Abstract draft: 1000 words or less due April 15

Main text (with in-text citations is 864 words) Will have to reduce to include Preiminary Results and Discussion

Title: TBD

**Introduction**

The field of landscape genetics commonly combines methods used in population genetics, landscape ecology, and spatial statistics. The primary goal of many landscape genetic studies is to associate spatial genetic structure (neutral and/or adaptive) with landscape structure and partition variance between predictors of interest (Balkenhol, Gugerli, et al., 2009; Balkenhol, Waits, et al., 2009; Holderegger & Wagner, 2008; Manel et al., 2003; Storfer et al., 2007). As a relatively new field, first defined in 2003 (Manel et al., 2003), landscape genetics is growing rapidly and has plenty of challenges. Landscape genetics as a relatively new field has been characterized by an abundance of review articles, followed by empirical case studies in various system. In order to advance the field and contribute to theory, we should now move to synthesize across studies and systems (Dyer 2015). Meta-analysis based on the strength of spatial genetic structure would be the way to go. However, it is unclear what is the best metric and under what conditions it can be compared between studies differing in sampling design or underlying population genetic structure, which may be more complex than isolation-by-distance.

~~These two factors are part of the larger problem within landscape genetics: the lack of comparability between studies (Balkenhol et al., 2015; Balkenhol & Landguth, 2011; Dyer, 2015; Holderegger & Wagner, 2008; Landguth et al., 2015).~~ Without the ability to compare between studies, one cannot search for common patterns or perform meta-analysis (Dyer, 2015).

And by extension, the lack of comparability between studies makes it exceptionally difficult to develop a unifying theory specific to landscape genetic, which is essential because landscape genetic studies routinely fail to meet many assumptions of existing population genetic theory (Balkenhol & Landguth, 2011; Landguth et al., 2015). This study addresses the fundamental question of comparability between studies using the program MEMgene (Galpern et al., 2014).

We selected MEMgene as the primary analysis approach for this study because it detects even cryptic spatial genetic structure, is robust to non-stationarity, and provides commonly used metrics of effect size, R2 and Moran’s *I* (Galpern et al., 2014)*.* Additionally, spatial filtering and Moran’s Eigenvector Maps (MEM) are highly relevant because they are common tools used to account for spatial genetic structure when trying to partition genetic variance correlated with predictors of interest (Epperson et al., 2010; Klinga et al., 2019; Manel et al., 2010; Manel & Holderegger, 2013; Richardson et al., 2016).

We used MEMgene for comparative analysis of a simulated dataset, created by Lotterhos and Whitlock (2014) to assess which response metric, R2, Moran’s *I,* or rescaled Moran’s *I*, is more robust and comparable between varied study conditions. The parameters of interest (factors) for this study were the strength of true Fst, spatial configuration of the sampling points, sample size, and the population demographic history. We hope that these simulations will provide a means for comparability between studies and serve as a step towards future meta-analysis and the development of unifying landscape genetic theory.

**Methods**

We are using a simulated dataset consisting of 10,000 haploid, biallelic loci (9,990 neutral, 100 selective) on a quasi-continuous, square landscape (360 x 360). Lotterhos and Whitlock (2014) designed the program to simulate a species with large geographic range, high effective population size, and rapid linkage decay. Loci were generated independently, and their starting allele frequencies were chosen at random from a normal distribution ranging between 0 and 1. Source code for the landscape simulator is located in the Dryad repository (doi: 10.5061/dryad.v8d05).

Variation in population demographic history is represented by replicated landscapes under the following models: island model (IM), isolation by distance (IBD), expansion of one refugium (1R) or two (2R) (Lotterhos & Whitlock, 2014). The IM model served as the null hypothesis and is non-spatial. The IBD model at equilibrium provides a spatial but stationary process, and the 1R and 2R models are spatial and non-stationary. The original dataset controlled for overall Fst by sampling at the time period when the overall Fst equal to ~0.05. We are interested if there are fluctuations in performance of R2, Moran’s *I*, and rescaled Moran’s *I*, over a range of Fst values. For this study,Fst ranges from 0, panmixia, to 0.05. The maximum true Fst  is 0.05 from the original dataset, and we decreased overall Fst values by randomizing of allele frequencies of an increasing number of loci until an Fst of 0, panmixia, was attained.

For all analyses we are using the 9,990 neutral loci and MEMgene to calculate and record of R2, Moran’s *I*, and rescaled Moran’s *I*. We also recorded the sample Fst to compare against the true Fst over a range of values because in landscape genetic studies, the source population does not meet the requirements for equilibrium. Therefore, it is unknown how well the sample Fst represents the true population. The parameters we varied and compared in replicated runs were the: strength of the true Fst and demographic population history, as described above. In addition, we also varied the spatial configuration of the sampling points, number of sampling points, and number of individuals per sampling point, to test the idea that increasing these parameters reduces the noise to signal ratio. We used multi-way ANOVA to …

**Preliminary Results**

????

**Discussion**

????

**Future Directions**

We are considering to expand this pilot study to include more variation in two additional parameters: the definition of the neighbor matrix and selective vs neutral loci.

**Acknowledgements**

**Literature cited**

Balkenhol, N., Cushman, S. A., Storfer, A. T., & Waits, L. P. (2015). Introduction to Landscape Genetics - Concepts, Methods, Applications. In *Landscape Genetics: Concepts, Methods, Applications* (pp. 1–8). Wiley Blackwell. https://doi.org/10.1002/9781118525258.ch01

Balkenhol, N., Gugerli, F., Cushman, S. A., Waits, L. P., Coulon, A., Arntzen, J. W., Holderegger, R., Wagner, H. H., Arens, P., Campagne, P., Dale, V. H., Nicieza, A. G., Smulders, M. J. M., Tedesco, E., Wang, H., & Wasserman, T. (2009). Identifying future research needs in landscape genetics: Where to from here? In *Landscape Ecology* (Vol. 24, Issue 4, pp. 455–463). Springer. https://doi.org/10.1007/s10980-009-9334-z

BALKENHOL, N., & LANDGUTH, E. L. (2011). Simulation modelling in landscape genetics: on the need to go further. *Molecular Ecology*, *20*(4), 667–670. https://doi.org/10.1111/j.1365-294X.2010.04967.x

Balkenhol, N., Waits, L. P., & Dezzani, R. J. (2009). Statistical approaches in landscape genetics: An evaluation of methods for linking landscape and genetic data. *Ecography*, *32*(5), 818–830. https://doi.org/10.1111/j.1600-0587.2009.05807.x

Dyer, R. J. (2015). Is there such a thing as landscape genetics? *Molecular Ecology*, *24*(14). https://doi.org/10.1111/mec.13249

EPPERSON, B. K., MCRAE, B. H., SCRIBNER, K., CUSHMAN, S. A., ROSENBERG, M. S., FORTIN, M.-J., JAMES, P. M. A., MURPHY, M., MANEL, S., LEGENDRE, P., & DALE, M. R. T. (2010). Utility of computer simulations in landscape genetics. *Molecular Ecology*, *19*(17), 3549–3564. https://doi.org/10.1111/j.1365-294X.2010.04678.x

Galpern, P., Peres-Neto, P. R., Polfus, J., & Manseau, M. (2014). MEMGENE: Spatial pattern detection in genetic distance data. *Methods in Ecology and Evolution*, *5*(10), 1116–1120. https://doi.org/10.1111/2041-210X.12240

Holderegger, R., & Wagner, H. H. (2008). Landscape Genetics. *BioScience*, *58*(3), 199–207. https://doi.org/10.1641/b580306

Klinga, P., Mikoláš, M., Smolko, P., Tejkal, M., Höglund, J., & Paule, L. (2019). Considering landscape connectivity and gene flow in the Anthropocene using complementary landscape genetics and habitat modelling approaches. *Landscape Ecology*, *34*(3), 521–536. https://doi.org/10.1007/s10980-019-00789-9

Landguth, E., Cushman, S. A., & Balkenhol, N. (2015). Simulation Modeling in Landscape Genetics. In *Landscape Genetics: Concepts, Methods, Applications* (pp. 99–113). Wiley Blackwell. https://doi.org/10.1002/9781118525258.ch06

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, *23*(9), 2178–2192. https://doi.org/10.1111/mec.12725

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. In *Trends in Ecology and Evolution* (Vol. 28, Issue 10). https://doi.org/10.1016/j.tree.2013.05.012

MANEL, S., JOOST, S., EPPERSON, B. K., HOLDEREGGER, R., STORFER, A., ROSENBERG, M. S., SCRIBNER, K. T., BONIN, A., & FORTIN, M.-J. (2010). Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, *19*(17), 3760–3772. https://doi.org/10.1111/j.1365-294X.2010.04717.x

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. In *Trends in Ecology and Evolution* (Vol. 18, Issue 4, pp. 189–197). Elsevier Ltd. https://doi.org/10.1016/S0169-5347(03)00008-9

Richardson, J. L., Brady, S. P., Wang, I. J., & Spear, S. F. (2016). Navigating the pitfalls and promise of landscape genetics. In *Molecular Ecology* (Vol. 25, Issue 4). https://doi.org/10.1111/mec.13527

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., Dezzani, R., Delmelle, E., Vierling, L., & Waits, L. P. (2007). Putting the “landscape” in landscape genetics. In *Heredity* (Vol. 98, Issue 3, pp. 128–142). Nature Publishing Group. https://doi.org/10.1038/sj.hdy.6800917