**Results:**

*Differences between t-SNE and PCA visualizations of population structure and pedigree data.*

t-SNE and PCA produced similar plots of population structure SNP and microsatellite data. In the SNP dataset, both methods clustered monarch individuals mostly into groups corresponding to their sampling location (Figure 1). The plot produced via t-SNE showed much greater variance in cluster size, with several clusters blown up significantly in comparison to others. This variation corresponds roughly to the genetic diversity observed in these populations: more diverse populations tended to be more loosely clustered than less diverse populations. PCA, however, tended to place populations with respect to their relatedness (individuals from the geographically adjacent islands of Rota, Guam, and Saipan were in adjacent clusters in the bottom-right of the PCA plot in question), which t-SNE did not reliably do. On the microsatellite data, in which the population sub-structuring was much weaker, t-SNE and PCA performed similarly, although PCA seemed to be more driven by 6 outlier alleles from different loci, creating a “streaked” pattern. These alleles were rare in general, and probably result from cutthroat trout (*Oncorhynchus clarkii*) introgression into the population (see Hemstrom et al. (2017) for more information).

Pedigree data, however, produced drastically different results when plotted via PCA or t-SNE. PCA produced large, undefined single clusters from both the SNP and t-SNE datasets, with no visible family structure. In contrast, t-SNE produced many cleanly delineated clusters corresponding to groups of siblings, although some individuals were still clustered in less well defined groups in the center of the plots. Interestingly, both plots reveal additional, useful details about the family structure of the populations at a glance: the microsatellite data shows many clusters containing a handful of individuals which share mothers (circle exteriors) but not fathers (circle centers), and only a few with the opposite whereas the SNP data shows several individuals that were incorrectly assigned parentage during spawning. We found little differences across multiple different perplexity and theta values for the microsatellite data, although very low perplexity values tended to obscure patterns (Figure S2).

We found that t-SNE was effective at visualizing data in formats other than called diploid genotypes. While there was little difference between the PCA and t-SNE plots produced from polyploid microsatellite population structure data, t-SNE seemed to better cluster polyploidy family groups (Figure S1). However, the difference between PCA and t-SNE with this dataset was less than in other datasets, likely due to the small number of family groups, and thus lower dimensionality, in the data. t-SNE was far better than PCA at clustering individuals into family groups when provided with genotype posterior probabilities or an IBS matrix, however (Figure S3). Interestingly, individuals with many missing genotype calls were erroneously clustered together by both methods. This was probably due to the fact that, since the prior genotype frequency we provided was uniform, all individuals with missing genotypes at a locus were assigned an expected allele count of one for each allele at that locus. If allele frequencies were instead used as priors instead, this issue would likely be mitigated.

*Visualizing multiple levels of population sub-structuring.*

Both methods depicted the same general clustering patterns when run across subsets of the stickleback SNP data (Figure 2). Sexes and populations were split when run on all chromosomes, populations were split when run without the sex chromosomes, and individuals with at least one copy of the linkage group IX chromosomal inversions were clustered separately when only diagnostic SNPs for that allele were run. As before, populations with greater genetic diversity were less tightly clustered by t-SNE but not PCA at all levels. Additionally, PCA primarily split the outgroup population (PAL) from the other samples, which t-SNE spread apart to a greater degree. In general, t-SNE results did not change at different perplexities, although the clusters were less well defined at a very low perplexity.

Interestingly, PCA and t-SNE depicted different patterns when only SNPs diagnostic for the linkage group IX chromosomal inversion were run. While both methods split apart individuals with at least one copy of the inversion, PCA separated homo and heterozygote carriers of the inversion. Additionally, at all perplexities above 2, t-SNE created additional well-defined clusters of individuals homozygous for the typical version of the chromosome. While these clusters were consistent across t-SNE runs, they were not constant across bootstraps of the original data (Figure S4).

Hemstrom, W., van de Wetering, S., & Banks, M. A. (2017). Fish ladder installation across a historic barrier asymmetrically increased conspecific introgressive hybridization between wild winter and summer run steelhead salmon in the Siletz River, Oregon. *Canadian Journal of Fisheries and Aquatic Sciences*. doi:10.1139/cjfas-2016-0411