

NANOPORE SEQUENCING OF THE CACHABI ROBBER FROG (*PRISTIMANTIS*
ACHATINUS) MITOCHONDRIAL GENOME

by

Toriann Ayzlynn Molis

HONORS THESIS

Submitted to Texas State University
in partial fulfillment
of the requirements for
graduation in the Honors College
May 2022

Thesis Supervisor:

David Rodriguez

Second Reader:

Utpal Smart

COPYRIGHT

by

Toriann Molis

2022

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, Toriann Molis, authorize duplication of this work, in whole or in part, for educational or scholarly purposes only.

DEDICATION

I dedicate this work to my ever-supportive partner, Joshua Rogalski, and our dog, Oatly.

Thank you for keeping the barks to a minimum.

ACKNOWLEDGEMENTS

I am tremendously grateful to Dr. David Rodriguez for making this project possible and for all the opportunities and invaluable guidance he has provided me. I am also thankful for everyone in the RDZ lab who has helped me along the way. I especially thank Stephen Harding for taking the time to teach me many of the lab skills that were instrumental to beginning this project. I also very much appreciate Dr. Utpal Smart for acting as my second reader and his contributions to making this document legible. Thank you as well to the Texas State Honors College for giving me the opportunity to present this research.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
LIST OF ABBREVIATIONS.....	x
ABSTRACT.....	xi
CHAPTER	
I. Introduction.....	1
II. Methods and Materials.....	8
Primer Testing and Sanger Sequencing.....	9
Nanopore Run without Treatment.....	11
Nanopore Run with WMGA.....	12
Nanopore Run with REPLI-g UltraFast WGA.....	14
Nanopore Run with Magnetic Bead Capture.....	15
Nanopore Run with Mechanical Shearing.....	16
III. Results.....	18
Primer Testing and Sanger Sequencing.....	18
Nanopore Run without Treatment.....	18
Nanopore Run with WMGA.....	19
Nanopore Run with REPLI-g UltraFast WGA.....	19

Nanopore Run with Magnetic Bead Capture	19
Nanopore Run with Mechanical Shearing	19
IV. Discussion.....	20
VI. Conclusion	26
APPENDIX SECTION.....	27
REFERENCES	30

LIST OF TABLES

Table	Page
1. Table 1	9
2. Table 2	10
3. Table 3	11
4. Table 4	12
5. Table 5	13
6. Table 6	13
7. Table 7	14
8. Table 8	14
9. Table 9	16
10. Table 10	16

LIST OF FIGURES

Figure	Page
1. Figure 1	7
2. Figure 2	11
3. Figure 3	12
4. Figure 4	14
5. Figure 5	15
6. Figure 6	16
7. Figure 7	18
8. Figure 8	18
9. Figure 9	20

LIST OF ABBREVIATIONS

B.d. --- *Batrachochytrium dendrobatidis*

bp --- base pairs

DNA --- deoxyribonucleic acid

ExoV --- Exonuclease V

kb --- kilobases

mitogenome --- mitochondrial genome

mtDNA --- mitochondrial DNA

ng --- nanograms

nucDNA --- nuclear DNA

NUMTs --- nuclear mitochondrial insertions

ONT --- Oxford Nanopore Technologies

PCR --- polymerase chain reaction

ssDNA --- single stranded DNA

tDNA --- template DNA

T7EI --- T7 Endonuclease I

μl --- microliter

WGA --- whole genome amplification

WMGA --- whole mitochondrial genome amplification

ABSTRACT

Due to the devastating impacts of chytridiomycosis caused by *Batrachochytrium dendrobatidis* on tropical amphibians, it is vital to understand their host-pathogen dynamics. These studies can be better informed by measuring genetic diversity in the hosts, especially where cryptic diversity is present, such as in the tropics. Measuring diversity can now be performed *in situ* with portable nanopore DNA sequencing technology. To facilitate this, universal PCR primers have been designed to generate mitochondrial DNA (mtDNA) fragments of ~3400 base pairs (bp) when applied to most amphibians; however, in *Pristimantis achatinus*, they only generate fragments of ~900 bp, likely due to a rearrangement in its mitogenome. To investigate the cause of the decreased amplicon size, we performed a series of tests using nanopore sequencing, wherein we attempted to sequence the entire mitochondrial genome of *P. achatinus*. Different methods of mtDNA extraction were tested as well as different means of whole genome amplification. The sequence data obtained from these nanopore runs will provide necessary insights into the genetic diversity of these amphibians, while the tests will help to optimize a protocol for nanopore sequencing of whole mitochondrial genomes of amphibians, furthering our ability to uncover genetic diversity and understand the dynamics between fungal pathogens and their amphibian hosts.

I. Introduction

Tropical rainforests are among the most biologically diverse and productive ecosystems on the planet. High levels of rainfall and nutrients, soaked up by the forest's vegetation, allow for the production of significant proportions of biomass and bioenergy. Due to the plethora of vegetation, these environments function as important carbon reservoirs, and they play a vital role in keeping the planet healthy through their absorption of carbon dioxide and release of oxygen via photosynthesis (Phillips, et al. 2017). Additionally, rainforests house numerous plants that are used as foods and in medicines, making these environments invaluable to both the indigenous communities within them as well as to people around the globe (Vandebroek, et al. 2004).

Much of what tropical rainforests can provide is due to the incredible biodiversity that exists within these ecosystems. A testament to this diversity is the number of amphibian species housed within these habitats. According to amphibiaweb.org, the rainforests of Ecuador alone host over 635 native species of amphibians, making it one of the most amphibian-rich countries in the world. Unfortunately, 57% of these amphibian species are under threat of extinction, making Ecuador the country with the third highest number of threatened amphibian species in South America and one of the highest in the world (Ortega-Andrade, et al. 2021). According to the International Union for Conservation of Nature Red List, amphibians are more threatened than any other vertebrate globally.

Of the 635 amphibian species in Ecuador, over 500 are frogs (amphibiaweb.org). Existing on most continents, frogs, members of the order *Anura*, are indicator organisms, which serve as vital gauges of the health and stability of the ecosystems they reside in.

Due to the gas and biomolecular exchange processes that occur through the thin membranes of their skin, frogs are uniquely susceptible to small changes in their environments. Their exposed, permeable skin readily absorbs toxins and other compounds from contact with the environment around them (Blaustein and Bancroft, 2007). Their unique life histories can also expose them to changes in both the aquatic and terrestrial portions of their habitat (Blaustein and Bancroft, 2007).

Thus, frogs can give researchers an idea of how an ecosystem will respond to contaminants and changes based on variation in their populations before other animals in the ecosystem are impacted. Considering how intimately frogs are tied with their habitat, the health of these organisms in threatened ecosystems, such as rainforests, is critical if we want to preserve such valuable resources. However, recent years have shown major declines in frog populations world-wide, with declines especially prevalent in South American countries, including Ecuador. If frogs are indicators of environmental response to change, this decline could have dire implications for the health of tropical rainforests at large.

Anthropogenic environmental destruction is a major contributor to declining amphibian populations around the globe. Because frogs conduct many biological processes through their skin, they require optimal conditions in their habitats to survive. Climate change, habitat loss, and increased use of pesticides and other chemicals that seep into the surrounding ecosystem have detrimental effects on frog populations. However, there is one threat to frog populations that has increasingly become of concern. Chytridiomycosis, a fungal disease caused by the chytrid fungus *Batrachochytrium dendrobatidis*, or *B.d.*, is responsible for the declines in over 500 amphibian species

(Scheele, et al. 2019). *B.d.* colonizes and feeds on the keratin in the skin of frogs. Because frogs conduct physiological processes through their skin, they are uniquely susceptible to infection. *B.d.* hinders frogs' ability to conduct these physiological processes, ultimately leading to death (Rosenblum, et al. 2010).

It is important that we understand the host-pathogen dynamics at play between *Anurans* and *B.d.* in order to potentially prevent further declines and extinctions. This dynamic describes which organism, host or pathogen, drives the infection, which can inform researchers of whether some aspects of the hosts contribute to greater rates or different presentations of infection. To characterize this dynamic and potentially mitigate the spread of infection, we must first be able to accurately identify the hosts. This is made challenging due to the sheer number of species that exist within amphibian genera. For example, the *Pristimantis* genus contains over 500 described species (Santiago, et al. 2020). Cryptic diversity, instances of distinct species with identical morphology, also presents a major roadblock to accurately determining host identity.

Although taxonomy has been historically determined by morphology, growing evidence from genetic studies suggests that this is often an inaccurate approach to determine taxon identity due to the high rates of cryptic diversity in under sampled regions, such as the tropics (Beheregary and Caccone, 2007). Cryptic diversity results from unrelated species living in similar environments that are difficult to distinguish due to their convergence towards very similar morphologies. According to Beheregary and Caccone (2007), cryptic species are distinct species that are morphologically identical, resulting in their classification as a single taxon. Typically, cryptic species exist within similar habitats, indicating that the occurrence of cryptic diversity is due to allopatric

speciation and niche conservatism (Guayasamin, et al. 2017). Because we can't rely on morphology to distinguish between species in these instances, genetic sequencing is the most accurate measure of the host's taxonomic identity.

In Eukaryotic organisms, deoxyribonucleic acid (DNA) can be placed into one of two categories: nuclear DNA (nucDNA), which resides in the nucleus, or mitochondrial DNA (mtDNA), which is located within the mitochondria. While nucDNA is linear and often very long, ~1 to 10 billion base pairs (bp), mtDNA is circular and much shorter, only about ~17 kb in amphibians (Cabaña, et al. 2020). Unlike nucDNA, which is passed on to an organism by both parents, mtDNA is largely maternally inherited (Merheb, et al. 2019). Therefore, mtDNA does not undergo the recombination that nucDNA does, resulting in an exact copy being passed from maternal parent to offspring. mtDNA generally has a higher mutation rate compared to nucDNA, making it a useful tool for determining the relationships between organisms that have recently diverged from a common ancestor (Rubinoff and Holland, 2005). These unique aspects of mtDNA allow researchers to use it as a means of discovering the exact lineage of an organism, allowing us not only to uncover host identity but also to more accurately examine the evolutionary relationships between species (Rubinoff and Holland, 2005).

Traditional sequencing methods, such as Sanger sequencing, or next-generation methods like massively parallel sequencing (MPS), could be used to sequence the entire mitochondrial genome (mitogenome) but are not always optimal. mtDNA exists in a sample at 0.1% of the total DNA present despite there being potentially tens of thousands of mtDNA copies within a cell (Robin and Wong, 1988). Thus, amplification of the mtDNA is required for Sanger sequencing. This can be achieved through Polymerase

Chain Reaction (PCR); however, PCR introduces the chance for PCR error and PCR bias. PCR also relies heavily on the efficacy of the primers used. This can be problematic for mtDNA amplification due to the existence of nuclear mitochondrial DNA (NUMTs), which is the transposition of mtDNA into the nuclear genome (Dhorne-Pollet, et al. 2020). The primers used in PCR may target the NUMTs in addition to sections of the mitogenome, confounding the sequencing data. MPS can be used for whole-genome sequencing, but short-read sequencing can also be hindered by NUMTs (Dhorne-Pollet, et al. 2020). Additionally, both Sanger and massively parallel sequencing are impractical for sequencing of organisms found in remote environments. Both methods require more equipment than is typically found in a field lab, and the storage and shipping of samples can introduce risk of contamination and considerable time to the process.

Emerging Oxford Nanopore Sequencing Technology (ONT) provides a more convenient alternative to traditional sequencing methods. Nanopore MinION sequencers are small, portable DNA long-read sequencers that perform real-time sequencing *in situ*. These sequencers use flow cells that contain 1200 pores corresponding to an electrode connected to a sensor that measures the electric current flowing through a pore. As molecules pass through the pore, the current is disrupted, and the sequencer characterizes this disruption as an individual base. This method of base-calling allows for much longer reads than traditional sequencing methods, resulting in the ability of nanopore sequencers to capture whole mitogenomes in a single read (Dhorne-Pollet, et al. 2020). It also allows for the collection of real-time sequence data, which no other sequencing method can provide. The portability, availability of real-time data, and long-reads make nanopore sequencing a convenient and possibly more useful method of sequencing mtDNA.

While nanopore sequencing has several advantages over other sequencing methods, there are a few drawbacks to consider. Compared to short-read sequencing, long-read sequencers have higher error rates, anywhere from 5% to 20% higher depending on multiple factors (Kono and Arakawa, 2019). While this error rate can be lowered with optimized library preparation and type of input, it is still a concern. Nanopore sequencers are also more prone to insertion and deletion errors, making it difficult to use them to look at single nucleotide variation (Kono and Arakawa, 2019). The utilization of voltage potential in base-calling can also result in homopolymers, or strings of the same nucleotide in a section of the genome where it doesn't exist in that quantity. However, due to the high volumes of reads that can be obtained with nanopore sequencers, adequate error-correction can be performed to mitigate the higher rates of error.

Nanopore sequencing can be used in tandem with PCR and universal primer pairs to quickly sequence regions of interest. Universal primer pairs are primer sets based on conserved regions in the mitogenomes of a wide range of organisms. These primers can be used on multiple species to target the same desired regions of the mitogenome for amplification in PCR. One such primer pair has been designed to target a region encompassing the 12S, 16S, and ND1 segments in the mitogenome of amphibians. When used in PCRs, this pair consistently generates amplicons of ~3400 bp in several types of organisms, including caecilians, salamanders, anurans (including those in the *Pristimantis* genus), and even lizards.

However, according to our preliminary data, when applied to the frog species *Pristimantis achatinus*, the same primer pair produced a fragment only ~900 bp in length.

The failure of this primer pair to produce a complete amplicon in *P. achatinus* could putatively imply rearrangement in the mitogenome of *P. achatinus* within this 12S-16S-ND1 region. There are several types of rearrangements that can occur in the mitogenome: an inversion, which is the 180-degree flip of a gene; a translocation, which is the movement of a gene; a deletion, or the removal of a gene; and a duplication, which is the duplication of a gene already present in the genome (Baris, et al. 2013). Preliminary sequence data from the ~900 bp amplicon fails to match any known *Pristimantis* sequence for this region of the mitogenome. When uploaded to MITOS, a web server that predicts gene arrangements based on sequence data, the server produced an arrangement of tRNAs unlike any of the other amphibian sequences (**Fig. 1**). Further sequencing of the entire mitogenome is necessary to verify these arrangements and to obtain data for the remainder of the 12S-16S-ND1 region in *P. achatinus*.

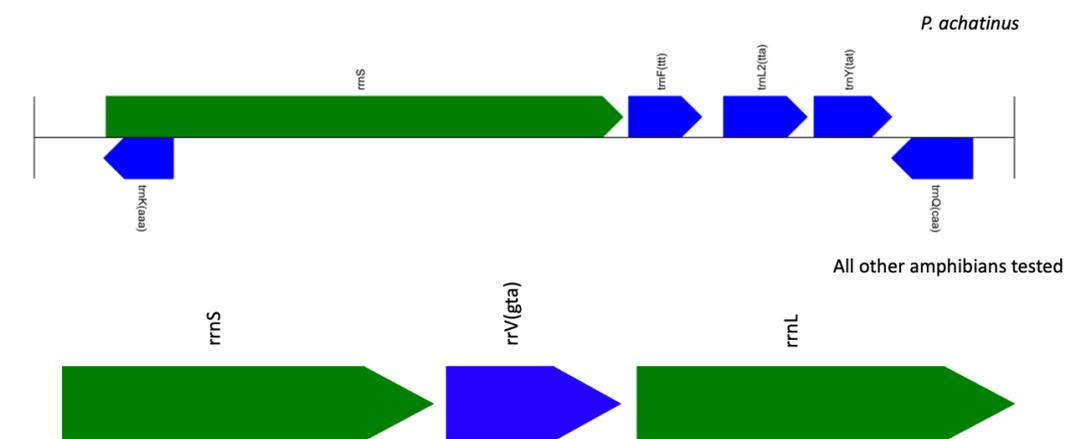


Figure 1. The predicted arrangement of the 12S region of the *P. achatinus* mitogenome generated by MITOS (top) is compared to the 12S region of other amphibians we tested (bottom).

P. achatinus presents a perfect opportunity to test ONT on an organism exhibiting signs of cryptic diversity. Following the protocol laid out in the Dhorne-Pollet, et al. (2020), I intend to use nanopore sequencing to retrieve the entire mitogenome of *P.*

achatinus. To this end, I will be testing two main approaches: (1) sequencing with a treatment protocol wherein nucDNA is depleted and mtDNA is amplified and (2) sequencing without any treatment or alteration to the template-DNA (tDNA) sample. These tests will be conducted on *P. achatinus* samples from two locations in Ecuador: the Jama-Coaque Reserve (0.1081° S, 80.1177° W) and Mashpi Reserve (0.1659° S, 78.8776° W). Once the sequence data are obtained, the mitogenome will be assembled *de novo*, given the absence of suitable sequence data on GenBank that can serve as reference material for *P. achatinus*. Because nanopore sequencers provide long-reads, a *de novo* assembly should be easier with it than alternative sequencing methods.

Obtaining sequence data for the mitogenome of *P. achatinus* will help develop new universal primer pairs to generate longer sequences. New primer pairs, combined with optimized nanopore sequencing protocols, should make it more feasible to accurately determine the identity of previously uncharacterized tropical amphibians. If applied to a wide range of amphibians, this approach could be useful in revealing more accurate phylogenies, allowing us to better describe frog species. Overall, this will be a step towards being able to determine the host-pathogen dynamics between amphibians and *B.d.* through the lens of host-identity, bringing us closer to protecting a vulnerable class of organisms and their valuable ecosystems.

II. Methods and Materials

The amplification and sequencing process was conducted using 10 nanograms (ng) of DNA per microliter (μl) of water. To select which samples would be used in the sequencing trials, a list was compiled of all *P. achatinus* tDNA samples stored in our lab. Those that were successfully amplified as identified via gel electrophoresis were

quantified using a Qubit on the high sensitivity dsDNA setting with 2 μ l of input material. The starting quantities for samples used in the tests are shown in **Table 1**.

Table 1. Starting quantities for tDNA samples.

Sample ID	DNA Concentration (ng/ μ l)
BB051	9.31
BB056	15.5
BB063	10.4
M025	9.64
M037	9.21
M085	10.1

1) Primer Testing and Sanger Sequencing

Prior to beginning the sequencing trials, I tested three different reverse primers to determine whether we could detect what rearrangements had occurred. Because the initial ~900 bp was obtained with the original primer set, it was hypothesized that the rearrangement had occurred within the 16S gene, preventing the original reverse primer from annealing to the sequence correctly. To determine this, I tested: (1) a reversed primer, (2) a complemented primer, and (3) a reversed-complemented primer on four *P. achatinus* samples.

The reversed primer consisted of a flipped sequence of the original reverse primer. Applying this set in a PCR would result in a longer amplicon if an inversion of the tRNA in that area of the sequence had indeed occurred. The complemented primer consisted of the complementary sequence to the original reverse primer, which would result in a larger amplicon if the rearrangement was a translocation. The reversed-complemented primer consisted of a flipped complementary sequence to the original reverse primer, which would target both a translocation and inversion. Should any of

these rearrangements have occurred in this area of the mitogenome, the resultant PCR amplicon would be expected to be ~3400 bp.

The PCRs were carried out in 25 µl reactions, which were conducted on four *P. achatinus* samples – two from the Jama-Coaque Reserve (BB051, BB063) and two from Mashpi (M025 and M085). The setup for the reactions is depicted in **Table 2**.

Table 2. PCR reaction for primer testing of rearrangements.

Volume (µl)	Reagent
1.0	tDNA
12.5	Dream Taq Master Mix
0.25	12SAL Primer
0.25	Reverse Primer (original, reversed, complemented, or reverse-complemented)
1.0	MgCl ₂ +
10.0	Nuclease-free water

PCRs were run using a long protocol (3.5 hours) with a 46° C annealing temperature. Gel electrophoresis, using a 1% agarose gel and TBE, was conducted on PCR products to determine amplicon sizes.

The products from the original, reversed, and reversed-complemented PCRs for each sample were run through a second 1% agarose TBE gel for the purposes of gel extraction. 15 µl of product was loaded into the gel with 3 µl of loading dye. The gel was set to run at 85 volts until bands were separated enough for extraction. The bands were visualized using a UV light and excised from the gel with a razor, then digested and cleaned using a GENEjet gel extraction kit by ThermoFisher Scientific.

After cleanup, the products were cycle sequenced and used in traditional Sanger sequencing. The reaction set up for the cycle sequence is shown in **Table 3**.

Table 3. Reaction set-up for cycle sequencing of gel-extracted amplicons.

Amount (µl)	Reagent
2.0	Gel extraction product

2.28	Nuclease-free water
1.0	Big Dye 3.1
0.60	Big Dye Buffer
0.12	Primer (forward or reverse)

2) Nanopore Run Without Treatment

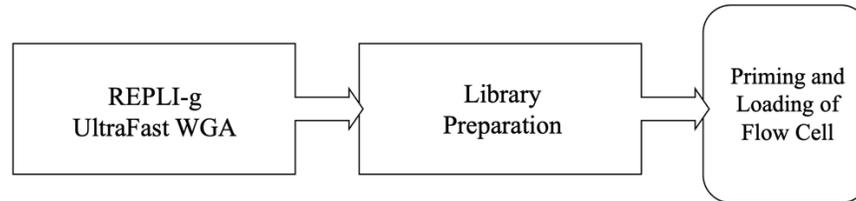


Figure 2. Workflow for nanopore run without treatment

For the nanopore run without treatment, one *P. achatinus* tDNA sample (BB063) was taken from the group of high concentration samples. This sample was whole genome amplified and library prepped according to the Nanopore Protocol for rapid whole genome amplification (Oxford Nanopore Technologies). The first step in this process is a denaturation step, in which all the DNA in the sample is broken into single stranded DNA. Subsequently, a specialized polymerase can be added to amplify all DNA in the sample. This elongation was allowed to take place for 2 hours, until the DNA concentration was >80 ng/ μ l.

Following amplification of all DNA in the sample, library preparation took place. A fragmentation mix was added to the amplified DNA, which enzymatically cleaved the double stranded DNA, leaving an overhanging base that the Rapid adapter could then ligate to. This adapter attaches the DNA to the flow cell, pulling it through the nanopores so that the voltage of the bases can be read.

After library prep, the flow cell was primed, and the sample was loaded dropwise according to protocol. Once loaded, the sequencer was set to run for 24 hours.

3) Nanopore Run with Whole Mitochondrial Genome Amplification

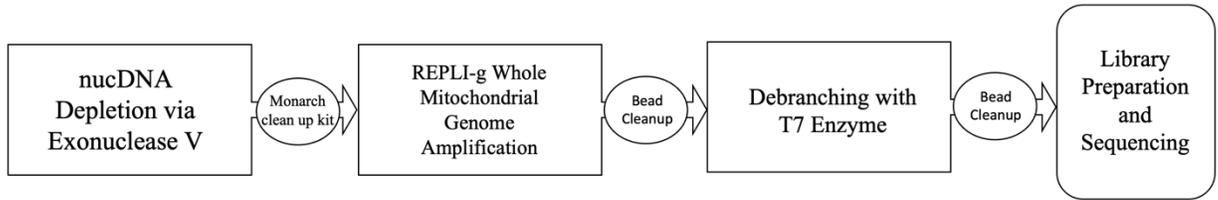


Figure 3. Workflow for nanopore run with whole mitochondrial genome amplification Treatment.

Following the protocol developed by Dhorne-Pollet, et al. (2020), we first used an Exonuclease V (ExoV) reaction to deplete the nucDNA concentration from our samples (BB056, BB063, M025, and M085), leaving behind mtDNA and trace amounts of nucDNA. Exonuclease V is an enzyme that degrades single stranded DNA (ssDNA) as well as linear DNA, which would effectively diminish the linear nucDNA concentration, leaving behind only the circular mtDNA. Following the suggestions of Dhorne-Pollet, et al. (2020), we ran a 2-hour incubation with 10 units of Exonuclease V. However, we ran a 100 μ l reaction as opposed to 50 μ l.

This nucDNA depletion step was necessary to sequence just the mitogenome without interference from the nuclear genome and possible NUMTs. Following the ExoV reaction, the samples were purified using a Monarch PCR & DNA Cleanup Kit (New England Biolabs inc.), following standard protocol. After purification, all four samples were quantified using a Qubit (ThermoFisher Scientific) on the dsDNA high sensitivity setting with 2 μ l of input.

Table 4. Quantities of mtDNA after Monarch purification.

Sample ID	mtDNA Concentration (ng/ μ l)
BB056	1.24
BB063	21.5
M085	14.8
M025	14.0

After depleting the nucDNA from our samples, we whole-genome amplified the mtDNA using the REPLI-g Mitochondrial DNA Kit (Oxford Nanopore Technologies). We added 5 μ l of each of our ExoV treated mtDNA samples to separate tubes along with 15 μ l of RNase-free water. The amplification mix was prepared as specified by the QIAGEN REPLI-g Mitochondrial DNA Kit with the exception the primer mix used, which was swapped out for amphibian-specific primers (Rodriguez, unpublished).

Following whole mitochondrial genome amplification (WMGA), the samples were quantified to check the amount of mtDNA in ng/ μ l using a Qubit on the high sensitivity dsDNA setting (**Table 6**):

Table 5. mtDNA concentrations following WMGA

Sample ID	mtDNA Concentration (ng/ μ l)
BB056	0.260
BB063	2.77
M085	1.82
M025	2.07

Although concentrations were lower than expected, we proceeded with AMPure XP bead purification utilizing a half volume of beads per volume of sample (Beckman Coulter Life Sciences). Following clean up, concentrations of the samples were taken again using the same parameters as used previously.

Table 6. mtDNA concentrations following AMPure XP bead purification

Sample ID	mtDNA Concentration (ng/ μ l)
BB056	0.304
BB063	3.40
M085	2.05
M025	2.24

After WMGA, we used T7 Endonuclease I (T7EI) (New England Biolabs inc.) in a debranching step. The samples were then quantified once more using the same parameters as before.

Table 7. mtDNA concentrations following debranching.

Sample ID	mtDNA Concentration (ng/ μ l)
BB056	Too low
BB063	1.67
M085	1.25
M025	1.43

This concentration of mtDNA was insufficient to load onto the MinION, and a different form of enrichment was required.

4) Nanopore Run with REPLI-g UltraFast WGA

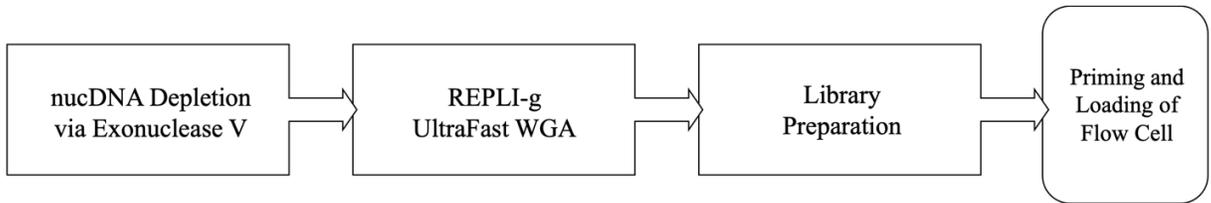


Figure 4. Workflow for nanopore run with REPLI-g UltraFast whole genome amplification

Utilizing the left-over product from the previous ExoV reactions, we performed a whole genome amplification (WGA) following the ONT rapid whole genome amplification protocol (Oxford Nanopore Technologies). After the ExoV reaction, the steps in this process were the same as the steps in the nanopore run without treatment. Following WGA, concentrations of mtDNA were taken as described previously.

Table 8. mtDNA concentrations following WGA.

Sample ID	mtDNA Concentration (ng/ μ l)
BB056	19.4
BB063	23.1
M085	33.1
M025	17.5

With concentrations high enough for sequencing, we proceeded with library preparation, following the same process as described in the nanopore run without treatment. However, we chose only to use samples with the highest mtDNA concentrations, namely BB063 and M085. Following library prep, the flow cell was primed, and the samples were loaded into the sample port following the same protocol as the nanopore run without treatment. The MinION was set to run for 24 hours.

5) Nanopore Run with Magnetic Bead Capture

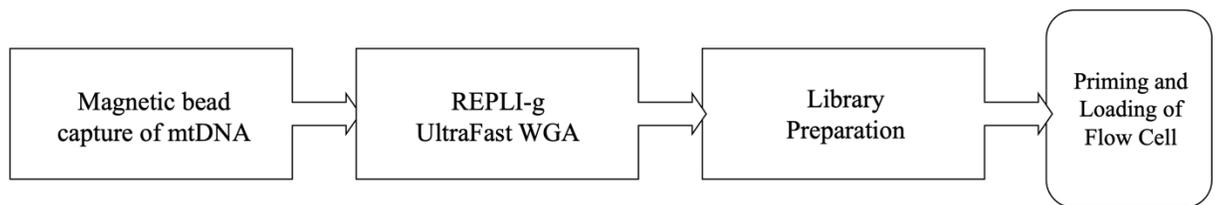


Figure 5. Workflow for nanopore run with magnetic bead capture of mtDNA

Following a protocol developed by Bardan (2019), we attempted an extraction of the mtDNA in our samples with specialized probes and magnetic beads, as opposed to degrading the nucDNA with ExoV. For this trial, we used samples BB037, BB058, M080, and M085, as well as a positive and negative control.

We hybridized the biotinylated 16S amphibian probes in the DNA samples and introduced the coated beads into the sample. After exposing the sample to a magnet, we extracted the supernatant containing the nucDNA, which was stored in separate tubes for use in gel electrophoresis verification. Bead decoupling and elution of mtDNA was performed according to Bardan (2019). The full separation of nucDNA from mtDNA was confirmed with PCR using mitogenome-specific primers and gel electrophoresis. The mtDNA product was quantified before moving on to WGA (**Table 9**).

Table 9. Concentration of mtDNA after magnetic bead capture.

Sample ID	mtDNA Concentration (ng/ μ l)
BB037	0.67
BB058	0.68
M080	0.62
M085	0.72
Em (pos. control)	0.64
n.f.H2O (neg. control)	0.72

Following quantification, the mtDNA samples were amplified following the rapid whole genome amplification protocol as described previously. After WGA, the quantities were taken again for all samples and both controls.

Table 10. mtDNA concentrations following magnetic WGA.

Sample ID	Concentration (ng/ μ l)
BB037	1.28
BB058	0.878
M080	4.52
M085	4.56
Em (pos. control)	0.358
n.f.H2O (neg. control)	0.172

The samples were library prepped and loaded onto the MinION, and sequencing was set to run for 24 hours.

6) Nanopore Run with Mechanical Shearing

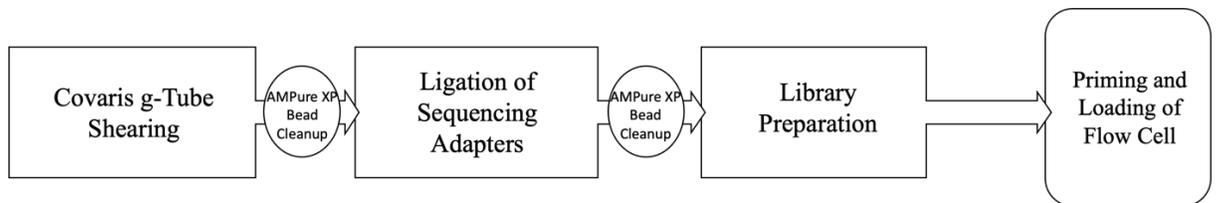


Figure 6. Workflow for nanopore run with mechanical shearing of DNA.

Rather than shear the DNA enzymatically for this trial, we decided to take a mechanical approach using the Covaris g-TUBE shearing protocol 520079 (Covaris) for 6 kilobase (kb) fragments. We loaded sample BB037 into a Covaris g-TUBE and spun it

down through a matrix that sheared all DNA strands into 6 kb fragments. For this trial, we did not deplete the nucDNA from the sample. Purification was performed using AMPure XP beads (Beckman Coulter Life Sciences).

We then followed the Nanopore protocol for sequencing by ligation (Oxford Nanopore Technologies) to ligate the sequence adapters to the sheared DNA. This process works by enzymatically repairing the ends of the sheared DNA and adding an A-tail that the sequence adapters can attach to. These adapters are what attaches the DNA to the flow cell, allowing the nucleotides to be pulled through the pores. After ligation, another purification step was carried out using AMPure XP beads (Beckman Coulter Life Sciences). The samples were then loaded onto the flow cell and sequencing was set to run for 24 hours on the sequencing by ligation setting.

The sequence data were then transported from the MinION to Geneious for visualization and mapping. 25-30 bp were trimmed from the ends of the sequences using Trimmomatic. Then, the sequence data was aligned to amphibian reference sequences for COI, 12S, CytB, 16S, and the control region in Geneious. Following alignment, the reads were exported as fastq files. The aligned reads were error corrected through Canu, and matching reads were assembled with a target size of 21 kb. The corrected reads and unitigs were then exported to Geneious where they were assembled de novo with the Geneious assembler. The consensus was generated and exported as a fasta file to MITOS.

II. Results

1) Primer Testing and Sanger Sequencing

The reversed primer yielded amplicons sizes of ~1500 bp. The complemented primer yielded amplicon sizes of ~1000 bp. The reversed-complemented primer yielded amplicon sizes of ~2000 bp. (Fig. 7)

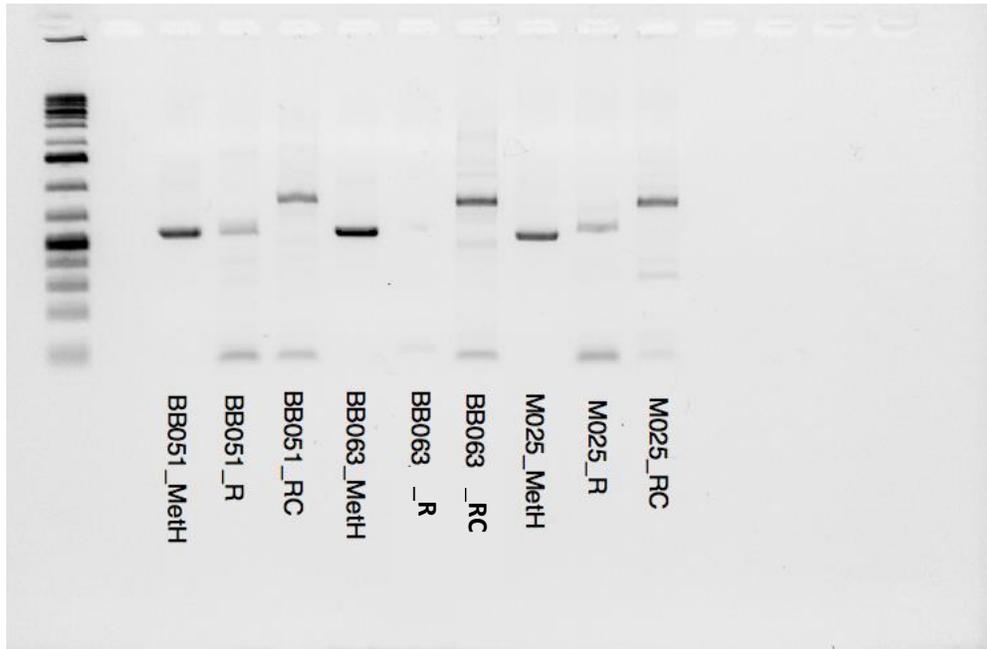


Figure 7. Gel electrophoresis bands for the primer test reveals, from left to right: ladder, 900 bp bands for the original primer set (_MetH), 1000 bp bands for the reversed (_R) primer, 900 bp bands for the complemented primer (_C), and 1500 bp bands for the reversed-complemented primer (_RC) consistent across all samples.

The Sanger sequencing reads were short and low quality as shown in Fig. 8.

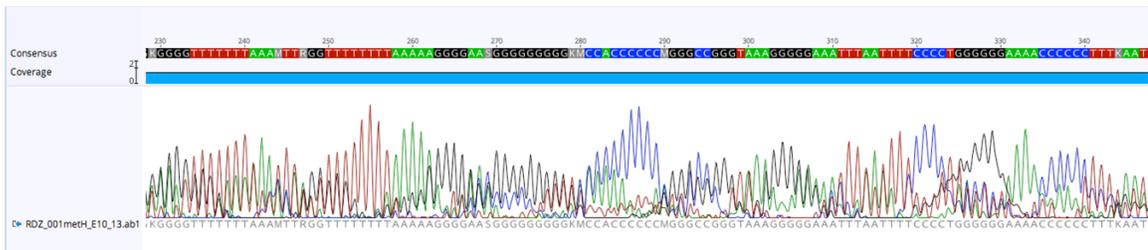


Figure 8. Overlapping base signals visualized using Geneious.

2) Nanopore Run Without Treatment

Due to a power outage, the sequencing run terminated prematurely, resulting in an inadequate number of reads.

3) Nanopore Run with WMGA

Concentrations of the mtDNA were too low to utilize in sequencing for this trial.

4) Nanopore Run with WGA

The mtDNA concentrations of mtDNA for the WGA trial were high enough to utilize in sequencing. This sequencing trial ran on the MinION for the whole 24-hour duration, but the sequence data failed to align to a reference.

5) Nanopore Run with Magnetic Bead Capture

mtDNA concentrations following the magnetic bead capture were low. Following WGA, the two samples from the Jama-Coaque Reserve were 1.28 ng/ μ l and 0.878 ng/ μ l respectively, while the Mashpi samples were both \sim 4.50 ng/ μ l. The data obtained from the 24-hour sequencing run did not align to a reference.

6) Nanopore Run with Mechanical Shearing

The MinION showed reads averaging 5 kb in length soon after sequencing commenced. The data from the 24-hour sequencing run were adequate for assembly. The predicted gene arrangements for *P. achatinus* revealed a switch in the placement of the 12S and 16S sections. There were also two potential duplications of the D-loop, or control region, one occurring between 16S and 12S, and another between 12S and ND1. There were also multiple duplications of the met-tRNA (**Fig. 9**).

suspected of harboring rearrangements, an approach to amplification not relying on specific primers is likely to be more effective for sequencing.

The WGA approach yielded much higher mtDNA concentrations than the WMGA method. This approach to amplification is less specific than WMGA. Instead of targeting the mitogenome via mtDNA-specific primers, REPLI-g UltraFast WGA amplifies all DNA within a sample. Because of this, the nucDNA depletion step was especially necessary to avoid amplification of the nuclear genome in addition to the mitogenome. Were the nucDNA not depleted, it is possible the frequency of NUMTs would be higher, resulting in more difficulty discerning between the mitogenome and nuclear genome.

The probes developed by Bardan (2019) were attached to a biotin molecule, which can hybridize to the desired DNA in a sample. The hybridized probes can then be exposed to streptavidin-coated magnetic beads, which couple with the biotin molecule, allowing for the capture and removal of the attached mtDNA when exposed to a magnet (Bardan, 2019). Although the magnetic beads should have captured the entire mitogenome from the samples, the reads from the sequencer were much shorter on average than expected, rendering them ineffective for assembling the mitogenome. The ability of the two amphibian 16S probes to anneal to the mitogenome may be impacted by rearrangements, resulting in only partial recovery of the mitogenome from the sample. This is supported by the low concentration of mtDNA following bead capture and amplification. It is possible this approach could introduce the same problems with the NUMTs as previous trials that did not include an enzymatic nucDNA depletion step. However, this problem would be to a lesser extent, as only 16S primers were used. With

primers meant to bind to only one region of the mitogenome, the chance of NUMTs being pulled is lessened.

One occurrence of note in the magnetic bead capture trial was the discrepancy between the Jama-Coaque Reserve samples and the Mashpi samples. While the Jama-Coaque samples had an mtDNA concentration of around 1.0 ng/ μ l following magnetic bead capture and amplification, the Mashpi samples were both around 4.0 ng/ μ l. This suggests that the probes were more successful at pulling out the mtDNA for the Mashpi samples as opposed to the Jama-Coaque samples. This could be due to a single-nucleotide polymorphism (SNP), which is a difference in a single base pair between individuals (Marth, et al. 1999). SNPs are the most prevalent type of genetic variation and exist within members of the same species (Marth, et al. 1999). A SNP could impact the way the probes bind to the mtDNA in a similar way that rearrangements impact the primer ability to bind. Potentially, this discrepancy could be an instance of cryptic diversity, given that the organisms these samples were collected from were geographically isolated. However, it is likely the divergence is not large enough to categorize them as distinct species. More sampling should be done on *P. achatinus* in these locations to determine if different arrangements exist between their mitogenomes.

There are only a few partial sequences of other *Pristimantis* in GenBank, and *P. achatinus* is not similar enough to other amphibians to utilize their mitogenomes as appropriate references. Therefore, a de novo assembly of the reads from the mechanical shearing and ligation trial had to be performed.

The predicted gene arrangement derived from our sequence data explains the smaller amplicons in our initial PCR tests. The universal primers were meant to target the

12S, 16S, and ND1 regions of the mitogenome. In amphibians with the typical vertebrate mitogenomic arrangement, the forward primer binds in the middle of the 12S region, while the reverse primer binds just beyond ND1 on a met-tRNA. This segment of the mitogenome is ~3,400 bp. In *P. achatinus*, the forward primer anneals to the same location, but the reverse primer targets a duplication of met-tRNA just beyond the 12S region, ~900 bp from where the forward primer sits (**Fig. 9**). There are two other met-tRNAs in the mitogenome of *P. achatinus*: one just beyond the new placement of the 16S region and one at the start of the original placement of the D-loop (**Fig. 9**). Though the reverse primer can anneal to these two locations, it preferentially targets the met-tRNA closest to the forward primer, resulting in far more of the smaller fragments than the much larger sequences the other two binding spots would produce.

Mitogenomes tend to be very conserved across most vertebrate species (Zhang, et al. 2020). Neobatrachia, more recently evolved frogs, make up 96% of all anuran species and tend to experience a higher frequency of rearrangements in the mitogenome than other vertebrates (Zhang, et al. 2020). Many Neobatrachian mitogenomes follow the “typical Neobatrachia arrangement,” though deviations from this arrangement have also been reported (Zhang, et al. 2020). It is not yet known why Neobatrachia have different rates of evolution in their mitogenomes compared to other vertebrates (Irisarri, et al. 2012). Some theories attempting to explain this higher rate of evolution involve the metabolic rate and life-history of neobatrachians, as well as higher occurrences of speciation events. However, there is little statistical backing for these theories (Irisarri, et al. 2012).

Previous research by Zhang, et al. (2020) has quantified the frequency of rearrangements of genes within the amphibian mitogenome. Zhang, et al. (2020) demonstrated that the control region, the D-loop, experienced the highest frequency of rearrangement compared to other sections of the mitogenome. The research team also characterized the 12S region as having a low frequency of rearrangements and the 16S region as having a very low frequency of rearrangements. Met-tRNAs were found to have a moderate frequency of rearrangement. Based on the findings in the Zhang, et al. (2020), it is likely the translocation resulting in the switch of 12S and 16S in *P. achatinus* occurred in the 12S region, due to the slightly higher frequency of rearrangement for 12S compared to 16S. The Zhang, et al. (2020) study also found that duplications of the control region were frequent in Neobatrachia and have even been observed in other organisms, such as parrots.

The propensity for Neobatrachia to experience mitogenomic rearrangements is noteworthy when it comes to attempting to determine their identity. As mentioned, cryptic diversity introduces a major roadblock to accurately distinguishing between different species when relying on morphology alone. Thus, DNA sequencing is a necessary step when categorizing frogs and other organisms. However, the high rates of mitogenomic evolution in Neobatrachia presents a challenge to sequencing methods that include a PCR step. Like the rearrangements discovered in the mitogenome of *P. achatinus*, other anurans may exhibit rearrangements that prevent universal primers from binding to the correct spots in the mitogenome. These cases would benefit from an efficient approach to whole mitochondrial genome sequencing. The approach outlined

here would eliminate the need for specific primers and could provide more data in a timely manner than short-read sequencing.

VI. Conclusion

With global amphibian populations on the decline, it is imperative that more efforts be directed towards more accurately delineating species-boundaries in amphibian taxa. Cryptic diversity and high frequencies of mitogenomic rearrangement present challenges to distinguishing between species. However, long-read sequencing with the use of ONT provides an opportunity to sequence entire mitogenomes quickly and efficiently in the field. In doing so, characterizing amphibian species diversity should become more feasible in remote regions. These characterizations are crucial for a better understanding of the dynamics between *Anura* and *B.d.* infection rates across different species. With the majority of threatened amphibians residing in South America, this work should be prioritized for the Neotropics, as these amphibians face threats from habitat loss and *B.d.* at more pronounced rates. With a plethora of factors accelerating the global decline of amphibians, it is critical that efforts be taken to mitigate the extinction of these valuable keystone organisms.

Appendix

Exonuclease V Treatment Protocol (New England Biolabs inc.)

Add 10.0 µl of tDNA to a tube with 10 µl (1X) of NEBuffer 4 (10X), 10.0 µl (1mM) ATP (10mM), 2.0 µl (10 units) Exonuclease V, and 77.0 µl nuclease-free water. Incubate the reaction for 2 hours. Add 1.1 µl of EDTA (to a 11mM concentration). Heat inactivate the reaction at 70° C for 30 minutes.

Whole Mitochondrial Genome Amplification Protocol (QIAGEN)

Add 5 µl of ExoV treated mtDNA to 15 µl RNase-free water in a PCR tube. Prepare amplification mix by adding 27 µl of REPLI-g mt Reaction Buffer to 2 µl amphibian primer mix. Vortex and centrifuge the reaction. Add amplification mix to the DNA. Vortex and centrifuge. Incubate the reaction at 75° C for 5 minutes, then allow to cool to room temperature. Add 1 µl of the REPLI-g Midi Polymerase to the DNA mixture. Gently mix via flicking and centrifuge. Incubate the sample at 33° C for 8 hours. Inactivate the reaction by heating it to 65° C for 3 minutes.

T7EI Debranching Protocol (New England Biolabs inc.)

Add 20.0 µl of WMGA DNA into a PCR tube. Add 1.0 µl of T7EI, 2.5 µl of 10X Buffer, and 1.5 µl of nuclease-free water to the DNA. Incubate the reaction at 37° C for 2 hours.

AMPure XP Bead Clean Up Protocol (Beckman Coulter)

Shake the bottle of AMPure XP Beads to resuspend. Add a full volume of beads to the DNA and mix by pipetting. Incubate at room temperature for 5 minutes. Expose reaction to a magnet. Discard the supernatant. Add 200 µl of 70% Ethanol to the DNA and bead mixture. Incubate for 30 seconds at room temperature, then discard the

supernatant. Repeat this step once more. Add 40 μl of Elution Buffer to the DNA, mix via pipetting, and incubate at room temperature for 2 minutes. Expose the reaction to a magnet, and transfer the supernatant to a new tube.

Rapid Whole Genome Amplification Protocol (Oxford Nanopore Technologies)

In a PCR tube, add 1.0 μl of DNA to 1.0 μl of D1 buffer. Vortex the solution, and centrifuge. Incubate the reaction at room temperature for 3 minutes. Add 2 μl of N1 Buffer to the tube. Vortex the reaction and spin down. In a separate tube, add 15 μl of REPLI-g UltraFast Reaction Buffer to 1 μl of REPLI-g UltraFast Reaction Polymerase. Gently mix the reaction via flicking, and spin down. Add the DNA from the previous reaction into the second tube. Gently mix, and spin down. Incubate the reaction at 30° C for 2 hours until DNA concentration is >80 ng/ μl .

In a PCR tube, add 2.5 μl of the amplified DNA to 5 μl nuclease-free water and 2.5 μl Fragmentation Mix. Gently mix, spin down, and incubate at 30° C for 1 minute, then 80° C for 1 minute. After incubation, add 1 μl of Rapid Adapter to the amplified DNA library. Gently mixed and spin down. Incubate the reaction at room temperature for 5 minutes.

Magnetic Bead Capture Protocol (Bardan, 2019)

In a 0.2 mL strip tube, add 10 μl of tDNA. Add 15 μl of 12X SSC, 1.5 μl nuclease-free water, 1.0 μl of bovine serum albumin, 0.5 μl of probe 1, and 0.5 μl of probe 2. Denature the reaction at 98° C for 30 minutes. Then, hybridize at 50° C for 2 hours. Allow the reaction to cool to room temperature.

Add 10 μl of magnetic beads to 30 μl of hybridization reaction. Incubate the reaction at 45° C for 30 minutes in a ThermoMixer while vortexing at ~ 1400 rounds per

minute. Allow the reaction to cool to room temperature. Expose the tubes to a magnet and elute the supernatant into separate tubes labeled nucDNA. Resuspend the magnetic beads in 20 µl of Elution Buffer, vortex, and incubate at 95° C for 10 minutes. Expose the tubes to a magnet, and elute the supernatant into separate tubes labeled mtDNA.

Covaris g-TUBE Mechanical Shearing Protocol (Covaris)

Place the g-TUBE into the load position and add 50 µl of sample into the top of the tube. Centrifuge the tube at 11,000 rounds per minute for 30 seconds. Invert the tube and repeated centrifugation. Transfer the tube to the unload position and retrieve sample from the cap.

Sequencing by Ligation Protocol (Oxford Nanopore Technologies)

Add 47 µl of tDNA into a reaction tube. Add 1 µl DNA CS, 3.5 µl NEBNext FFPE DNA Repair Buffer, 2 µl NEBNext FFPE DNA Repair Mix, 3.5 µl Ultra II End-prep reaction buffer, and 3 µl Ultra II End-prep Enzyme Mix. Mix tube via flicking and centrifuge briefly. Incubate reaction at 20° C for 5 minutes, then 65° C for 5 minutes. Following incubation, perform an AMPure XP bead clean up.

Add 60 µl of the cleaned DNA reaction to a reaction tube with 50 µl Blunt TA Master Mix and 10 µl AMX. Incubate the reaction at room temperature for 10 minutes. Perform an AMPure XP bead clean up with 40 µl of beads. Incubate the beads and DNA mixture at room temperature for 5 minutes. Then, add 250 µl of Long Fragmentation Buffer x2. Discard the supernatant. Resuspend in 15 µl of Elution Buffer. Incubate at 37° C for 10 minutes and move supernatant to a new tube.

References

- AmphibiaWeb. AmphibiaWeb: Information on amphibian biology and conservation: Berkeley, California; 2020 [05/10/2020]. 2020:[Available from: <http://amphibiaweb.org/>].
- Andrew R. Blaustein, Betsy A. Bancroft, Amphibian Population Declines: Evolutionary Considerations, *BioScience*, Volume 57, Issue 5, May 2007, Pages 437-444, <https://doi.org/10.1641/B570517>
- Barış, Ö., Karadayı, M., Yanmış, D., & Güllüce, M. (2013). Genomic Rearrangements and Evolution. In (Ed.), *Current Progress in Biological Research*. IntechOpen. <https://doi.org/10.5772/55456>
- Beheregaray, L. B., & Caccone, A. (2007). Cryptic biodiversity in a changing world. *Journal of biology*, 6(4), 9. <https://doi.org/10.1186/jbiol60>
- Cabañas, N., Becerra, A., Romero, D. *et al.* Repetitive DNA profile of the amphibian mitogenome. *BMC Bioinformatics* **21**, 197 (2020). <https://doi.org/10.1186/s12859-020-3532-8>
- Daniel J. Hocking and Kimberly J. Babbitt. (2014) Amphibian Contributions to Ecosystem Services. *Herpetological Conservation and Biology* 9(1):1–17.
- Daniel Rubinoff, Brenden S. Holland, Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference, *Systematic Biology*, Volume 54, Issue 6, December 2005, Pages 952-961, <https://doi.org/10.1080/10635150500234674>
- Guayasamin JM, Hutter CR, Tapia EE, Culebras J, Peñafiel N, Pyron RA, et al. (2017) Diversification of the rainfrog *Pristimantis ornatissimus* in the lowlands and

Andean foothills of Ecuador. PLoS ONE 12(3): e0172615.

<https://doi.org/10.1371/journal.pone.0172615>

Ina Vandebroek, Patrick Van Damme, Luc Van Puyvelde, Susana Arrazola, Norbert De Kimpe, A comparison of traditional healers' medicinal plant knowledge in the Bolivian Andes and Amazon, *Social Science & Medicine*, Volume 59, Issue 4, 2004, Pages 837-849, <https://doi.org/10.1016/j.socscimed.2003.11.030>.

Irisarri, I., Mauro, D.S., Abascal, F. *et al.* The origin of modern frogs (Neobatrachia) was accompanied by acceleration in mitochondrial and nuclear substitution rates. *BMC Genomics* **13**, 626 (2012). <https://doi.org/10.1186/1471-2164-13-626>

IUCN. The IUCN Red List of Threatened Species. Version 2021–1. 2021 [04/09/2021]. Available from: <https://www.iucnredlist.org>.

Kono, N, Arakawa, K. Nanopore sequencing: Review of potential applications in functional genomics. *Develop Growth Differ.* 2019; 61: 316– 326. <https://doi.org/10.1111/dgd.12608>

Marth, G., Korf, I., Yandell, M. *et al.* A general approach to single-nucleotide polymorphism discovery. *Nat Genet* **23**, 452–456 (1999). <https://doi.org/10.1038/70570>

Maude, H., Davidson, M., Charitakis, N., Diaz, L., Bowers, W. H., Gradovich, E., Andrew, T., & Huntley, D. (2019). NUMT confounding biases mitochondrial heteroplasmy calls in favor of the reference allele. *Frontiers in Cell and Developmental Biology*, 7. <https://doi.org/10.3389/fcell.2019.00201>

Merheb, M., Matar, R., Hodeify, R., Siddiqui, S. S., Vazhappilly, C. G., Marton, J., Azharuddin, S., & Al Zouabi, H. (2019). Mitochondrial DNA, a Powerful Tool to

Decipher Ancient Human Civilization from Domestication to Music, and to Uncover Historical Murder Cases. *Cells*, 8(5), 433.

<https://doi.org/10.3390/cells8050433>

Ortega-Andrade HM, Rodes Blanco M, Cisneros-Heredia DF, Guerra Arévalo N, López de Vargas-Machuca KG, Sánchez-Nivicela JC, et al. (2021) Red List assessment of amphibian species of Ecuador: A multidimensional approach for their conservation. *PLoS ONE* 16(5): e0251027.

<https://doi.org/10.1371/journal.pone.0251027>

Phillips, O.L., Brienen, R.J.W. & the RAINFOR collaboration. Carbon uptake by mature Amazon forests has mitigated Amazon nations' carbon emissions. *Carbon Balance Manage* **12**, 1 (2017). <https://doi.org/10.1186/s13021-016-0069-2>

Robin, E. D., & Wong, R. (1988). Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *Journal of cellular physiology*, 136(3), 507–513. <https://doi.org/10.1002/jcp.1041360316>

Ron SR, Carrión J, Caminer MA, Sagredo Y, Navarrete MJ, Ortega JA, Varela-Jaramillo A, Maldonado-Castro GA, Terán C (2020) Three new species of frogs of the genus *Pristimantis* (Anura, Strabomantidae) with a redefinition of the *P.*

lacrimosus species group. *ZooKeys* 993: 121

155. <https://doi.org/10.3897/zookeys.993.53559>

Rosenblum EB, Voyles J, Poorten TJ, Stajich JE (2010) The Deadly Chytrid Fungus: A Story of an Emerging Pathogen. *PLoS Pathog* 6(1): e1000550.

<https://doi.org/10.1371/journal.ppat.1000550>

Scheele, B. C., Pasmans, F., Skerratt, L. F., Berger, L., Martel, A., Beukema, W., Acevedo, A. A., Burrowes, P. A., Carvalho, T., Catenazzi, A., De la Riva, I., Fisher, M. C., Flechas, S. V., Foster, C. N., Frías-Álvarez, P., Garner, T. W., Gratwicke, B., Guayasamin, J. M., Hirschfeld, M., ... Canessa, S. (2019). Amphibian fungal panzootic causes catastrophic and ongoing loss of biodiversity. *Science*, 363(6434), 1459–1463. <https://doi.org/10.1126/science.aav0379>

Zhang, J., Miao, G., Hu, S., Sun, Q., Ding, H., Ji, Z., Guo, P., Yan, S., Wang, C., Kan, X., & Nie, L. (2021). Quantification and evolution of mitochondrial genome rearrangement in Amphibians. *BMC ecology and evolution*, 21(1), 19. <https://doi.org/10.1186/s12862-021-01755-3>