GENOMICS IN THE JUNGLE V - PERU

Course Syllabus

Instructors:

Gideon Erkenswick

Senior Scientist, Field Projects International

Ph.D., Ecology, Evolution and Systematics

Mrinalini Watsa

Scientist, San Diego Zoo

Ph.D., M.A., Biological Anthropology

Course Introduction

Photo: Tim Paine

Biology has turned to genetic research methods for a deeper look into the biological factors that encode behavior and physiology. We use genetic techniques to determine species delimitations, define populations, understand mating systems, explain behavioral differences in foraging efficiency, screen for disease, conduct paternity studies, evaluate immune status and functioning, and explore microbiome diversity... and these are just a few examples applied to wildlife biology. High throughput genetics is revolutionizing biological research. In the past decade we have witnessed the successful deployment of instruments that enable molecular work to be conducted 'on-the-fly' and in the field. These tools are minimizing the hassles and barriers associated with transporting samples around the world to distant laboratories that possess the equipment and resources to extract, amplify, and sequence DNA. In many ways, this technology is democratizing wildlife research by empowering field scientists all around the world with genetic tools to directly advance their research and conservation aims. This course is taught by 4 researchers who have dedicated their careers to just

Stefan Prost

Faculty, University of Oulu

Ph.D., Biology

Zane Libke

Researcher, In Situ Lab Initiative

Jhakelin Reyes

Lab Manager, ACCA

that. Together, they form part of the In Situ Lab Initiative, a program that promotes the development of on-site wildlife genetics labs all over the world. They have sequenced DNA in the jungles of South America, the African savannah, the Arctic, and more. Collectively, they have more than 20 years of experience teaching Nanopore sequencing in remote settings.

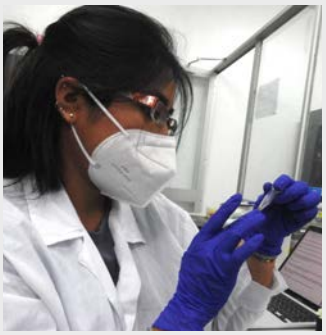
This course will take place in the Peruvian Amazon, where participants will benefit from interactive, informative lectures from our expert instructors. The course has an emphasis on practical skills, students will participate in everything from sample collection, wet lab processing, sequencing, data analysis, and interpretation of results. It will take place at the Los Amigos Biological Station, the site of the Wildlife Conservation Laboratory - a one-of-a-kind biosafety level 2 genetics laboratory in the heart of the Amazon rainforest. This course will provide dedicated biologists with a firm theoretical and practical understanding of biodiversity-focused Nanopore sequencing applications. Upon completion of the course, you will be equipped with ample knowledge and experience to design & execute successful sequencing experiments in non-traditional lab environments.

Course Research Objectives

This is a course that repeats annually. Because of the rapidly evolving nature of genomics, topics can vary from year to year, although some topics & methods are always covered (i.e. DNA barcoding). Others may be a short-term focus (e.g. ddRADseq, nanopore adaptive sequencing, eDNA). We work from samples collected opportunistically or as part of FPI's long-term vertebrate research projects in order to achieve the following long-term aims (class size and time permitting):

Objective 1: Recording Species Richness & Diversity

Biodiversity is a vital indicator of environmental health, and worldwide biodiversity is on the decline. Today, at least 15,000 species are threatened with extinction (Smithsonian), but this number is likely much larger. This is because measuring diversity is extremely difficult, especially in biodiversity hotspots. Despite centuries of scientific exploration, an estimated 17% of total species remain undescribed (Liu et al. 2022). To inventory and understand this biodiversity, we must answer a range of difficult questions, and develop faster, more effective, and efficient tools to do so.



Gabriela Caceres

Lab manager, ACCA

Location:

Los Amigos Biological Station (EBLA), Madre de Dios, Peru

In this course, we record species with sequence data collected directly from specimens or indirectly from environmental samples. Traditional sequence targets for eukaryotes and prokaryotes include 16S, 18S, CO1, CytB. These records, which are ultimately deposited to publicly accessible databases become invaluable in future assessments of diversity at this site. Constructing a reference DNA library for this site is an extremely valuable case study for the field station and an essential task that most field biologists will exercise in the future.

Objective 2: Metabarcoding

If you have read any recent literature on diet or microbiome analyses then you are already familiar with the concept of METABARCODING, not to be confused with metagenomics. Metabarcoding is when you enrich for the same gene target from a mixed species sample. For example, we amplify the 16S gene from all bacteria in a sample because it has conserved primer regions and lots of variability in-between. With these sequences, we compare microbiome richness, abundance, and diversity across samples. eDNA and diet research has the same approach, but with differences in extraction protocol or gene targets depending on sample type and taxonomic interests. In all courses, we discuss the methodology underlying metabardocing projects, practice the majority of the requisite laboratory techniques, and provide basic data analysis pipelines that participants can use with data generated on-site or with the multitude of publicly available datasets. Whether we generate metabardocing sequence data from our practical exercises varies from course to course.

Objective 3: Whole Genome Sequencing

Not very long ago whole genome sequencing was only available high-throughput research and commercial labs. Now, with portable sequencing technology, it is completely feasible to sequence entire genomes. We strive to sequence entire genomes because they are the basis for studying population genetics and differentiating large and unresolved species complexes, among many other things. For applied research, we have to ask ourselves how much data coverage is needed at each position throughout the genome, or do we need to know the entire nuclear genome versus just the mitochondrial genome. In this course, we will experiment with low coverage whole genome sequencing with the goal of assembling entire mitochondrial genomes to differentiate species.

Objective 4: Nanopore Adaptive Sequencing

A common challenge for forensic science, pathogen research, and many other genetics projects is enriching for target DNA. Environmental samples are full of bacterial and eukaryotic DNA, but our sequencing capacity is much more limited. That is to say, the answer to your research question is present, but sometimes the target you need is proportionally so small that your success depends on your ability to enrich for what you are interested in, avoiding all the other noise. Culture and PCR are well known methods for enriching for specific targets, but others include ChIPseq, hybridization capture, and CRISPR. In this course we will experiment with another method call nanopore adaptive sequencing (NAS), in which active sequencing pores receive information in real-time that instructs the pore to continue with a particular DNA fragment or reject it in search of something else.

Across these objectives this course will give participants a fundamental understanding of the following topics:

Experiment design - sample collection and storage - sterile technique - DNA extraction - PCR amplification - gel electrophoresis - DNA clean-up - library preparation - nanopore sequencing - high throughput sequence data analysis - bioinformatics

A few additional questions with all case studies:

- To what extent can we increase our efficiency by multiplexing samples and even projects on a single flow cell?
- Can we effectively wash a flowcell and then use it for an entirely different purpose (AKA two experiments for the price of one)?
- How much data do we need to minimize sequencing error?

Course Goals

The broad goals of this course are to give participants advanced training in field techniques important to the collection of biological samples from wildlife, and their prey, their parasites, their environment all the way to sequencing DNA from these sources. In more detailed form, this includes:

- To engage in both independent and team-based data collection
- To teach sample collection techniques from invertebrates to megafauna
- To learn sample storage and clean-lab protocols tailored for fieldwork
- To extract DNA in a portable field laboratory
- To test DNA quality and quantify it
- To run basic PCRs for a range of markers using multiple protocols (smaller, lighter, lower-scale and more rugged than typical lab-based instruments)

- To explore metagenomics in the field using case studies
- Whenever possible, to involve all participants in the publication of data collected on the field course.

Bibliography

Liu, J., Slik, F., Zheng, S., & Lindenmayer, D. B. (2022). Undescribed species have higher extinction risk than known species. *Conservation Letters*, *15*(3), e12876.

https://naturalhistory.si.edu/education/teaching-resources/paleontology/extinction-over-time

Course Topics

Topic of Study	Activity	Description
I. Introduction		
> Threats to the Amazon in the Madre de Dios Department of Peru; conservation efforts of the Amazon Conservation Association (ACA); conservation efforts of FPI	Lecture	A review of the major conservation approaches in the MDD, including the conservation and research efforts of ACA and FPI.
> Field ethics, safety precautions, rules,and useful tips.	Discussion	Keeping your footprint to a minimum while working with wildlife in the tropics, and ensuring your safety and that of the wildlife around you.
> DNA sequencing and genomics	Lecture/Discussion	An introduction to genomics and its practical applications in the field
> Genomics and Ethics	Discussion	What should we be thinking about when applying genomics to wildlife research? Whose rights come into question? Do different countries regulate genomics in different ways?
II. Navigation and Space Use		
> Basic functions of a handheld GPS and compass	Demonstration	Getting familiar with the most important pieces of equipment you will have in the field.
> Waypoint and track data and how to use them	Exercise	Recording key features of the research station with waypoints and tracks
III. Collecting Biological Specimens		
> Field methodology: indirect observation and biological sampling	Lecture/Practical Exercise	Tracking primates, identifying plants based on botanical features, and collecting non-contaminated samples from each.
> Field methodology: sample storage and preparation	Lecture/Practical Exercise	Sample collection methods in the field, time-scales for DNA deterioration, and field laboratory sterile technique
IV. Basic Field Laboratory Techniques		

Topic of Study	Activity	Description
> Basic laboratory techniques	Practical Exercise	Regardless of participant background, we will spend a moment practicing good pipetting techniques, sample volume calculations, and gel loading techniques. Repetition is key, getting everyone onto the same skill level
> Laboratory Safety	Lecture/Practical Exercise	Even in the field, lab safety is critical. We will go over protocols, cautionary tales, and specific examples of how not to hurt yourself in a field laboratory.
> Lab recipes	Lecture/Practical Exercise	We will learn the processes behind the various lab protocols or recipes we will be following in this course
V. Genomics methods		
> DNA barcoding	Lecture	The history of species identification using barcodes, including some of the most exciting applications in wildlife science
> Metabarcoding & Metagenomics	Lecture	Microbiome, eDNA, Diet, Trace DNA - where they overlap and where they differ.
> High-throughput sequence analysis overview	Lecture	Basecalling, filtering, OTU and consensus forming, alignments, classification
VI. Genomics in a Jungle Lab		
> DNA extraction	Practical Exercise	Whole genome amplifications and genomic DNA extraction
> DNA quantification	Practical Exercise	How much DNA do you have?
> PCRs	Practical Exercise	Amplifying markers for all three case studies using PCRs in a field laboratory
> Gel electrophoresis	Practical Exercise	Testing if your PCR worked using gel electrophoresis and PCR product quantification
> Library prep	Practical Exercise	Creation of DNA libraries from samples

Topic of Study	Activity	Description
> Multiplexing and indexing	Practical Exercise	Minimising sequencing costs by running multiple samples in a single run - multiplexing and indexing samples to tell them apart afterwards
> Using the MinION	Practical Exercise	Learn to run sequences on a MinION device - cutting edge genetic sequencing using nanopore flowcells
> Basic bioinformatics	Practical Exercise	How to interpret data from the MinION - both real-time and post-hoc
V. Excursions and Activities (time and weather permitting)		
> On-trail and off-trail forest hikes	Excursion	
> Night surveys	Excursion	
> Oxbow lake and palm swamp	Excursion	
> Los Amigos River (tentative)	Excursion	

Daily Schedule

Days begin at 6am and conclude around 7:30pm at dinner time. On some days, activities will resume after dinner, for example, to carry out a nighttime survey for amphibians and reptiles or to complete am extended laboratory protocol. A detailed day-by-day schedule will be provided towards the end of July in time for the participants to be well aware of each days activities prior to departure to the field site.

Course Work

Lab notebook (250 pts)

Paper notebooks in a lab are soon becoming a thing of the past. We will instead keep a digital laboratory notebooks to keep up with the times, and to conserve paper in a field environment where paper copies are not reliable. For the purposes of this course, students can use any word processing software that they prefer and have available on their own laptops.

Sample List (100 pts):

All participants will contribute towards the course sample list, documenting a range of metadata for each sample that will help us identify it in the future. The ongoing sample list will be judged as a group at the end of the course and individual effort to locate and collect samples will also be appreciated.

Lab Exercises (250 pts):

Every case study is going to involve a series of lab sessions, some which may have reports due at the end. Your ability to complete the exercises, remain enthusiastic and a good team player will determine your evaluation for the lab exercise. Evaluations are not intended to stress or discourage any participants, they are provided to you in the hope they will be helpful for you to decide how to prioritize your ongoing training/education.

Navigation (100 pts)

All field courses, even one that focuses on genetics, requires that students learn basic forest orienteering skills. Participants must learn and demonstrate how to use a handheld GPS and compass to navigate a trail system, find their way back to the trail system, and ultimately back to basecamp.

Quizes (100 pts, each)

Short quizes (usually 2 per course) are not intended to frighten or stress any participants they are merely a way for participants and instructors to understand what information is being absorbed.

Reading List (subject to change)

Pomerantz, A., Penafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L.A., Barrio-Amoros, C.L., Salazar-Valenzuela, D. and Prost, S., 2017. Real-time DNA barcoding in a remote rainforest using nanopore sequencing. bioRxiv, p.189159.

Yildirim, S., Yeoman, C.J., Sipos, M., Torralba, M., Wilson, B.A., Goldberg, T.L., Stumpf, R.M., Leigh, S.R., White, B.A. and Nelson, K.E., 2010. Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities. PloS one, 5(11), p.e13963.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Douglas, W.Y. and De Bruyn, M., 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution, 29(6), pp.358-367.

Janjua, S., Fakhar-I-Abbas, William, K., Malik, I.U. and Mehr, J., 2017. DNA Mini-barcoding for wildlife trade control: a case study on identification of highly processed animal materials. Mitochondrial DNA Part A, 28(4), pp.544-546.

Pomerantz, A., Sahlin, K., Vasiljevic, N., Seah, A., Lim, M., Humble, E., ... & Prost, S. (2022). Rapid in situ identification of biological specimens via DNA amplicon sequencing using miniaturized laboratory equipment. *Nature Protocols*, *17*(6), 1415-1443.

Karin, B. R., Arellano, S., Wang, L., Walzer, K., Pomerantz, A., Vasquez, J. M., ... & McGuire, J. A. (2023). Highly-multiplexed and efficient long-amplicon PacBio and Nanopore sequencing of hundreds of full mitochondrial genomes. *BMC genomics*, *24*(1), 1-12.

Krehenwinkel, H., Pomerantz, A., Henderson, J. B., Kennedy, S. R., Lim, J. Y., Swamy, V., ... & Prost, S. (2019). Nanopore sequencing of long ribosomal DNA amplicons enables portable and simple biodiversity assessments with high phylogenetic resolution across broad taxonomic scale. *GigaScience*, *8*(5), giz006.

Watsa, M., & Wildlife Disease Surveillance Focus Group. (2020). Rigorous wildlife disease surveillance. *Science*, *369*(6500), 145-147.

Kipp, Evan J., et al. "Nanopore adaptive sampling for targeted mitochondrial genome sequencing and bloodmeal identification in hematophagous insects." *Parasites & Vectors* 16.1 (2023): 68.

De Vivo, Mattia, et al. "Utilisation of Oxford Nanopore sequencing to generate six complete gastropod mitochondrial genomes as part of a biodiversity curriculum." Scientific Reports 12.1 (2022): 9973.

Grading Criteria

Individual and group assignments will be assessed according to the following point schedule:

Assessment Item	Date Due	Points Possible	Total Points Possible
Navigation	NA	100	
Lab notebook	END OF COURSE	250	
Lab exercises	END OF EACH EXERCISE	250	
Sample List	END OF COURSE	100	1000
Quiz 1	TBD	100	
Quiz	TBD	100	
Enthusiasm	END OF COURSE	100	