loci in the RAD data. Reseq data would be a better method for delimiting the AUs. In small populations, genetic drift may inflate the number of detected MUs, however, in large populations MUs may be falesly suppressed due to lingering ancestral alleles. This paper outlines a way to idenify functional corridors for species acounting for gene flow between populations, as well as, habitat. Essentially, combine environmentla and genetic connectivities to find overlaps in where gene flow is occuring and how it correlates wtih suitable habitats.

Mussmann et al.: BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data **mussmann2019ba3snps**

Steven M Mussmann, Marlis R Douglas, Tyler K Chafin, and Michael E Douglas. "BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data". In: *Methods in Ecology and Evolution* 10.10 (2019), pp. 1808–1813.

Annotations: This program is a modification of BayesAss3 allowing SNPs to be utilized for detection of isolated popualtions. BA3-SNP can be used to delimit management units based on the ability of population to maintain and grow in population size due to its demographic characterstics. It also calculates a dispersal parameter to identify the rate of immigration

from the population. I need to check the BayesAss3 manual to see if some of the future directions they proposed have been addressed in more recent versions.

Petkova et al.: Visualizing spatial population structure with estimated effective migration surfaces **petkova2016EEMSDescription**

Desislava Petkova, John Novembre, and Matthew Stephens. "Visualizing spatial population structure with estimated effective migration surfaces". In: *Nature genetics* 48.1 (2016), pp. 94–100.

Annotations: This paper describes the progarm EEMS. EEMS is based on a stepping stone model and estimates the rate in which genetic simlarity decays between samples. This scheme has benefits over using PCAs or clusting methods because EEMS accounts for the geographic proximity of samples when comparing to genetic similarities. Although PCAs have been used to investigate the effects of IBD, these analyses have to be done *posthoc*. Clustering methods are suceptible to k estimation deviation caused by sampling bias. EEMS is unable to detect gene flow directionality because the base model assumes symetric equal gene flow between popualtions. In the same vein that effective population size represents the idealized population size needed to maintain the levels of genetic diversity observed in a population, EEMS calucaltes effective migration rates that idealize the migration rate parameters under a stepping stone model that would produce the genetic decay or simlarity between the given samples. EEMS also calculates a effective diversity parameters within each deme that reflects the expected genetic dissimiarility between two individuals within a deme. This is somewhat analagous to expected heterozygosity within the deme. Demes with higher effective diversity are going to have higher levels of heteorzygosity. For the resistance grids, their size is arbitrary so multiple scales should be tried based on the sampling density.

Wilson et al.: Bayesian inference of recent migration rates using multilocus genotypes **wilson2003BayesAss**

Gregory A Wilson and Bruce Rannala. "Bayesian inference of recent migration rates using multilocus genotypes". In: *Genetics* 163.3 (2003), pp. 1177–1191.

Annotations: This is the base paper for BayesAss3 . There are several parameters that can be estimated using the MCMC including: migration rate (m), allele frequencies, inbreeding coefficiants, the proportion of non-immigrant, first generation immigrants, and second generation immigrants within a population, and finally the probability that an individuals belongs to one of these classifications. Again, sampling is an imporatnt consideration, if the source population for migrants is not sampled then the program will not be able to accurately estimate the migration rates.

Zimmerman et al.: An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern **zimmerman2020empirical**

Shawna J Zimmerman, Cameron L Aldridge, and Sara J Oyler-McCance. "An empirical

comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern". In: *BMC genomics* 21 (2020), pp. 1–16.

Annotations: This paper identified the diffferences between SNP and microsatellite markers for estimating demographic parameters (Ho,He,Fis,AR) and population differentiation (fst, gst, jostD). Generally they found that the demographic parameters were correlated with each other except for Ho. Again found a high level of correlation between the population differentiation statistics, note that Fst and Gst are considered fixation indicies and JostD is considered an allelic differentiation metric. Gst estimates were higher in the microsatellites, however, the theoretical max heterozygosity for microsats is double that of SNPs, and given Gst is based on heterozygosity, it makes sense that this metric is higher. Finally, the authors found that there was similar ability of SNPs from a smaller sub-set of individuals to cluster genetically similar samples, compared to a full microsatellite dataset. There should be some caution used when placing to much emphasis on the ESUs or conservation units. Adaptive variation may be adaptive for a specific contemporary environment, but focusing on purifying the genetics of ESUs may lead to loss of future adaptive potential. There is the potential to consider the clustering of populations based on the presence of putatively adpative alleles.