BIO594 – Summary

Topic: Correlation between phenotype and genotype

February 19, 2019

Kevin Wong

Understanding the correlations between phenotype and genotype, and its associations with fitness is fundamental to describe patterns of speciation and predictions against future climate change conditions. The three studies discussed in this summary, Barret et al. 2019 (BA), Bosse et al. 2017 (BO), and Nadeau et al. 2016 (NA), describe changes in specific genes that correlate with phenotypic variation among populations. With the same overarching goal, each study utilized different methods do describe both genomic and phenotypic variation within their study system.

BA describe changes in the dorsal coats of *Peromyscus maniculatus* in relation to substrate color through a cross transplant field experiment where mice collected from both ancestral light and dark populations were introduced into isolated enclosures with dark or light substrates. The overall experimental design was to: 1) identify SNP variants among the *Agouti* gene, 2) test the functional link between the SNP variants, phenotype (i.e. pigmentation), and fitness (i.e. survival), and 3) understand the mechanisms in which changes in a variant can change the phenotypic trait. The authors described five different physiological (i.e. pigmentation) traits that were not correlated with each other, inferring independent selection. To quantify phenotypic traits, hairs were examined under the microscope and phenolmelanin content was extracted through high-performance liquid chromatography. Using SNP data that included the *Agouti* region and regulatory elements, 2442 high-quality sites were identified in the *Agouti* gene and 53,507 in the genome. Demographic effects were controlled by identifying ~2100 unlinked genome wide regions. 353 and 549 non-random genotype allele frequencies were calculated at each *Agouti* variant site independently for the surviving mice in each light and dark enclosure, respectively. 31 *Agouti* SNPs were identified as candidates associated with dorsal brightness, and were subsequently compared under a model with and without selection. Seven SNPs associated with ΔSer, which is the deletion of serine in an amino acid. Thus ΔSer was tested further to determine its functional link with survival and pigmentation, which was highly correlated with dorsal brightness and had high levels of genetic differentiation. Additionally, ΔSer influenced the production of pheomelanin, which is the primary component in changing hair color, mechanistically determined by quantifying real-time binding interactions between *Agouti* and attraction proteins. By using both molecular and biochemical data, the BA authors were able to describe the functional changes in *P. maniculatus* dorsal coloration attributed in ΔSer in the *Agouti* gene altering pheomelanin production.

BO utilized spatio-temporal data to understand the adaptive evolution of *Parus major* across Europe. Populations from the United Kingdom, Netherlands, and Veluwe were used in this study with 2322 individuals genotyped and 485,122 SNPs identified after filtering. *P. major* across these three regions had low genetic structure, large effective population size, and high levels of inferred gene flow, subsequently making it an excellent study system to determine the potential for evolutionary adaptation. A genome-wide association study (GWAS) was used to identify the loci under divergent selection, which produced outliers that were strongly associated with candidate genes related to skeletal development and morphogenesis. From the extracted gene ontology terms, bill morphology was determined to be the main contributor to differentiation between UK and Dutch populations. By using two complementary approaches (mixture model analysis and sliding window approach), the genetic architecture of bill length in the UK population was described to be highly polygenic and that the gene associated are under divergent selection between populations. The allele associated with bill length, COL4A5-C, was compared to annual reproductive success (i.e. number of fledglings) to determine how natural selection influenced bill morphology. In the UK population, the BO authors predict that COL4A5-C predicts a fitness advantage as it produces larger bills. By using high-resolution spatio-temporal data, this study was able to encapsulate how natural selection is the main driver in bill morphology and how this attributes to population divergence.

NA integrated a fine-scale mapping population genomics approach with a gene expression analyses to describe the functional changes in the gene, *cortex*, which is responsible for altering wing coloration and patterning in *Heliconius* butterflies. 108 SNPs were identified in *H. erato favorinus* to be introns of *cortex* and highly correlated to “yellow bar” patterning on the wing if they were fixed or non-fixed. 15 SNPs were identified in *H. erato demophoon* with the same fixed/non-fixed relationship to wing patterning. However these SNPs did not overlap with the *H. erato favorinus* SNPs, suggesting independent acquisition of this phenotype from separate populations. The alleles at the *Yb* locus are associated with the yellow banding phenotype in *H. melpomene, H. timareta*, and *H. elevates* were inferred to be regulatory variants rather than coding. Therefore, gene expression analysis with specific *Yb* probes for *in situ* hybridization was used to determine differential expression relating to patterning phenotypes and races across developmental stages. The NA authors suggest that the *cortex* is mostly related to wing patterning in multiple races of *Heliconius* butterflies, with the recruitment of this gene being the major constituent to the diversification of Lepidoptera lineages.

All three studies use different methods to describe phenotype-genotype associations, all resulting in different conclusions. The study system used in the BA article has strong environmental selection associations with phenotype; therefore a cross-transplant was necessary to determine the functional changes in phenotype that related to genotype. BO used a dataset with a large spatio-temporal rage, allowing for a GWAS to compare bill morphology over space and time. NA used genomic and gene expression data to isolate the patterning of *Heliconius* butterflies, and made strong inferences relating to the recruitment of the *cortex* gene as a driver for early speciation of Lepidoptera.