

SMBE Satellite Meeting Scientific Report:
“Molecular Evolution in Small Populations”
Princeton University, Princeton, New Jersey, USA
June 28-30, 2023

Overview. The scientific programming for our conference centered around four key themes: (1) theoretical approaches for studying the evolution of small populations, (2) understanding the impact of complex demography on the evolutionary process, (3) approaches for measuring fitness and natural selection in wild populations, and (4) conservation applications. Below, we summarize the talks and discussions relevant to each of these themes that occurred during the meeting. The full meeting program, including talk and poster abstracts, and talk recordings, can be found on the meeting website: <https://smbe-smallpops2023.com/>

Theoretical approaches for studying small populations. A central component of understanding how evolution operates in small populations is developing theoretical and methodological approaches. Theory can inform us about what we might expect to happen in small populations under an evolutionary model, and methods can help us infer important parameters to test hypotheses based on theory.

In our talk session on this topic, we heard from researchers working at the cutting edge of developing and applying new theoretical and methodological approaches for inferring the demographic and selective forces operating in small populations. Our session kicked off with a talk from Dr. Krystyna Nadachowska-Brzyska (Jagiellonian University) on the complexities of effective population size. Dr. Nadachowska-Brzyska’s talk highlighted the division in the literature between methods that estimate “contemporary” vs “historical” N_e and discussed the importance of considering the entire temporal spectrum of N_e when thinking about small populations. Our next talk in this session came from Dr. Zachary Szpiech (Penn State University), who presented a new method for detecting positive selection from unphased data. Dr. Szpiech demonstrated how his method for detecting selective sweeps using unphased data can perform nearly as well as methods that require phased data—a major advance given the many challenges often posed by phasing. Finally, we also heard from Dr. Diego Ortega del Vecchyo (National University of Mexico), who detailed how we might be able to use the ancestral recombination graph (ARG) to infer the action of natural selection on new deleterious mutations. Dr. del Vecchyo provided a great introduction to the ARG and described how it can be leveraged to estimate the impact of new mutations on fitness, a key parameter when modeling the impact of deleterious mutations in small populations.

In addition to these oral presentations, we also had several posters that highlighted theoretical work on small populations. Some highlights included a poster from Ulises Hernandez (University of Arizona) where he described an analytical model for exploring the impact of epistasis on mutational meltdown, as well as a poster from Patrick McKenzie (Columbia University) describing the expected distance across a chromosome for a particular genealogy under the Sequential Markovian Coalescent.

Inferring complex demography from genomic variation data. Small populations are ultimately defined by their demography, and inferring demography using genetic variation datasets has long been a central aim in population genetics. One key takeaway from studies in natural populations is that demographic trajectories are often complex and most species are far from equilibrium. In this session, we heard from researchers working to understand complex demography across recent and ancient timescales and across a wide array of taxa, from humans to woolly mammoths.

Dr. Sergio Nigenda-Morales (California State University, San Marcos) described his work to model the demographic history of two fin whale populations: fin whales in the Eastern North Pacific and the isolated Gulf of California population. Using the site frequency spectrum, he and his team were able to detect the signature of the whaling-induced bottleneck in the Pacific population, and a long-term history of small population size and limited gene flow in the Gulf population that was not impacted by whaling. This fine-scale demographic inference was then used to design complex genetic simulations to understand the impacts of human activities and gene flow on genetic load. Dr. Marianne Dehasque (Stockholm University) then took us further back in time, describing her inference of the population history of isolated woolly mammoths on Wrangel Island using ancient mammoth genomes. Her analyses indicate that the Wrangel Island mammoths experienced an initially severe bottleneck during the founding of the population, but then persisted at demographically stable levels for the next ~6,000 years, experiencing purging of the most highly deleterious mutations. This snapshot of an ancient small population provides important insights to understanding the dynamics and possible fates of bottlenecked and isolated populations today.

In the second half of the session, Dr. Aaron Shafer (Trent University) used empirical examples from his work on the demography of white-tailed deer to highlight different metrics, including Tajima's D , mutational load, and F_{ROH} for measuring genetic diversity and historical demography. He discussed the uncertainties inherent in the N_e/N_c ratio and the importance of linking genomic metrics to population viability and phenotypes in order to inform management decisions. Dr. Shyamalika Gopalan (Duke University) took us through her fascinating work on the demographic history of the Chabu hunter-gatherer people of Southwest Ethiopia. Using identity-by-descent (IBD) tract demographic inference methods (IBDNe), she has found that the Chabu population has been declining over the past 1,400 years, with a divergent population history from neighboring agriculturist populations which showed signals of exponential growth in the same time period as farming spread through the region. Her talk demonstrated the power of demographic inference methods based on numerous whole genomes to detect relatively recent demographic changes, and she discussed important caveats, such as gene flow, which may confound results.

In addition to these outstanding talks, there were numerous posters applying cutting-edge demographic inference methods to address questions in small population biology, in species ranging from orchids (Brett Morgan, Smithsonian Environmental Research Center) to mouflon sheep (Eren Yüncü, Middle East Technical University), to white-lipped peccaries (Anna Carolina Mazará, Universidade Federal do ABC (UFABC)).

Novel approaches for measuring fitness in small populations. A key feature of small populations is that elevated genetic drift can overwhelm the effects of positive and negative natural selection. However, quantifying the effects of selection and identifying the mutations controlling fitness in natural populations has long posed a major challenge in the field. Our second plenary speaker, Dr. Yvonne Willi (University of Basel), framed this interplay of drift and selection in the context of the extinction vortex. In her talk, Dr. Willi reviewed the empirical evidence for aspects of the theoretical extinction vortex but emphasized the need for a meta-analysis investigating aspects of the extinction vortex using genomic data. Dr. Willi then shifted to discuss her own research in the effects of population size using the rock/sand cress (*Arabidopsis lyrata*). Through both population genetic analyses and common garden experiments, Dr. Willi's work showed that populations on the range edges with small effective population sizes have lower fitness, higher mutational load, and lower capacity to adapt.

In our session on approaches to studying fitness, Dr. Henrique Figueiró (Smithsonian-Mason School of Conservation) presented an in-depth analysis of genomic diversity and mutational load in the

captive breeding population of black-footed ferrets. Dr Figueiró showed that breeding black footed ferrets have extremely low genome-wide diversity, and that elevated mutational load can be found in genes related to immunity and reproduction. Dr. Nancy Chen (University of Rochester) then discussed her work using a multi-generational pedigree of Florida Scrub Jays to infer the fitness of individuals and haplotypes, as well as her use of genome data to study complex trade-offs between fecundity and survival. Such trade-offs may contribute to the maintenance of fitness-related variation in small populations. Finally, Dr. Giorgio Bertorelle (University of Ferrara) presented a study that started by identifying putatively deleterious alleles in an endangered brown bear population in Italy and proceeded to use gene editing technology to test the phenotypic effect of these alleles in cultured cells. Despite this mutational load, Dr. Bertorelle's simulations showed that genetic rescue is not required for the survival of this population.

In addition to these four oral presentations, there were several submissions to our poster session that dealt with different angles of studying fitness in small populations. For example, Dr. H. Lisle Gibbs (The Ohio State University) presented on differences in functional and neutral genomic diversity in an endangered rattlesnake, and Anna Maria Calderon (Pennsylvania State University) presented a poster highlighting inbreeding and mutational load in *Setophaga* warblers. Dr. Julia Höglund (Stockholm University) took a theoretical approach to the topic and presented new methods to predict deleterious variants in non-model organisms. To show how some species may be able to escape the extinction vortex, Sabhrina Aninta (Queen Mary University of London) showed that smaller populations of island ungulates purged highly deleterious load more quickly than larger populations.

Conservation applications. A key application of research on molecular evolution in small populations is informing conservation management strategies for threatened or endangered species. The consequences of small population size, including greater genetic drift and inbreeding, can often contribute to fitness declines and further exacerbate extinction risk.

Our meeting began with a plenary talk that was highly relevant to this topic, from Dr. Uma Ramakrishnan (National Centre for Biological Sciences, Bangalore) on her work using genomics to inform the conservation of Indian tiger populations. Dr. Ramakrishnan highlighted that Indian tigers have the highest genetic diversity of any tiger species, though many populations now exist in small and isolated patches. Dr. Ramakrishnan also showcased her work aiming to quantify the deleterious mutation load across Indian tiger populations, and discussed her plans for trying to better understand the fitness consequences of deleterious mutations using captive tiger populations.

On the second day of our meeting, we hosted a talk session focused on conservation applications. This session began with a talk from Dr. Klaus Koepfli (Smithsonian-Mason School of Conservation), who talked about his work aiming to use genomics to inform captive breeding programs of the Dama gazelle. Specifically, Dr. Koepfli discussed the role of genomics in guiding a potential genetic rescue between highly diverged subspecies of the Dama gazelle. Next, we heard from Dr. Jana Wold (University of Canterbury) on her project developing genomic resources for the kākāpō, a critically endangered parrot native to New Zealand. Dr. Wold highlighted some of the challenges in assaying structural variants in the species using long reads and graph-based assembly approaches. Dr. Fuwen Wei (Jiangxi Agriculture University) gave a wonderful talk summarizing his work applying genomics to several flagship species in China, focusing on the giant panda. Finally, our session concluded with a talk from Dr. Eduardo Eizirik (Pontifical Catholic University of Rio Grande do Sul) on his genetic and genomic research on jaguars. Dr. Eizirik showed us how population structure in jaguars has been transformed by human influence, highlighting the many small and isolated populations of this species that now exist in Brazil.

Our poster session similarly featured numerous projects aiming to use genomics to inform conservation management. For instance, David Tian (UC Berkeley) presented his work on the endangered Devil's Hole pupfish, demonstrating the genomic consequences of extremely small population size in the species. Similarly, Hana Thompson (Queen's University) gave a poster on using genomics to inform conservation of the Eastern Massasauga Rattlesnake and Erangi Heenkenda (Purdue University) gave a poster on her work using genomics to inform translocation strategies in the Northern Bobwhite. In general, we found that a large fraction of our attendees were interested in conservation applications, such that many stimulating conversations were had on this topic during coffee breaks, poster sessions, and dinner.

Conclusions & Takeaways. Together, the plenary, talks, posters, and informal discussions showed that we are making great progress in realizing the potential for using genomics to understand how evolution proceeds in small populations. However, these discussions also helped identify some key areas that are ripe for future research, leaving us with several overarching takeaways from the meeting. These include:

1. **More standardization of metrics for deleterious variation and inbreeding is needed.** The use of various metrics of deleterious variation (such as “masked load” and “realized load”) and inbreeding (based on proportions of runs of homozygosity of varying lengths) were widespread in the meeting. However, many discussions on this topic centered on the lack of standardization of these metrics as well as a broader discussion of how meaningful such genomic measures of putatively deleterious variation are for fitness. Thus, a critical area for future work in the field is to better standardize how we measure genetic load and inbreeding and validate such genomics-based metrics using more direct measures of fitness to improve our interpretation of these metrics.
2. **Understanding the complex demographic history of a population is essential.** Numerous oral and poster presentations employed approaches for inferring the historical demography of a population using genomic data, and highlighted the importance of complex demography in determining evolutionary dynamics. Yet, at the same time, numerous pitfalls in implementing and interpreting various methods for demographic inference remain, particularly when trying to infer the very recent changes in population size that are most relevant for management. Thus, a key takeaway from the meeting is that more work is needed to better validate existing demographic inference methods and identify areas where they may succeed or fail.
3. **A large gap remains between genomics research and conservation practice.** Given that the majority of attendees had interests in conservation applications, many discussions centered around how to better bridge the gap between academic genomics research and conservation. One key area of disconnect that was identified is that, whereas academic researchers seem most interested in exploring novel genomic technologies and theoretical frameworks, conservation practitioners are primarily in need of cheap and easily employed tools such as SNP arrays. How to better align the objectives of academic researchers and conservation practitioners remains a key area for future discussion.
4. **Computational simulation tools have not yet been adopted by the field.** Computational simulation modeling approaches were notably absent from our meeting. Although these approaches for modeling population dynamics have been made increasingly accessible by user-friendly software such as SLiM, almost none of the talks or posters at our meeting featured a simulation analysis. This appears to be a major area for future growth in the field, particularly given the challenges in interpreting genomics-based metrics of deleterious variation outlined above.