SMBE Satellite Meeting on the Genetics of Admixed Populations

May 18-20, 2016
San Antonio, Texas
Program: SMBE Satellite Meeting on the Genetics of Admixed Populations
San Antonio, Texas, May 2016

Wednesday, May 18
8:15-8:45am Check in
8:45-9am Welcome

Session 1: Empirical studies of population histories
9:00-9:35am Bridgett von Holdt, Whole genome sequence analysis shows two endemic species of North American wolf are admixtures of the coyote and gray wolf
9:35-10:10am Torsten Günther, Sampling populations before, during and after migration events – using ancient DNA to study European prehistory

10:10-10:30am Coffee break

10:30-10:50am Alicia Martin, The history of admixture in the 1000 Genomes Project
10:50-11:10am Francesco Montinaro, Complex ancient genetic structure and cultural transitions in southern Africa populations
11:10-11:30am Anna Linderholm, A Novel MC1R allele for black coat colour reveals the Polynesian ancestry and hybridisation patterns of Hawaiian feral pigs

11:30am-1:00pm TXBRI lab/primate center tours (optional)
1:00-2:00pm Lunch

Session 2: Novel methods to untangle admixture histories
2:00-2:35pm Alisa Sedghifar, The Spatial Mixing of Ancestries in Secondary Contact Zones
2:35-3:10pm Simon Gravel, Population structure in African-Americans

3:10-3:30pm Coffee break

3:30-3:50pm Amy Goldberg, Mechanistic models of admixture
3:50-4:10pm Joseph Lachance, Adaptive introgression and the evolutionary genetics of hybrid fitness effects
4:10-4:30pm Shaila Musharoff, Modeling ancestry-dependent phenotypic variance reduces bias and increases power in genetic association studies

4:45-6:15pm Poster session with cocktails & snacks
Odd numbers 5:00-5:30pm
Even numbers 5:30-6pm
Thursday, May 19  
Session 3: Admixture and selection: medical and phenotypic consequences  

9:00-9:35am  
**Esteban Parra**, Genome-wide studies in admixed populations: Challenges and opportunities  

9:35-10:10am  
**Chris Gignoux**, Insights from array design and genotyping of over 50,000 diverse individuals from the Population Architecture using Genomics and Epidemiology Consortium  

10:10-10:30am  
Coffee break  

10:30-10:50am  
**Ricardo Verdugo Salgado**, Genomic identification of recent positive selection in populations from Andean highlands and Southern Chile  

10:50-11:10am  
**Michael Dannemann**, The functional interpretation of Neanderthal introgression into modern humans  

11:10-11:30am  
**Marquitta White**, Whole Genome Sequence Analysis of a Multi-ethnic Pediatric Asthma Cohort Reveals Ethnicity-specific Rare Variants Associated with Bronchodilator Drug Response  

11:30-11:50am  
**Noah Zaitlen**, Partitioning Phenotype-Ancestry Correlations  

12:00-2:00pm  
Ethics Panel, including lunch  
**Andrés Moreno-Estrada, Deborah Bolnick, Uma Ramakrishnan.**  
Moderated by **Shelly Cole.**  

Session 4: Admixture as a mechanism for and against speciation  

2:15-2:50pm  
**Katerina Guschanski**, Taking a comparative approach to study the role of hybridization in speciation  

2:50-3:25pm  
**Laurent Frantz**, Gene flow, speciation and domestication in pigs  

3:25-3:45  
Coffee break  

3:45-4:05pm  
**Joanna Malukiewicz**, Anthropogenic Hybridization: A Driver of Despeciation in Callithrix Marmosets?  

4:05-4:25pm  
**Alexey Yanchukov**, Copy number and high-resolution SNP genotyping across the hybrid zone reveal genome-wide associations beyond the annotated positions of CNV genes  

4:25-4:45pm  
**Emily Puckett**, Phylogeographic analyses of American black bears (Ursus americanus) suggest four glacial refugia and complex patterns of post-glacial admixture  

Conference dinner, shuttles provided at 6:30pm
Friday, May 20

Session 5: Admixture as a dynamic process

9:00-9:35am  Joshua Akey, *Excavating archaic hominin DNA from the genomes of modern humans*

9:35-10:10am  María A. Nieves-Colón, *Migration, admixture and selection in the population history of the Caribbean*

10:10-10:30am  Coffee break

10:30-10:50am  C. Eduardo Guerra Amorim, *Long-distance dispersal suppresses introgression of local alleles during range expansions*

10:50-11:10am  Amanda Pierce, *The role of hybridization in range expansion*

11:10-11:30am  Cesar A. Fortes-Lima, *Unraveling the genome legacy and sex-biased ancestry in African descendants from South America*

11:30-11:40am  Closing remarks
Session 1: Empirical studies of population histories

Whole genome sequence analysis shows two endemic species of North American wolf are admixtures of the coyote and gray wolf
Bridgett vonHoldt, University of California Los Angeles

Protection of populations comprising admixed genomes are a challenge under the Endangered Species Act (ESA), which is regarded as the most powerful species protection legislation ever passed in the US but lacks specific provisions for hybrids. The eastern wolf is a newly recognized wolf-like species that is highly admixed and inhabits the Great Lakes and eastern US, a region previously thought to be included in the geographic range of only the gray wolf. The US Fish and Wildlife Service (USFWS) has argued that the presence of the eastern wolf, rather than the gray wolf, in this area is grounds for removing ESA protection (delisting) of the gray wolf across its geographic range. In contrast, the red wolf from the Southeast US was one of the first species protected under the ESA and was listed for protection under the ESA despite admixture with coyotes. For the first time, we use whole genome sequence data to demonstrate a lack of unique ancestry in eastern and red wolves that would be expected if they represented long divergent North American lineages. These results suggest that arguments for delisting the gray wolf are not valid. Our findings demonstrate how a strict designation of a species under the ESA that does not consider admixture can threaten protection of endangered entities. We argue for a more balanced approach that focuses on the ecological context of admixture and allows for evolutionary processes to potentially restore historical patterns of genetic variation.

Sampling populations before, during and after migration events – using ancient DNA to study European prehistory
Torsten Günther, Uppsala University

Population migrations do not always follow the most parsimonious explanation; even the most likely model inferred using state-of-the-art methodology might not be a true representation of what actually happened. Ancient genomics provides us with the opportunity to sequence individuals from the populations that were involved in the actual migration and admixture events. This allows us to avoid the effects of later migrations, population replacement or drift on a population's gene pool. I will present some case studies from the field of ancient genomics. The examples are mainly focused on the Neolithic transition in Europe which was subject of a long standing question of whether farming practices spread as an idea or a population. Samples from different parts of the continent showed a concordant pattern of European hunter-gatherers and early farmers as two highly differentiated groups who then mixed over the course of at least two millennia. In particular, I am going to show what Neolithic individuals from Scandinavia, Anatolia and Iberia have told us about major population movements across Europe. I will highlight their relationship to modern-day individuals from the same region as well as their connection to population isolates such as Sardinians and Basques.
The history of admixture in the 1000 Genomes Project
Alicia Martin, The Broad Institute

The definition of admixture requires some care, as the notion is historically charged meaning and can vary by researcher. Given the complexity of human history, we expect that all humans can be described as admixed for some choice of source populations. Necessarily this has led to approaches where an analyzed individual is assumed to descend from a mixture of pre-defined discrete populations. However, these methods rely on sufficient source population allelic differentiation and representative sampling strategies. Therefore, results from these simple models need to be viewed with an awareness of the complex demographic events shaping genomic variation.

Here, we explore the ‘admixed’ history of populations in Phase 3 of the 1000 Genomes Project. We compare recent admixture inference in populations from the Americas (≤20 generations, e.g. African American and Hispanic/Latino populations) and methods for inferring demographic history for older admixture events (e.g. in South Asian populations). By leveraging haplotype sharing, linkage disequilibrium decay, and deconvolving ancestry across chromosomes, we also gain fine-scale insight into ancestral origins and sex-biased demography. We quantify previously unpublished results on complex admixtures including East Asian ancestry in some of the Peruvian samples, and substructure within Native American ancestry in the African American samples. Both findings are supported by historic and demographic events, and we discuss the aspects of sampling in the project that also contributed to this complexity. Lastly, we discuss how admixture may impact phenotype architecture, with important implications for how patterns of admixture may be leveraged for disease risk discovery.

Complex ancient genetic structure and cultural transitions in southern Africa populations
Francesco Montinaro, University of Oxford

The characterisation of the structure of southern Africa populations has been the subject of numerous genetic, medical, linguistic, archaeological and anthropological investigations. The subcontinent diversity is the complex result of episodes of genetic admixture and cultural contacts involving the early inhabitants and the migrants which arrived in the region over the last 2,000 years, with some of the variation present in the past being now lost as the result of cultural and demographic assimilation by surrounding populations. Here we analyzed 1,856 individuals from 91 populations, comprising novel and available genotype data to characterise the genetic ancestry profile of Southern African populations. By the combination of Local ancestry and allele frequency analyses we identified a tripartite, ancient, KhoeSan-related genetic structure, which correlates with geography but not with linguistic affiliation or subsistence strategy. The fine mapping of these components in Southern African populations revealed admixture dynamics and episodes of cultural reversion involving several KhoeSan groups and highlighted different mixtures of ancestral components in Bantu speakers and Coloured individuals.
A Novel MC1R allele for black coat colour reveals the Polynesian ancestry and hybridisation patterns of Hawaiian feral pigs
Anna Linderholm, Texas A&M

Pigs (Sus scrofa) have played an important cultural role on Hawaii since Polynesians first introduced them around 1200 AD. Additional lineages of pigs were introduced following Captain Cook’s arrival on Hawaii in 1778 and it has been suggested that the current pig population may descend primarily, or even entirely from European pigs. Although modern populations of feral pigs are an important source of recreational hunting on all of the major islands, they also negatively impact numerous native plants and animals. As a result, understanding the origins of the feral pig population has significant ramifications for discussions concerning both conservation management and cultural continuity and identity on the islands. Here we analyse a neutral mitochondrial marker and a functional nuclear coat colour marker in 57 feral Hawaiian pigs. Through the identification of a new mutation in the MC1R gene that results in black coloration, we demonstrate that, though there is evidence for admixture, Hawaiian feral pigs remain largely descended from those originally introduced by Polynesian settlers.

Session 2: Novel methods to untangle admixture histories

The Spatial Mixing of Ancestries in Secondary Contact Zones
Alisa Sedghifar, Princeton University

Past admixture events leave a population-genomic signature of co-ancestry. These patterns can be used to learn about the timing of admixture, but the results of inference can depend on the assumed demographic model. We present a model of admixture that assumes a geographic component by incorporating diffusive local migration of lineages. Focusing on patterns of linkage disequilibrium, I will provide examples of our model applied in an inference framework to human populations. Inferences under this spatial model can result in substantially different results to the commonly used model of panmictic admixture, highlighting the sensitivity of admixture timing inference to the choice of demographic model. This model can also been extended to describe ancestry block lengths in hybrid zones, and I will present predicted distributions of block lengths and how these may be influenced by selection against hybrid genotypes.
Population structure in African-Americans
Simon Gravel, McGill University

We present a detailed population genetic study of 3 African-American cohorts comprising over 3000 genotyped individuals across US urban and rural communities: two nation-wide longitudinal cohorts, and the 1000 Genomes ASW cohort. Ancestry analysis reveals a uniform breakdown of continental ancestry proportions across regions and urban/rural status, with 79% African, 19% European, and 1.5% Native American/Asian ancestries, with substantial between-individual variation. The Native American ancestry proportion is higher than previous estimates and is maintained after self-identified Hispanics and individuals with substantial inferred Spanish ancestry are removed. This supports direct admixture between Native Americans and African Americans on US territory, and linkage patterns suggest contact early after African-American arrival to the Americas. Local ancestry patterns and variation in ancestry proportions across individuals are broadly consistent with a single African-American population model with early Native American admixture and ongoing European gene flow in the South. The size and broad geographic sampling of our cohorts enable detailed analysis of the geographic and cultural determinants of finer-scale population structure. Recent identity-by-descent analysis reveals fine-scale geographic structure consistent with the routes used during slavery and in the great African-American migrations of the twentieth century: east-to-west migrations in the south, and distinct south-to-north migrations into New England and the Midwest. These migrations follow transit routes available at the time and are in stark contrast with European-American relatedness patterns.

Mechanistic models of admixture
Amy Goldberg, Stanford University

Admixture histories are often complex, with contributions after the founding of the hybrid population, differences in male and females migration rates, or non-random mating within the hybrid population. Statistical models of admixture consider hybrid populations as linear combinations of their source populations, with equal contributions of males and females to the autosomes and a 2/3—1/3 relationship for the X chromosome. Using a set of related mechanistic admixture models that capture complex behavior of admixed populations, we show that the linear combination model is only valid for a simple admixture model unlikely to represent most populations. Under the model, we study the distribution of ancestry on the autosomes and X chromosome within the admixed population as a function of sex-specific contributions and various admixture models, and discuss the interacting effects of sex-biased admixture and ancestry-assortative mating.
Adaptive introgression and the evolutionary genetics of hybrid fitness effects
Joseph Lachance, Georgia Institute of Technology

Ancient introgression appears to be relevant to all global populations, and an important question is how difficult it is for introgressed alleles to persist in human genomes. Here, I extend population genetics theory to include hybrid fitness effects. Additive and epistatic models are considered, and hybrids are allowed to have either increased or decreased fitness. I find that hybrid fitness effects persist over multiple generations, and these effects are dampened over time due to repeated backcrossing. This changes the probability of fixation, and introgressed alleles effectively enter populations with either a head start or a handicap, depending on whether hybrids have increased or decreased fitness. Because of this, classic equations from theoretical population genetics can easily be modified to include hybrid fitness effects – a finding that is verified by extensive computer simulations.

Modeling ancestry-dependent phenotypic variance reduces bias and increases power in genetic association studies
Shaila Musharoff, University of California San Francisco

Many complex human phenotypes vary dramatically in their distributions between populations. Genetic association studies typically use estimates of ancestry, such as principal components (PCs), as fixed-effect covariates to prevent confounding caused by a dependence of phenotypic mean on ancestry. However, the standard approach of including PC covariates in linear regression models assumes that different populations have the same phenotypic variance, which may not hold for recently admixed populations. In this work we consider the possibility that populations with differences in phenotypic mean also have differences in phenotypic variance. First we show this is the typical case under an additive genetic architecture. Then we develop a new likelihood-based method, based on double generalized linear models, to account for relationships between ancestry and phenotypic variance in genetic association studies. In simulations, our test has better power than several linear regression tests that assume equal variance across groups. We observe power increases of 12-66% and obtain unbiased parameter estimates for data simulated with realistic effect sizes and population minor allele frequency differences of 0.45. Furthermore, we show that the current gold standard approach of linear regression with PC covariates can lead to inflation or deflation of p-values for tests of genetic association when population phenotypic variances differ. For example, simulated populations with minor allele frequencies of 0.05 and 0.5 produce test statistics with a lambdaGC value of 1.56, which our method fixes. We further apply our method to several association studies of admixed populations and identify associations between genetic ancestry and phenotypic variance.
Session 3: Admixture and selection: medical and phenotypic consequences

Genome-wide studies in admixed populations: Challenges and opportunities
Esteban Parra, University of Toronto

Populations resulting from recent admixture events between continental groups are a unique resource for genetic studies. This is the case of many populations throughout the Americas, which are primarily the result of admixture between indigenous American, African and European populations, a process that started approximately five centuries ago. Genome-wide studies in recently admixed populations have unique challenges and opportunities. One of the major challenges is the presence of genetic stratification, which can cause false positives in association studies. One of the major opportunities in these populations is the potential to apply methods that are driven by their recent history of admixture, such as the admixture mapping method, which evaluates association of locus ancestry with phenotypic traits. I will provide relevant examples illustrating these challenges and opportunities based on genome-wide studies of diabetes and lipid traits in admixed samples from Mexico City.

Insights from array design and genotyping of over 50,000 admixed individuals from the Population Architecture using Genomics and Epidemiology (PAGE) Consortium
Chris Gignoux, Stanford University

The success of the first decade of genome--wide association studies (GWAS) has been largely predicated on a model of discrete, homogeneous populations. However, as we scale to mega-scale genomics in diverse populations, we need to address the reality where most populations are on a heterogeneous continuum and can represent a mixture of ancestries. Modeling this process requires improved tools and methods for discovery. Here we will describe the development of arrays tailored to diverse populations, and genomic discovery efforts in >50K deeply phenotyped admixed individuals from the Americas in the PAGE consortium, with insights for the next decade of biomedical discovery.

First, we designed the Illumina Infinium Multi-Ethnic Genotype Array (MEGA) range using a novel cross-population tag SNP selection strategy to design a GWAS array with balanced coverage across the globe. This design improves the capture of untyped variation given the mixture of ancestries found in the Americas where we evaluated per-ancestry imputation performance via admixture deconvolution. We genotyped MEGA across >30 populations in the Americas in the PAGE study. We infer an extraordinary breadth of population structure, admixture, and differential relatedness. The large sample size allows us to genetically characterize historically admixed populations as new candidates for admixture-based analysis, including European ancestry in Filipinos and substantial South Asian ancestry in parts of the Caribbean and South America. Finally, we will comment on implications for genomic discovery in the coming deluge of very large-scale multi-ethnic medical studies in research and healthcare settings.
Genomic identification of recent positive selection in populations from Andean highlands and Southern Chile
Ricardo Verdugo Salgado, *University of Chile*

Human migrations into South America reached the southern area of the continent by at least 14,800 BP. Therefore, populations inhabiting different regions in the continent may have developed local adaptations to their environment, adding to the genetic variation among indigenous populations. We evaluated genomic evidence of microevolutionary processes in three Amerindian populations from the Andes highlands in Peru (77 individuals from Puno) and Southern Chile (13 Pehuenche and 9 Huilliche) by genotyping with the Axiom LAT1 platform. We found sizable levels of genetic divergence between these populations. The Fixation Index Fst between the Highlands and Southern Chile was 0.046 and between Pehuenche and Huilliche was 0.025. To identify recent selective sweeps (less than 30,000 BP), we calculated iHS and XP-EHH scores in 7600~ windows (200 kb). We found that windows with the highest scores in these tests (p-value <0.05) are enriched in genes related with biological processes as Vitamin D metabolism in Andean individuals, and MHC class II protein complex in Southern Chile. Among these high scored windows, we highlight a region on chromosome 6 (51.8-52.6 Mb) that showed strong evidence of selection in all the three groups analyzed, as previously reported in other Amerindian populations. This region harbors genes associated with regulation of inflammatory responses and calcium homeostasis, among others. These results suggest that selective pressures on these biological mechanisms may have been present during the recent history of Native South American populations and represent a potential source of genetic variation in admixed individuals from South America.

The functional interpretation of Neanderthal introgression into modern humans
Michael Dannemann, *Max Planck Institute for Evolutionary Anthropology*

Between 1% and 6% of the genomes of present-day non-Africans derive from Neanderthals and Denisovans. A number of studies have now provided evidence that this introgression from the archaic humans contributes to important human phenotypes such as high altitude adaptation, skin and hair physiology and immunity. An emerging pattern is there is a substantial contribution of archaic humans to variation at immune loci in modern humans. To determine the impact of archaic alleles on immunity in modern humans, we have explored the extent to which these alleles contribute to differences in gene expression in immune-related tissues between present-day humans with and without the archaic alleles. We have also analyzed whether those alleles that influence gene expression are associated with particular phenotypes via genome-wide association studies in humans. We find a number of immune loci where archaic ancestry has consequences for modern human phenotypes, and show that some of these have been targets of selection in modern human populations.
Whole Genome Sequence Analysis of a Multi-ethnic Pediatric Asthma Cohort Reveals Ethnicity-specific Rare Variants Associated with Bronchodilator Drug Response
Marquitta White, University of California San Francisco

Asthma is a complex disease characterized by chronic inflammatory of the airways, and is the most common chronic disease of childhood, worldwide. There are ethnic disparities to albuterol bronchodilator drug response (BDR) among children with asthma. Latino and African populations display the highest asthma morality and the lowest drug response. Recent heritability estimates of BDR suggest a sizeable genetic component, however, previous GWAS studies have failed to reveal common variants associated with BDR that account for the majority of the variation in BDR. One hypothesis is that ethnicity-specific rare variants account for a portion of the “hidden heritability” in BDR. We analyzed whole genome sequence (WGS) data to evaluate the impact of rare variant effects on response to asthma therapies in a multi-ethnic pediatric asthma study. We performed rare-variant SNP-set tests using SKAT-O within each racial/ethnic group to identify population specific rare variants associated with drug response. We used a “sliding window” approach to look at the combined effects of rare variants within overlapping 5kb regions (windows) throughout the entire genome. We identified population specific rare variants associated drug response in each racial/ethnic group (Mexican, Puerto Rican, and African American). We performed fine-mapping of the top “windows” in each population to determine the exact set of rare variants responsible for the pharmacogenetic association. This is the first study to interrogate, genome-wide, the effects of rare variants on pharmacogenetic associations among racial/ethnically diverse children with asthma.

Partitioning Phenotype-Ancestry Correlations
Noah Zaitlen, University of California San Francisco

Over 14,000 medical billing codes (ICD-9) are used to classify diseases, injuries, and health encounters. ICD-9 codes are stored in electronic health records (EHR) for billing purposes and have been recently used to define phenotypes. Differences of these phenotypes between individuals will be a result of both genetic and environmental factors. In admixed populations, which contain genetic information from multiple ancestral populations, the extent to which genetics drives phenotypic differences between ancestral populations will induce a correlation between phenotype and genetic ancestry. Genetic ancestry has also been shown to be correlated with environmental covariates and thus it is unclear how much of a given phenotype-ancestry correlation will be driven by genetics. Partitioning the correlation into genetic and environmental components has important implications for precision medicine and global health. In this work, we present a method that partitions the phenotype-ancestry correlation into genetic and environmental components and apply our method to BioMe to investigate the relationship between genetic ancestry and disease risk. The BioMe Biobank at Mt. Sinai is composed of >32,000 multi-ethnic patients, including admixed patients such as African Americans (n=3,705) and Hispanic/Latinos (n=5,104), with phenotype data from EHRs linked to DNA samples. We show analytically how to partition the correlation and show via simulations that our approach provides unbiased estimates of effect-sizes, heritability, and correlation components. In BioMe, we find 16 correlations between ancestry and ICD-9 based phenotypes (permutation p-value<0.05), such as asthma and hypertension, and show that the components of the phenotype-ancestry correlations vary between them.
Session 4: Admixture as a mechanism for and against speciation

Taking a comparative approach to study the role of hybridization in speciation
Katerina Guschanski, Uppsala University

Hybridization has traditionally been considered an evolutionary dead-end. This view is being increasingly challenged, as evidence for the importance of hybridization in speciation accumulates, mostly from studies of ongoing gene flow between incipient species. However, because the outcome of hybridization may be manifold, gauging the end-result of hybridization from ongoing processes is problematic. One solution is to compare related species groups that differ in species richness and to study instances of past gene flow, which can provide valuable insights into the role of hybridization in generating species diversity. The species-poor African great apes and the species-rich guenons (tribe Cercopithecini) provide an ideal study system. Both groups diverged within the last 10 million years in Sub-Saharan Africa, ruling out differences in evolutionary time and climatic conditions. Rich genomic data exist for all African great ape species and subspecies. Guenons, on the other hand, the most diverse group of African primates with up to 80 recognized taxa, have been little studied genetically. Inter-specific hybridization is unknown in great apes and genomic data suggest low levels of historical secondary gene flow between subspecies. In contrast, guenons hybridize in captivity and the wild, producing fertile offspring between species that diverged millions of years ago and differ cytogenetically. This suggests that past hybridization was likely in guenons. Establishing genomic resources for this species group will provide the means to study ancient gene flow and explore the importance of hybridization and admixture, chromosome evolution, sexual selection, population size and other life history traits on speciation.

Gene flow, speciation and domestication in pigs
Laurent Frantz, University of Oxford

Phenotypic differentiation can arise due to the cessation of gene flow between two populations. However, recent genomic studies have shown that such model often fails to capture the complexity of modern genomic data sets. Instead these studies have demonstrated that speciation or domestication can occur even when gene flow persists. The persistence of genetic exchange during domestication and speciation raises question regarding the mechanisms that maintains phenotypic differentiation. On the other hand, genetic introgression, between phenotypically diverged populations, can also be seen as beneficial, providing the mean for rapid adaptation. Here I will discuss two examples taken from recent studies on the genomics of Sus (pigs and related species). I will show how natural (during speciation) and artificial selection (during domestication) have counteracted the exchange of key structural and single nucleotide genetic variants between diverging populations, allowing for the maintenance of strong phenotypic differentiation. Lastly, I will also provide a few examples in which gene flow between divergent domestic and wild populations may have facilitated adaptations to novel environments.
Anthropogenic Hybridization: A Driver of Despeciation in *Callithrix* Marmosets?
Joanna Malukiewicz, Federal University of Viçosa

Brazilian *Callithrix* marmosets are emerging as an important primate hybridization model. Natural hybridization occurs at contact points between the six historically allopatric *Callithrix* species. However, a smorgasbord of anthropogenic *Callithrix* hybrid populations is now widespread throughout Brazil. These hybrids are the result of interbreeding between introduced *Callithrix* species, particularly those of *C. penicillata* and *C. jacchus*, with one another or native species. Both natural and anthropogenic hybrids serve as evolutionary laboratories that better inform us on hybridization’s role in (de)speciation. Through online surveys, interviews, and site visits we have collected a large number of reports of on-going anthropogenic *Callithrix* hybridization in SW Brazil. We are examining the genetic and phenotypic consequences of both anthropogenic marmoset hybridization, including that of *C. aurita* x *C. jacchus/C. penicillata*, *C. penicillata* x *C. geoffroyi*, *C. penicillata* x *C. jacchus* and natural hybridization (e.g. *C. penicillata* x *C. jacchus* and *C. flaviceps* x *C. geoffroyi*). In contrast to natural hybridization, anthropogenic hybridization induces patterns of localized collapse of both native and introduced *Callithrix* species populations with replacement by hybrid swarms. Phenotypic variation of pelage and morphometric features of the anthropogenic *Callithrix* hybrids examined thus far exceeds that of parental species, and also features striking, novel phenotypes. Notably, these anthropogenic *Callithrix* hybrids may be experiencing a loosening of evolutionary constraints present in the parental species. Given the widespread occurrence of anthropogenic *Callithrix* hybridization, our data suggest that this type of hybridization may be driving localized instances of despeciation of the *Callithrix* genus.

Copy number and high-resolution SNP genotyping across the hybrid zone reveal genomewide associations beyond the annotated positions of CNV genes
Alexey Yanchukov, Bülent Ecevit University

The hybrid zone between two subspecies of the house mouse (Mus musculus musculus and Mus m. domesticus) is a unique mammalian model system to study fine-scale interactions of recently diverged genomes. In the first ever (to our knowledge) study of gene Copy Number Variation (CNV) across a hybrid zone, we assayed, using digital PCR, copy numbers of seven protein-coding genes that showed extensive CNV both within the hybrid zone and out in the allopatric ranges of the two subspecies. These copy number data were tested for association with either musculus or domesticus ancestry state of ~ 1/3 million SNPs covering the entire genome. The copy numbers of five genes correlated broadly with hybrid index, one locus showed no significant correlation at all, and one was highly significant only within a narrow region of introgression from musculus into domesticus. In three cases the highest and/or outlying levels of association were observed at or close to the position of the gene amplicon as annotated in the reference, demonstrating a surprising power of our approach in mapping the reference locations of copy number variants. Several other reference locations were recognized as positive outliers in the association with particular CNV genes: de novo assembly in these regions could be used to test the presence of these putative ‘distant’ copies.
Phylogeographic analyses of American black bears (Ursus americanus) suggest four glacial refugia and complex patterns of post-glacial admixture
Emily Puckett, Fordham University

Studies of species with continental distributions continue to identify intraspecific lineages despite continuous habitat. Lineages may form due to isolation by distance, adaptation, divergence across barriers, or genetic drift following range expansion. We investigated lineage diversification and admixture within American black bears (Ursus americanus) across their range using 22k SNPs and mitochondrial DNA sequences. We identified three subcontinental nuclear clusters which we further divided into nine geographic regions: Alaskan (Alaska-East), eastern (Central Interior Highlands, Great Lakes, Northeast, Southeast), and western (Alaska-West, West, Pacific Coast, Southwest). We estimated that the western cluster diverged 67 kya, before eastern and Alaskan divergence 31 kya; these divergence dates contrasted with those from the mitochondrial genome where clades A and B diverged 1.07 Mya, and clades A-east and A-west diverged 169 kya. We combined estimates of divergence timing with hindcast species distribution models to infer glacial refugia for the species in Beringia, Pacific Northwest, Southwest, and Southeast. Our results show a complex arrangement of admixture due to expansion out of multiple refugia. The delineation of the genomic population clusters was inconsistent with the ranges for 16 previously described subspecies. Ranges for U. a. pugnax and U. a. cinnamomum were concordant with admixed clusters, calling into question how to order taxa below the species level. Additionally, our finding that U. a. floridanus has not diverged from U. a. americanus also suggests that morphology and genetics should be reanalyzed to assess taxonomic designations relevant to the conservation management of the species.

Session 5: Admixture as a dynamic process

Excavating archaic hominin DNA from the genomes of modern humans
Joshua Akey, University of Washington

Anatomically modern humans overlapped and mated with archaic hominins including Neandertals and Denisovans. We have developed statistical methods to identify surviving archaic hominin DNA in the genomes of modern humans, and applied these methods to whole-genome sequences from 1,523 geographically diverse individuals, including 35 new Island Melanesian genomes. In aggregate, we recovered 1.34 Gb and 303 Mb of the Neandertal and Denisovan genome, respectively. We leverage these maps of archaic sequence to show that Neandertal admixture occurred multiple times in different non-African populations, characterize genomic regions that are significantly depleted of archaic sequence, and identify signatures of adaptive introgression. These data provide new insights in hominin evolutionary history and genomic regions that may harbor substrates of uniquely modern human phenotypes.
Migration, admixture and selection in the population history of the Caribbean
María A. Nieves-Colón, Arizona State University

This research examines how migration and admixture shaped the genetic diversity of ancient and modern Caribbean populations. These topics are addressed through two approaches: an ancient DNA analysis of individuals from pre-contact archaeological sites in Puerto Rico, and an analysis of whole genome variants from modern Caribbean islanders. Our ancient DNA analysis focuses on tracing the origin and number of pre-contact migrations to Puerto Rico, as well as testing archaeological hypotheses regarding the admixed origins of indigenous Caribbean groups. Ancient DNA is obtained from a large sample of human skeletal remains, dated between A.D. 500–1300, from three archaeological sites: Tibes, Paso del Indio and Punta Candelero. Genotypes generated from the remains are then compared, through unsupervised and model based approaches, with genetic variation from modern and ancient Native American populations. Preliminary data from ancient complete mitochondrial genomes supports previous findings from ancestry research in modern, admixed Puerto Ricans, and suggest a complex and continuous admixture history for pre-contact island communities. We also explore how indigenous populations contributed to the genetic diversity of modern Caribbean islanders. We especially focus on using autosomal SNP data from admixed individuals to examine how adaptive processes may have affected distributions of Native American, African and Eurasian ancestry in these groups. Here we present preliminary results from these two investigations. We discuss the difficulties of detecting ancient admixture with tropical ancient DNA and illustrate how evolutionary forces like gene flow and selection have shape genetic variation in past and present Caribbean groups.

Long-distance dispersal suppresses introgression of local alleles during range expansions
C. Eduardo Guerra Amorim, Columbia University

During range expansions, even low levels of interbreeding can lead to massive introgression of local alleles into an invaders genome. Nonetheless, this pattern is not always observed in human populations. For instance, European Americans in North America are barely introgressed by Amerindian genes in spite of known contact and admixture. With coalescent spatially explicit simulations we examine the impact of long distance dispersal (LDD) events on introgression of local alleles into the invading population using a set of different demographic scenarios applicable to a diverse range of natural populations and species. More specifically, we consider two distinct LDD models: one where LDD events originate in the range core and targets only the expansion front and a second one where LDD events can occur from any area to any other. We find that LDD generally prevents introgression, but that LDD events specifically targeting the expansion front are most efficient in suppressing introgression. This is likely due to the fact that LDD allows for the presence of a larger number of invader alleles at the wave front, where effective population size is thus increased and local introgressed alleles are rapidly outnumbered. We postulate that the documented settlement of pioneers directly on the wave front in North America has contributed to low levels of Amerindian admixture observed in Northern Americans of European ancestry and that this phenomenon may well explain the lack of introgression after a range expansion in natural populations without the need to evoke other mechanisms such as natural selection.
The role of hybridization in range expansion
Amanda Pierce, University of North Carolina

The evolution of species ranges is fundamental to understanding how biodiversity is distributed and maintained; however, we still do not fully know how species geographic ranges evolve and what factors fuel range expansions when they do occur. To colonize a new location, a species must either expand to areas in which it is pre-adapted, or it must quickly adapt to a novel habitat. One mechanism for adaptation involves hybridizing with a species native to the novel area. We investigate this using the Plains spadefoot toad, Spea bombifrons. S. bombifrons is thought to be ancestral to the central plains region of the United States and is believed to have expanded its range northward into a similar grassland habitat following the most recent glacial retreat. Interestingly, S. bombifrons has also expanded its range southward into a novel desert habitat and this expansion appears to have increased within the last 200 years and is still ongoing. It is hypothesized that S. bombifrons has hybridized with a closely related native toad species (S. multiplicata) that is already adapted to desert conditions. Using microsatellite markers, we investigated the route of the range expansion of S. bombifrons. We also performed population genetic analyses to determine the degree of population structure and gene flow occurring throughout the range. Finally, we determined the degree to which hybridization has played a role in the range expansion of S. bombifrons. This research gives insight into the evolution of species range as well as the potential role of hybridization in adaptation.

Unraveling the genome legacy and sex-biased ancestry in African descendants from South America
Cesar A. Fortes-Lima, University Paul Sabatier

The transatlantic slave trade enacted the most traumatic long-distance migration in human history. It changed the genetic composition of the entire New World, for instance, in French Guiana and Suriname most of enslaved Africans escaped and established new independent African settlements since then, called Noir Marron. To obtain a fine-scale genetic perspective of African descendants, we analyzed whole mitochondrial genomes (mtDNA), Y chromosome markers (47 Y-STRs and 96 Y-SNPs), and genome-wide data for the Illumina HumanOmni5 BeadChip (4.5 million variants) from six African-American populations from South America (French Guiana, Brazil, and Colombia), and six putative West-African ancestors. To uncover admixture patterns, we compared them with continental reference population included in the 1000 Genomes Project Phase 3. The Afro-Brazilian and the Afro-Colombian population presented different demographic migration models and admixture timing associated to their different colonial pasts. MtDNA genome, Y chromosome and X chromosome uncovered sexual asymmetric admixture patterns in admixed African-American populations, consistent with an excess of European male contribution and elevate African female ancestry. Besides, all genetic systems revealed strong African ancestry (above 98%) in Noir Marron communities. We also applied a historical framework to provide new insight into haplotype sharing between African-Americans and African geographic groups across the continent. In Noir Marron, both global and local ancestry inferences highlighted remarkably genome-wide ancestry linked to the populations residing today in the historical Bight of Benin region. Therefore, this study sheds new light on complex admixture events in African descendants, and reconstructs broken African ancestral links throughout the Atlantic world.
1 The genetic structure of Mexicans of African descent
Maria C. Ávila-Arcos¹,², Andres Moreno-Estrada³, Karla Sandoval-Mendoza³, Marcus W Feldman¹,⁴, Carlos Bustamante¹
¹Center for Computational, Evolutionary and Human Genomics, Department of Genetics, Stanford University
²International Laboratory for Human Genome Research, National Autonomous University of Mexico
³Laboratorio Nacional de Genómica para la Biodiversidad, Centro de Investigación y Estudios Avanzados, Instituto Politécnico Nacional, México
⁴Department of Biological Sciences, Stanford University

Between the 16th and 19th centuries, over 12 million people were kidnapped mainly in West and West Central Africa and transported to the Americas as a result of the trans Atlantic slave trade. The way enslaved Africans adapted to their imposed new places of residency affected how they established their identities, their social networks, their communities, and, ultimately, how the interactions of the Afro-descendant population with resident peoples contributed to the current diversity and cultural identity of today’s Countries in the Americas. This major demographic shift has been studied from a genetics perspective focusing mostly on ancestry of African Americans, and to a lesser extent on Caribbean populations. Despite Mexico having received 200,000 Africans during the slave trade, no study has focused on the study of African genetic ancestry and its implications in Mexico. In this study we worked together with Afro-Mexican communities living in the coastal regions of the States of Oaxaca, Guerrero and Veracruz to characterize their genetic ancestry. Using genome-wide genotype data we analyzed the African genetic ancestry that exists in these groups to complement our knowledge about the dynamics of the trans Atlantic slave trade to Mexico. Importantly, as the current situation of Afro-Mexicans, is one of poverty, marginalization and lack of official recognition as a vulnerable minority, this study should contributes to their appreciation as part of Mexico’s mosaic of diversity and is a positive step towards their recognition.

2 Painting by evolutionary history: Inference of local ancestry in admixed genomes
Ali Berens, Joseph Lachance
Georgia Institute of Technology

Local ancestry inference is important for understanding demographic history and predicting hereditary disease risks in admixed populations, which are increasingly becoming the norm as human migration around the world accelerates. Almost all current methods require reference sequences from ancestral (or at least closely related proxy) populations. Unfortunately, some populations, including Native Americans, are underrepresented in the publically available global population-level genomic datasets. For the few methods that are capable of reference-free local ancestry inference, information about the historical relationship between these ancestral populations is lost. We propose a new chromosome painting method that uses the sequence identity matrix of all samples to infer local ancestry without the need for reference sequences. From whole genome sequence data of admixed individuals, ancestral populations are formed top down iteratively by splitting a single population into two daughter populations. Following each split, local haplotypes are repainted by ancestry to include these new daughter populations. This hierarchal approach of painting chromosomes captures evolutionary relationships between ancestral populations.
3 Estimating Archaic Admixture in Modern Humans
Ryan Bohlender, Alan Rogers
University of Utah

Recent advances in the study of archaic hominin DNA have resulted in the rapid development and application of new methods designed to test our relationship to our nearest relatives. These methods have been applied with archaic samples in contexts with ghost admixture, sparse sampling, ascertainment bias, and poorly understood historical events. They have also been applied to modern samples with complex relationships which will exacerbate the same problems. Here, we introduce a new method for estimating the admixture fraction, and test it and several previous methods to determine their sensitivity to the above problems. Finally, we apply these methods to Single Nucleotide Polymorphism (SNP) microarray and whole genome data, and compare our estimates to those published previously.

4 Weasling around - Molecular evolution and gene-flow among diverging North American ermine (Mustela erminea)
Jocelyn Colella1, Tina Lan2, Charlotte Lindqvist2, Sandra Talbot3, Joseph Cook1
1University of New Mexico
2SUNY Buffalo
3US Geological Survey

Secondary contact zones provide an opportunity to investigate introgression, reinforcement and divergence. This study examines genetic variation (amplicon and whole-genome sequencing) to understand the evolutionary histories and contemporary consequences of climate-induced range shifts on North American Ermine (Mustela erminea). Ermine - the most widespread terrestrial carnivore in the arctic - are represented in North America by four genetically and morphologically distinct clades. Extended isolation in glacial refugia (last glacial maximum; 21kya), followed by post-glacial range expansion has led to secondary contact and hybridization at multiple locations within this Holarctic species’ range, providing independent tests of the consequences of hybridization on both divergence and diversity. First, we use complete mitochondrial genomes (>16,600bp; Illumina MiSeq) and a suite of nuclear amplicons (8 independent loci) to generate robust phylogenies, refine distributional boundaries, quantify inter-populational divergence, and detect signatures of hybridization. Second, preliminary analysis of whole-genome sequences (N=10; 3.8gigabases, Illumina HiSeq) corroborate our historical understanding of demography, but provide new perspective on the timing and extent of divergence and hybridization in this system. A genomic perspective on speciation, among divergent clades facilitates investigation of semi-permeable species boundaries, isolating mechanisms, and hypothesized ‘speciation genes’. Genomic investigations into non-model systems builds understanding of how introgression impacts the genome, enables delimitation of genomic commonalities across species, and can highlights diagnostic, non-introgressed genomic regions for use in downstream diagnostics and management.
5 Admixture mapping of TB in a complex admixture scenario
Michelle Daya¹, Lize van der Merwe¹, Christopher R Gignoux², Paul D van Helden¹, Marlo Möller¹ and Eileen G Hoal¹
¹Stellenbosch University
²Stanford University

The admixed South African Coloured population is ideally suited to the discovery of tuberculosis susceptibility genetic variants and their probable ethnic origins, but previous attempts at finding such variants using genome-wide admixture mapping were hampered by the inaccuracy of local ancestry inference. In this study, we infer local ancestry using the novel algorithm implemented in RFMix, with the emphasis on identifying regions of excess San or Bantu ancestry, which we hypothesize may harbour TB susceptibility genes.

Using simulated data, we demonstrate reasonable accuracy of local ancestry inference by RFMix, with a tendency towards miss-calling San ancestry as Bantu. Regions with either excess San ancestry or excess African (San or Bantu) ancestry are less likely to be affected by this bias, and we therefore proceeded to identify such regions, found in cases but not in controls (642 cases and 91 controls). A number of promising regions were found (overall p-values of 7.19×10⁻⁵ for San ancestry and less than 2.00×10⁻¹⁶ for African ancestry), including chromosomes 15q15 and 17q22, which are close to genomic regions previously implicated in TB. Promising immune-related susceptibility genes such as the GADD45A, OSM and B7-H5 genes are also harboured in the identified regions. In conclusion, admixture mapping is feasible in the South African Coloured population and a number of novel TB susceptibility genomic regions were uncovered.

6 Interspecific genome admixture in Old World mice
Eslam Elshahat, HM Kohn
Rice University

The house mouse group (Mus musculus) is commonly used animal model in biomedical and evolution research. Genome wide data is valuable to understand the importance of intra- and inter- population admixture and genome evolution in mammals. Inter-species genome admixture importance emerged from a recent case of adaptive introgressive hybridization between a subspecies of the house mouse ssp (Mus musculus domesticus) and the Algerian mouse (Mus spretus). A 10 mega-bases region of chromosome 7 was moved from M. spretus to M. m. domesticus, including anticoagulants poisons resistance gene Vkorcl. In this project we use Mouse Diversity Arrays (MDA) of 7 M. spretus, covering the geographical range of the species, were used to detect introgression in a dataset of 60 M. m. domesticus samples. We used PhyloNet–HMM a recently developed Bayesian statistical method to detect introgression. We report how widespread admixture is across the mouse genome, and across geographical space and evolutionary time. We examine the directionality of admixture, test for functional importance of introgressed and non-introgressed admixed genomic regions. Intra-specific Admixture is known process affecting genomic diversity of species, but genome diversity studies should consider the role of inter-specific gene flow also.
Blood lipid levels, including low-density lipoprotein (LDL), high-density lipoprotein (HDL), total cholesterol (TC), and triglycerides (TG) are important factors in determining risk for cardiovascular disease (CVD). These traits are also known to be highly heritable and show differences between ethnic groups. Here, admixture mapping was used in the African American cohort of the Family Heart Study (FamHS) to identify genetic regions that are associated with LDL, HDL, TC, and TG in African Americans. Our experiment resulted in genome-wide significant associations on chromosomes 2, 4 and 6 with TG. We also show one genome-wide significant association with TC on chromosome 19. To verify the results seen in FamHS, we sought replication of our findings using the African American cohort of the HyperGen Network. Currently, we are able to show replication for the association result on chromosome 6 with TG. The TG-associated region on chromosome 6 and the TC-associated region on chromosome 10 both contain loci involved in lipid biology. The locus on chromosome 6 contains HULC, which encodes a long non-coding RNA that is associated with the deregulation of lipid metabolism and the locus on chromosome 19 includes CAPN12, which is a member of a gene family whose members are known to influence cholesterol levels. Although these results are preliminary, our data suggest that admixture mapping can be a useful hypothesis-generating tool to identify genomic regions that contribute to complex diseases, like CVD.

Mexicans are a recent admixture of Amerindians, Europeans, and Africans. We performed local ancestry analysis of Mexican samples from two genome-wide association studies obtained from dbGaP, and discovered that at the MHC region Mexicans have excessive African ancestral alleles compared to the rest of the genome, which is the hallmark of recent selection for admixed samples. The estimated selection coefficients are 0.05 and 0.07 for two datasets, which put our finding among the strongest known selections observed in humans, namely, lactase selection in northern Europeans and sickle-cell trait in Africans. Using inaccurate Amerindian training samples was a major concern for the credibility of previously reported selection signals in Latinos. Taking advantage of the flexibility of our statistical model, we devised a model fitting technique that can learn Amerindian ancestral haplotype from the admixed samples, which allows us to infer local ancestries for Mexicans using only European and African training samples. The strong selection signal at the MHC remains without Amerindian training samples. Finally, we note that medical history studies suggest such a strong selection at MHC is plausible in Mexicans.
9 Variation in the DNA methylation of skeletal tissues in a population of pedigreed baboons
Genevieve Housman¹, Lorena M. Havill², Ellen Quillen², Anne C. Stone¹
¹Arizona State University
²Texas Biomedical Research Institute

DNA methylation is an epigenetic mechanism that regulates gene expression and may contribute to variations in complex skeletal traits. Characterizing DNA methylation patterns in the skeletal tissues of model organisms is a necessary first step towards understanding the role of this mechanism in skeletal development and maintenance. For this study, a captive colony of baboons was used because their known pedigree allowed for the inclusion of genetic relatedness in analyses. Genetic relatedness is important to consider because relatedness effects variation in underlying DNA sequences, and this, in turn, influences resulting DNA methylation patterns. We investigated the relationship between skeletal tissue DNA methylation variation and skeletal phenotype variation in baboons. Skeletal tissue DNA methylation patterns were assessed in right distal femur trabecular bone from adult female baboons (n=12) using the Illumina HumanMethylation 450K BeadChip. Skeletal phenotypes included twenty-nine linear measurements of the femur and the individual’s osteoarthritis disease state (either no or severe). Several loci were significantly differentially methylated across phenotypes. Specifically, out of over 450,000 positions, approximately 2.06% were differentially methylated between those with no and those with severe osteoarthritis; 5.67% across femoral anatomical neck heights, 5.75% across femoral anatomical neck depths, and 8.53% across femoral medial condyle widths. These measurements produced the highest number of differentially methylated loci. These findings provide evidence of DNA methylation variation in skeletal tissue from one non-human primate species. They also provide insight into the degree to which this epigenetic variation relates to variation in skeletal phenotypes.

10 The history of recent admixture events in Island Southeast Asia
Georgi Hudjashov¹, Hie Lim Kim², Tatiana Karafet³, Herawati Sudoyo⁴, J. Stephen Lansing⁵, Michael Hammer³, Murray P. Cox¹
¹Massey University
²Nanyang Technological University
³University of Arizona
⁴Eijkman Institute for Molecular Biology
⁵Nanyang Technological University

Human populations in Island Southeast Asia and Indonesia have a particularly complex evolutionary history. This region is a melting pot of multiple prehistoric admixture episodes. The genetic history of these events has been largely described using haploid loci and generally confined to the expansion of Austronesian speakers out of Taiwan. Only a limited number of studies have exploited patterns of variation within autosomal loci and questions of post-Austronesian admixture remain largely unaddressed.

Here we report new genome-wide genotyping data from a large number of Indonesian and Melanesian populations, and we explore post-Neolithic genetic differentiation and population history in the region. Three major layers of ancestry are observed across Indonesia – Melanesian, mainland East Asian and Austronesian. The proportions of different components vary between western and eastern parts of the archipelago, approximately demarcated by Wallace’s biogeographic line. We use the LD-based GLOBETROTTER technique to explore fine-scale patterns of admixture in Indonesia and detect traces of a major mixture event between
western and eastern Indonesian sources in eastern Indonesia approximately 2000 years ago. We hypothesize that this West to East migration may have been triggered by an increase in human population size and search for the new cultivation landscapes associated with the adoption of irrigated rice horticulture in western Indonesia. In addition, Sumatra and other western Indonesian islands experienced gene flow from South Asia around 1000 years ago, in a process which is probably related to the Indian maritime trade network.

11 Genomic signatures of diversification in woodland eucalypt species
Jasmine Janes\textsuperscript{1}, Rebecca Jordan\textsuperscript{2}, Justin O Borevitz\textsuperscript{3}, Rose L Andrew\textsuperscript{1}
\textsuperscript{1}The University of New England
\textsuperscript{2}University of Melbourne
\textsuperscript{3}Australia National University

The importance of ancient standing variation and introgressed alleles in recently diverged species is becoming an increasingly popular area of research. Teasing apart the patterns of historical versus contemporary gene flow, and new mutations, is fundamental to our understanding of adaptive and evolutionary processes. For example, the source of raw genetic material for evolution can have significant impacts on the speed and success with which a population can adapt, and on the genomic signatures resulting from selection. Our work addresses the importance of introgression in the diversification of an iconic Australian genus, Eucalyptus. Using whole-genome shotgun sequencing, we will identify genomic regions corresponding to historical introgression and adaptive divergence. Quantifying these genomic signatures will facilitate improved species resolution in this challenging group.

12 How gene flow has shaped morphological, genetic, and behavior aspects of a wild nonhuman primate hybrid system
Mary Kelaita
St. Philip’s College

In 2007, hybridization among two distinct species of howler monkeys was confirmed for the first time using genetic methods. With individual genetic ancestry known, the growing area of research on the topic has made morphological, ecological, and behavioral comparisons of the parent species with their hybrids to address questions about the extent of gene flow and its affect on individual fitness and the maintenance of species boundaries. Comparisons with other putative howler monkey hybrid zones has shed further light on the possible cellular and/or molecular mechanisms responsible for the degree of hybridization observed. Lastly, the Mexican howler monkey hybrid zone has allowed researchers to investigate how heritable certain traits are (behavior and vocalizations) using data from hybrid individuals and purebreds while being able to control for ecological variation. I review these findings in addition to analyses I conducted on morphological data to highlight the significance of hybridization in shaping primate genetic diversity.
13 Genomic Characterization of Nahuatl-Speaking Populations in Mexico
Erika Landa, Karla Sandoval-Mendoza, Andrés Moreno-Estrada, Rafael Montiel-Duarte
LANGEBIO

Nahua people are the most widespread and numerous ethnic group in Mexico with 1.5 million speakers and nowadays they are culturally very heterogeneous in comparison with other groups. In this project we analyzed ~60k SNPs as a product of merging three different genotyping platforms, in which we obtained a set of 271 individuals belonging to seven different ethnic groups. A Yoruba (YRI), Northern and Western European Ancestry individuals (CEU) dataset was included to determine admixture within Native populations. We compared genomic profiles of historical and linguistically related indigenous groups to Nahua people. Additionally we could detect a contribution of European populations to Native Mexican genomes, it was more evident in Nahua and Purepecha populations. We continue working to infer migrations between our seven ethnic groups, compare the diversity between these ethnic groups, and among Nahua populations. Can genomic data provide strong evidence to infer if their cultural and genetic diversity is due to admixture with other Native Mexican groups?

14 Genomic analysis of hybridization in schistosome parasites
Winka Le Clec’h1, Frederic Chevalier1, Marina McDew-White1, Bonnie Webster2, David Rollinson2, Tim JC Anderson1
1Texas Biomedical Research Institute
2Natural History Museum, London, United Kingdom

There are 25 recognized species of blood flukes in the genus Schistosoma. These parasites are diploid, with separate sexes, a ZW sex determination system, and ~360-385 Mb genomes (2n=16). Hybridization is particularly common in nature among eight closely related schistosome species (the “haematobium” group) that infect man and domestic animals. Many of these schistosome species can also be maintained in the laboratory, and interspecies laboratory crosses result in viable progeny. These parasites are experimentally tractable and biomedically important but several problems must be overcome for effective genomic analysis of hybridization in this system: (i) as adult parasites live in the human blood vessels, only larval stages collected from human feces or urine are available for analysis – we have therefore developed a robust exome capture approach to sequence the ~15 Mb exome from microscopic miracidia larvae following whole genome amplification (ii) the S. haematobium genome assembly consists of 29,834 scaffolds, which must be assembled to define introgressed regions of the genome - we will improve the genome assembly by conducting laboratory genetic crosses, resequencing the F2 progeny and reassembling the genome using linkage information, (iii) interspecific genetic crosses are complicated by the lack of sex specific markers - we are developing quantitative PCR approaches to identify parasite gender using copy number information from Z-linked regions and W-specific markers. Our long-term goals are to characterize introgression where multiple species live in sympatry, and to determine the genetic basis of species specific parasite traits through controlled interspecific genetic crosses.
15 **pcadapt: an R package for genome scans based on Principal Components Analysis**
Keurcien Luu¹, Michael Blum², Eric Bazin¹
¹CNRS, Grenoble
²CNRS, Paris

We present the R package pcadapt that performs genome scan of divergent selection based on principal component analysis. Compared to common approaches based on measures of differentiation (Fst) between populations, the method based on PCA works at the scale of individuals. Because it is an individual-based approach, pcadapt can be applied even when there is admixture in the sample. Using data simulated with an island model and a model of population divergence, we show that the PCA-based approach is more powerful than Fst-based approaches when sampled individuals result from recent admixture. We also discuss how genome scans based on PCA can be applied to detect adaptive introgression with large population genomic data.

16 **Leveraging whole genome sequencing data in racially diverse children to advance precision medicine for asthma therapies**
University of California San Francisco

Asthma, the most common chronic disease of childhood, worldwide, is a complex disorder characterized by chronic inflammation of the airways. There are significant racial/ethnic disparities in response to asthma therapies. Latino and African populations display the highest asthma death rates but the lowest drug response. The goal of President Obama’s Precision Medicine Initiative is to improve care for all patients by engaging in research efforts involving more diverse study populations. To this end, we analyzed whole genome sequencing (WGS) data from three racial/ethnic groups of children with asthma and drug response. We leveraged three complementary analytical techniques (common and rare variant analyses, and admixture mapping) to identify novel pharmacogenetic associations. Common variant analyses identified several ethnic-specific loci associated with drug response. In contrast, a trans-ethnic meta-analysis across populations revealed only one locus. Rare variant burden tests revealed an abundance of ethnic-specific effects. A meta-analysis of rare-variant effects did not reveal any universal associations; strengthening the argument that rare variant effects are more likely to be ethnic-specific. Admixture mapping identified several ancestry-specific loci associated with drug response. Our results highlight the benefits of including multi-ethnic populations to advance our understanding of the racial/ethnic variation in drug response to asthma therapies among multi-ethnic pediatric populations.

17 **Big Genomic Data and Ancestry**
Tes Mersha
University of Cincinnati

Given the huge amount of single nucleotide polymorphism (SNP) data available from high-throughput sources such as 1000 Genomes Project, data mining using various methods is a reasonable approach to identify SNPs that are informative for genetic ancestry. We extensively investigated the distribution and density of the SNPs across the genome of African and European
populations available within the framework 1000 Genomes Project database to prioritize potential candidate SNPs useful for ancestry mapping in an admixed population. About 88 million SNPs were compared between Africans and Europeans using various measures of ancestry informativeness in use today viz. absolute allele frequency differences (δ), Fisher Information Content (FIC), Shannon Information Content (SIC), informativeness for assignment (I), and Fixation Index (FST). Each method provides different sets of candidate ancestry informative markers (AIMs). The selected SNPs represent valuable resources for both controlling population structure and gene mapping studies. The overlap and non-overlap between selected AIMs by different measures of informativeness, data filtering strategies and the accuracy of each measure in classifying ancestral populations are discussed.

18 Fine-scale genetic structure of Chilean populations
Javier Blanco¹, Andrés Moreno Estrada¹, Julian Homburger², Christopher Gignoux², Esteban G Burchard³, Ricardo Verdugo⁴, Mauricio Moraga¹
¹LANGEBIO
²Stanford University
³University of California San Francisco
⁴Universidad de Chile

Chile covers the largest area of the Andean South American region and is home to a multiethnic population formed by the mixing of peoples from many different cultural and ethnic backgrounds. Previous genetic studies in the country have explored either classical or uniparental markers, but fine-scale patterns of human genome-wide variation remain largely uncharacterized. We use genome-wide SNP data from over 450 Chilean admixed individuals to explore the population structure and demographic history of Chile. Combining these data with population reference panels from Africa, Asia, Europe and the Americas, we perform global ancestry analysis and infer the subcontinental origin of European and Native American components of the admixed individuals. Similar to other Latin American populations, our ancestry-specific PCA analyses show that most of the European ancestry in Chilean Latinos comes from the Iberian Peninsula. We also find a strong gradient in the Native American component of Chilean Latinos that correlates with geography. In agreement with previous studies, our ancestry tract length analysis reveals that the onset of intercontinental admixture in Chile is one of the youngest migration events in Latin America during European colonization. By understanding the patterns of these genetic footprints it is possible to develop a more accurate view of population-level differences in biomedical traits and, therefore, inform the design of next-generation medical genetic studies in the country. Finer models for reconstructing local demographic history are being developed to increase the level of resolution of nation-wide diversity surveys.

19 Fine-scale genetic structure of Chilean populations
Ellen E Quillen¹, Anne Sheldrake¹, Jaydee Turner¹, Mark Shriver²
¹Texas Biomedical University
²Pennsylvania State University

The evolutionary history of human skin pigmentation has been studied extensively but the role of evolution in shaping tanning, a major response to UV exposure, has not. Quantitative analysis of persistence of tanning response was undertaken with 82 Mexican-American volunteers in San
Antonio, TX. Participants received four controlled exposures to ultraviolet radiation (UVA and UVB) in increasing doses. Erythemal and melanogenic dose-response were measured at 24 hours, 7 days, and 28 days post-exposure using a DSM-II Colorimeter (Cortex Technology). Minimal erythemal dose (MED, the amount of UV exposure necessary to induce visible reddening of the skin) and minimal melanogenic dose (MMD, the amount of UV exposure necessary to induce visible darkening of the skin) were determined at 24 hours and 7 days, respectively. Persistence of tanning response was calculated for each exposure as the melanin value at day 7 divided by melanin at day 28. Contrary to expectations, basal melanin level was only weakly predictive of either MED (R² = 0.026) or MMD (R² = 0.031), suggesting additional contributors. Biogeographic ancestry (BGA) was estimated from approximately 600,000 single nucleotide repeats (SNPs) assayed on the 23andMe custom Illumina platform. Population structure was evaluated using fastSTRUCTURE with inference constrained to two clusters reflecting the predominantly Indigenous American and Western European ancestry of the sample. The inclusion of BGA in the models for prediction of tanning intensity and persistence improve prediction indicating a potential population-level difference.

20 Evidence of selective constraint and sex-biased migration of human populations from X chromosome-autosome comparisons
Melanie H. Quiver¹, Kristin Mullen², Matthew E. B. Hansen², Michael A. Chen², PingHsun Hsieh³, Krishna R. Veeramah⁴, Sarah A. Tishkoff², Joseph Lachance¹
¹Georgia Institute of Technology
²University of Pennsylvania
³University of Arizona
⁴Stony Brook University

Because the number of X chromosomes differs for men and women, comparisons between sex-linked and autosomal genetic loci reveal sex-biased patterns of human demography. Using 44 high-coverage whole genomes from a diverse global set of 11 human populations we quantified the strength of selective constraint on X chromosomes and autosomes, found evidence of sex-biased colonization, and determined whether recent migrations are matrilocal or patrilocal. For each set of chromosomes, ratios of genic to intergenic diversity are similar across all studied populations regardless of subsistence pattern or geography. The strength of selective constraint on genes is greater for X-linked loci compared to autosomal loci - a pattern that is consistent with selection against deleterious recessive alleles. Using the pairwise sequential Markovian coalescent approach, we found evidence of large historic population sizes for West African Pygmies, but not Hadza or Sandawe populations. The ratio of X chromosome to autosome diversity is greater than the null expectation of 0.75 for African populations and less than 0.75 for non-African populations, with lower values for populations located farther from Africa. This pattern is consistent with a male-biased serial founder effect model, and computer simulations suggest a plausible out-of-Africa bottleneck size of 320-340 males and 40-70 females. Relative genetic distances at X-linked and autosomal loci reveal an evidence of female-biased movement between some populations and male-biased movement between other populations. These results are largely independent of each population’s mode of subsistence, which calls into question the idea that patrilocality is coupled with the emergence of agriculture.
21 The peopling of South America: new information from the phylogenetic refinement of human Y-chromosome haplogroup Q
Alessandro Raveane¹, V Grugni¹, L Ongaro¹, V Battaglia¹, B Trombetta¹, O Semino¹
¹University of Pavia
²University La Sapienz

The peopling of the Americas is a topic addressed by different disciplines. Archaeologists and anthropologists were the first to hypothesize an initial entry of Native American ancestors from Siberia across Beringia. Geneticists later confirmed this and provided new data on the timings and routes followed by America’s first colonizers. As for the peopling of South America, recent analyses of maternally inherited mitogenomes have allowed the identification of the first South American-specific subclades (B2i2, D1g, D1j, C1b13) within the founding “pan-American” haplogroups (A2, B2, C1b,c,d, D1 and D4h3a), providing new clues concerning the possible main entry routes from Mesoamerica. Less informative has been, so far, the male specific region of the Y chromosome, allowing the identification of two main founding lineages, Q-L54 and Q-M3, in Central and South America. Although both lineages carry a high STR haplotype variation, suggestive of an extensive internal differentiation, their current resolution level (too low) did not allow to obtain detailed information from the male perspective. In order to disentangle the nested haplogroup Q variation, 34 out of the 463 Y-chromosomes analysed in Battaglia et al. were selected taking into account both their geographic origin and STR haplotype status for a high coverage re-sequencing. On the whole, more than 800 new SNPs were discovered, identifying at least three L54 and five M3 sub-branches. The information obtained will be of great help to shed light on the first demographic processes that occurred in Central and South America.

22 Reconstructing recent admixture dynamics of post-colonial history in Mexico
Juan Esteban Rodriguez-Rodriguez¹, Javier Blanco¹, Alexander Ioannidis², Andrés Moreno Estrada¹
¹LANGEBIO
²Stanford University

In the past 500 years the genetic profile of Mexico has notably changed because of the considerable admixture of the distinct Native ethnic groups with incoming populations. Nevertheleess, it did not result in a homogeneous mixture all over the country. Mexico has a considerable population substructure due to historical events and distinct amounts of admixture between ethnic groups such as Africans, Europeans and Native Americans. Additionally, the migrations within Mexico have played a notable role on such substructure. Using genome-wide SNP array data from indigenous and admixed Mexican populations, we are exploring the ancestry tract length distribution from 7 different States across Mexico to infer the timing of admixture in each region. To that end we applied Tracts, a method that examines the length of all ancestry blocks in a given admixed population under the assumption that shorter blocks will be the result of older migrations while larger lengths will suggest earlier ones. The analysis will shed light on the time and size of the recent migration waves. Furthermore, we expect to correlate these admixture occurrences with historical events probably related to conflicts, slavery, natural disasters or other phenomena.
23 **Allele Sharing between Archaic and Modern Humans**

Alan Rogers
University of Utah

Rates of archaic admixture can be estimated from patterns of allele sharing between modern and archaic hominins. These estimates may be strongly biased when there is more than one archaic source population. In addition, current methods use only a fraction of the information in such data. We describe a new method, which uses all available site patterns and deals gracefully with complex patterns of phylogeny and admixture.

24 **Amerindian contribution in the pathogenesis of neuromyelitis optica in Mexico**

Sandra Romero-Hidalgo¹, José Flores-Rivera², Julio Granados³, Edmond J. Yunis⁴,⁵, Barbara Antuna-Puente¹, Marisela Villalobos-Comparán¹, Víctor Acuña-Alonzo⁶, Samuel Canizales-Quinteros¹,⁷, Teresa Corona²

¹Instituto Nacional de Medicina Genómica
²Instituto Nacional de Neurología y Neurocirugía Manuel Velasco Suárez
³Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán
⁴Dana-Farber Cancer Institute
⁵Harvard Medical School
⁶Escuela Nacional de Antropología e Historia
⁷Facultad de Química, Universidad Nacional Autónoma de México

Neuromyelitis Optica (NMO) is an autoimmune disease consisting in inflammation and demyelination of the optic nerve and the spinal cord. The NMO has a multifactorial etiology involving genetic factors. In contrast to multiple sclerosis, previous studies suggest that it is more common in non-Caucasian populations. In Mexico, the proportion of patients with NMO seems to be higher than in the rest of North America or Europe suggesting an Amerindian origin.

Objective: To study the Amerindian contribution of NMO patients at genome-wide level and to evaluate the role of the class I and II HLA genes.

Methods: 83 Mexican Mestizo NMO adult patients according to Wingerchuk criteria and 97 control individuals participating in the CANDELA-Mexico project. HumanOmniExpress microarray were used to estimate global, chromosomal and local ancestry. Data from Mexican indigenous, European and African population were used as reference. HLA genes class I and II were typed using high-resolution techniques.

Results: The results showed that NMO patients have a greater proportion of indigenous component compared to the control group (65% vs 59%; p = 0.05). Chromosomal and local ancestry estimations showed a significant indigenous enrichment particularly in 1q23.3 region where genes belonging to SLAM family are located. Regarding the HLA system, patients with NMO have a significantly higher frequency of HLA-DRB1*1602 as compared to controls (OR = 2.6, 95% CI: 1.4–4.9, p = 0.001).

Conclusions: These data suggest that NMO Amerindian contribution is particularly crucial in Mexican mestizos. Results from HLA system and chromosome 1 contribute to understand the genetic architecture that underlies the NMO.
Admixture Mapping Bronchodilator Drug Response in Latino and African American Children with Asthma

Zachary A. Szpiech*, Marquitta J. White*, Angel C.Y. Mak*, Donglei Hu, Esteban G. Burchard†, Ryan D. Hernandez‡
University of California San Francisco

Asthma is a common chronic inflammatory disorder of the airways with a large disparity in prevalence and mortality in the U.S. Bronchodilators are the most commonly prescribed asthma medication worldwide, regardless of disease severity. Recent heritability estimates suggest that genetic factors play a significant role in the inter-individual variation in bronchodilator drug response (BDR). Strikingly, variation in drug response is also seen at the population-level. Specifically, Latino and African American children have lower BDR than their European American counterparts. Latino and African Americans are highly admixed.

We leveraged the ethnic-specific disparity in BDR with recent admixture, characteristic of Latinos and African Americans, to uncover genetic variants associated with variation in BDR using admixture mapping. We analyzed whole genome sequence (WGS) data from 1,500 Latino and African American children with asthma and bronchodilator drug response to albuterol. We performed admixture mapping and identified several loci where drug response was associated with local ancestry. We then fine-mapped these associated regions using the full WGS data to identify novel ethnic-specific variants associated with BDR. To our knowledge, this is the first study to incorporate WGS data from a study of ethnically diverse children to explore the underlying genetic architecture of asthma drug response. Our results have the potential to lead to the development of new and more effective asthma therapies and serve to further the goals of the Precision Medicine Initiative.

Neanderthal ancestry in Yemeni populations

Deven N. Vyas¹, Ali Al-Meeri², Connie J. Mulligan¹

¹University of Florida
²University of Sana’a

Analyses of the Neanderthal genomes are strongly consistent with the idea that early modern human populations interbred with Neanderthals in the Near East soon after their successful dispersal out of Africa. However, it is unclear which part of the Near East this dispersal first crossed through, and few studies have assessed Neanderthal ancestry in modern populations.

We assayed 90 Yemenis sampled from across the country using the Affymetrix Human Origins array. We merged our data with previously published datasets from around the world as well as data from Neanderthal and Denisovan genomes in order to estimate archaic ancestry in modern populations. We also conducted ADMIXTURE analyses on a more regionalized dataset in order to infer ancestral components in Near Eastern populations.

Consistent with other studies, we found that North African and Near Eastern populations (including the Yemeni) generally have less Neanderthal ancestry than other western Eurasian populations. However, our ADMIXTURE results indicate that a subset of Yemeni samples from the Mahra governate share a very high level of ancestry (~85%) with a single Near Eastern component. Interestingly, these individuals have Neanderthal ancestry estimates that are greater than estimates from almost all Near Eastern and North African populations and are more consistent with estimates from European and South/Central Asian populations, suggesting that
eastern Yemen may be an area of elevated Neanderthal introgression in the Near East. Greater sampling of Near Eastern populations is needed to better understand variation in Neanderthal ancestry and the site(s) where modern humans and Neanderthals interbred.

### 27 Characterizing Patterns of Admixture and Inheritance across the Hybrid Genome of the Africanized Honeybee (*Apis mellifera scutellata*)

Daniela Zárate, Thiago Lima, Ron Burton, Joshua Kohn  
University of California San Diego

The honeybee subspecies *Apis mellifera scutellata* was brought from southern Africa to Brazil in the 1950s in an attempt to breed honeybees better suited to tropical conditions. Escaped swarms soon hybridized with existing European honey bees (EHB) (*Apis mellifera mellifera*) creating a hybrid known as the Africanized honeybee (AHB) (the so-called killer bee). AHB have since spread north from Brazil, largely replacing feral EHB as their range expands. AHB reached as far northward as San Diego in 1994. At present the majority (65%) of honeybees in San Diego County are feral (non-managed) and the majority of these carry the African mitochondrion.  

interest in preventing further loss. Here we propose to apply next-generation sequencing technology to map the portions of the genome of Africanized honeybees (AHB) that are derived from their African vs. European ancestors. The project aims to: 1) Measure how consistent the amounts and locations of genomic contributions from each ancestral lineage have been along the entire route of hybridization from Brazil to the southern United States. 2) Identify genomic regions under strongest selection. 3) Search for genes that may underlie the aggressive tendencies of AHB. The long-term goal is to find regions of the AHB genome that might provide traits useful for the honeybee industry facing hive mortality pressure.
Organizing Committee
Anne Stone, Arizona State University
Amy Goldberg, Stanford University
Heather Norton, Cincinnati University
Ellen Quillen, Texas Biomedical Research Institute

With thanks to
Abigail Bigham, University of Michigan
Joanna Malukiewicz, Federal University of Viçosa
Richard Hagan, University of Oklahoma

American Association of Anthropological Genetics Student Travel Awardees
Erika Landa & Juan Esteban Rodriguez-Rodriguez

Conference Travel Awardees
Alicia Martin
Amanda Pierce
C. Eduardo Guerra Amorim
Cesar A. Fortes-Lima
Daniela Zarate
Emily Puckett
Felicia Gomez
Francesco Montinaro
Keurcien Luu
Zachary Szpiech