Detecting Locally Extinct Mesic and Riparian Species in an Urban Forest

Leandre Aguilar[1], Kenya Davis[1], Nicholas Valdez[1], Zachary Vetsch[1]
Mentors: Alison Granberry[1], Damon P. Little[2]
[1]Hostos Lincoln Academy of Science [2]New York Botanical Garden

Funded by the Thompson Family Foundation

Abstract

Environmental DNA (eDNA) focuses on the extraction and analysis of DNA from environmental samples such as soil, water, and air. By using eDNA metabarcoding, more species can be detected in an area than if other methods were used. A 2015 inventory of the old–growth forest at the New York Botanical Garden found that five fern species present in 1899 had gone locally extinct. We planned to compare eDNA sequence data to the 1899 and 2015 inventories to identify any differences. Ultimately, we hoped to learn if the forest is continuing to decline or recovering—evident by the reemergence of locally extinct fern life. We were able to design and test three specific primer sets targeting diagnostic positions in *Botrychium*, *Equisetum*, and *Dryopteris*. PCR amplification of eDNA was, however, unsuccessful.

Introduction

Environmental DNA (eDNA) metabarcoding is a powerful tool used to detect species within communities. The method involves the extraction and analysis of DNA from environmental samples such as soil, water, and air. It is particularly useful in cases where certain species are not visible to the naked eye—making traditional survey methods impractical. The ability to detect a wide range of species from a single collection has made eDNA metabarcoding a valuable tool for conservation and management. In a recent review (Banerjee et al., 2022), eDNA methods were shown to be more time and cost efficient than other methods: more species can be detected in an area than if other methods were used.

A recent study collected airborne eDNA to document the seasonal biodiversity of airborne flora and funga in northern Italy (Banchi et al., 2020). The study took into account both site location and time because the release of pollen and spores by plants changes by time of year. In order to find locally extinct species, we sampled in late fall, when ferns produce spores, so there should have been many airborne spores to collect in addition to eDNA from vegetative sources. Applying airborne eDNA technology to the study of plant biodiversity will allow us to contribute towards techniques for rapid biodiversity monitoring.

The New York Botanical Garden's (NYBG) old growth forest is 20 hectares and contains rapirian (Bronx River) and mesic habitats. It is a broadleaf deciduous forest that is home to approximately 429 extant species (Atha et al., 2016). A 2015 inventory determined that the fern species *Botrychium dissectum, Dryopteris cristata, D. intermedia, Equisetum fluviatile,* and *E. hyemale* which were present in 1899 had gone locally extinct as of 2015. We collected airborne eDNA from the forest in hopes of finding these locally extinct ferns. In order to identify the eDNA of these species, we located unique diagnostic positions within their genomes, designed primer sets targeting these positions, and tested the primers with known DNA samples.

Table 1. Novel PCR primer sets designed for this study.

Genus	Marker	T_m (°C)	Left primer	Right primer
Botrychium	*trnL–trnF*	52	TCCTATACTGGATGAC ACTAGCAT	ATGACCCCATAATCCAT TCGAGA
Dryopteris	*matK*	49	TAGATATCAACCACT GCCTCATCC	ACAACTGGTAGATAAC AAATCGCT
Equisetum	*rbcL*	53	AGATGATTCTGTATTA CAATTTGGTGG	TCCATTTAGCAGCTTCA CGAATAA

Fig 1. eDNA extraction.

Table 2. Diagnostic *Botrychium dissectum trnL–trnF* SNPs.

	n	3134569091347452456880135623565755789469424682361262945623634731281458899026879457
B. angustisegmentum	7	GNNAGNCNNNNNNNNNNNNGGTCAATGGAGGCCGGACCTTTTAWATGCTGTAWTGKAATTCACATCTCCCCCKACTTGGCARAAT
B. dissectum	1	.AGG.G.AGAAACCAGGAC.AC.CG.A.GACGAA..T.CC.GAC.A.C.C.T..TGGC..GTGC.C.TT..T.C.A...A.TG
B. matricariifolium	26	.GG..A.CTAAACTA.GAC.....A.....T..T.....T.....NN.N
B. multifidum	3	.AGG.GSAGSAAACCAGGAC.AC.CG.A.GASSRR..T.CC.GAC.A.C.C.T..TGGCC.GTGC.C.TT..TG.C.A...A.WK
B. oneidense	1	--------GAAACCAGGAC.AC.CG.A.GA...A..T.CC.GAC.A.C.C.T-.TGGCC----------
B. simplex	21	KGG.RA.ATAAACTG.AAC..C.TGC.....K..TY...-A.K....YGT-AT...TG.K.M.....T....K.KTG..
B. virginianum	6	..A..G.AGAAACCA.GACAACTCG.-A.A....MAT.C.C.ACGAT.A..TC.TGG..YGTG...T.TTYTGT.CA.Y.T..

Table 3. Diagnostic *Dryopteris intermedia matK* SNP.

	n	597883068962378917148343820072568603098084656383601294894719698709241767166784791217022111 48
D. campyloptera	2	GGTTGCCATGTTCCTCGTAGCTCACCCGACTTGCACTCACCCCGAGCGGAATGCAATGACCCGCTCTAGCCCTTCACTCCGAATGGGCTA---
D. carthusiana	4	AA.GCT..C...TT.....ATCT..T...T.C.....T.T.G.T..G.....A..T..AC.....G.....C..A..RATT
D. celsa	3	...GC.T.C.CYTYYYAKR.T.TR.TY..YYCR.R.KYRYTYMR.R.A..W.KYRGCARY.YA..YYRR.TGYYYR.YY.RRRC.AWYCRAC.
D. clintoniana	2	AA.GCT..C...TT.....ATCT..T...T.C.....T.G.T..G.....A..T..AC.....G.....C..A...AT.
D. cristata	7	AA.GCT..C...TT.....ATCT..T..RT.C.....T.G.T..G.....A..T..AC.....G.....Y...C..A...AT.
D. goldieana	5	..CGC..GCACCT...A..AT.T.TT.C.T.CAG.A.T..T.....AT.....CGCA.T.TA..T....TG....TCT.AGGC.A..C.GC.
D. intermedia	6	.....Y.....T.....R....GTC
D. marginalis	1	-------.C.C.TTC.A.G.T.T..T....CC..G...GTTTA....A....T.GGCAG...A...CGA.TGCCTG.....C.AT.C-...

Table 4. Diagnostic *Equisetum fluviatile* and *E. hyemale rbcL* SNPs.

	n	89028251585636945958572592541018403802357397397523316479232840362058725125897503062629940 1
E. arvense	6	GTACCTAAACGCGCAATCTTAACCCTCTAAAGTCTTGTCTTGTCCTACAAGTAGTCCCTATCCCCCTAATACAACTCACTTTAGGTRTWK
E. fluviatile	6	.....C.....G.AT
E. hyemale	21	...TGGG.TSSC.CGCT.GC.TTT.TGRT.ACT...A...A..TG.T...G.....YC...TT..CGT..TC...T..TT...CC...G.AT
E. laevigatum	2	A..TTGGG.T..C.CGCT..C.TTT.TG.T.ACT...AAG.ACTTG.T..A.G.....CG...T.TCGT..TC...T..TT...CCA..G.AT
E. palustre	4	.....T.....T...GA.....M.W.....R.....T..C..T.T.....C.....A.....S.G.AT
E. pratense	4	.....TC.....T...GA.T..T...A...T...C..C.....C...TT.....GT..A.....Y.....G.AT
E. scirpoides	6	..R.....T....CGCT..C.TTTCT.GT.AC....A..WA..T.....CG...Y.C....T.TGT..T...T...C...G.AT
E. sylvaticum	4	.....T...T..T...GA...CT.A.....A.....G...TT.....A...GC.....AGGA-
E. variegatum	10	.W..TGGGGT..CTCGCT..CRTTYYTRGT.ACTC..A...A..TG...M.CGSY...CRA.YT.T.GT..T.M..TT...CC...G.AT

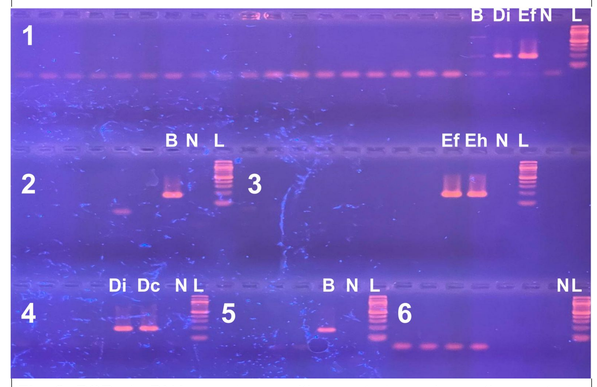

Fig. 3. PCR of eDNA samples, positive controls, and negative controls. For all: B = *Botrychium dissectum*; Dc = *Dryopteris cristata*; Di = *Dryopteris intermedia*; Ef = *Equisetum fluviatile*; Eh = *Equisetum hyemale*; L = 100 bp ladder; N = negative control; unlabeled lanes = eDNA extracts. **(3.1)** *rbcL* D. **(3.2)** *Botrychium trnL–trnF*. **(3.3)** *Equisetum rbcL*. **(3.4)** *Dryopteris matK*. **(3.5)** *Botrychium trnL–trnF*. **(3.6)** *Dryopteris matK*.

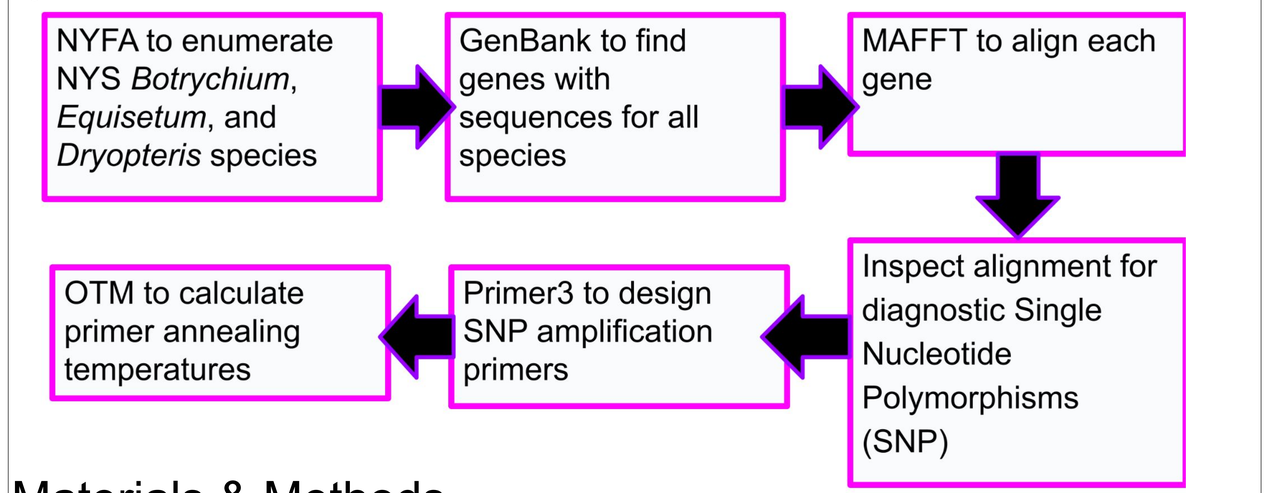

Fig 2. Flow diagram describing diagnostic primer design.

Materials & Methods

The Botanical Garden's old growth forest contains *Botrychium* habitat (mesic) as well as *Dryopteris* and *Equisetum* habitat (riparian). For each habitat, we collected samples from two sites with 18.84 m spacing using a 0.22 µm Sterivex filter (SVGV010RS; Merk) to collect 1.5 L of air over 6 minutes. Negative control filters were handled identically, but without any air.

To recover eDNA from the Sterivex filters, we transferred the membranes into 5 mL tubes, using pliers, under a positive pressure hood (Fig. 1). The membranes were incubated at 42°C for 1 hour with 2 mL of extraction buffer (Alexander et al., 2007). The remainder of the extraction followed Alexander et al. except purified DNA was eluted as 2 × 25 µL.

We identified the species of *Botrychium*, *Equisetum*, and *Dryopteris* present in or near Bronx County using the New York Flora Atlas. We then identified genes with sequences for all of these species available in GenBank: *trnL–trnF* for *Botrychium; rbcL* for *Equisetum*; and *matK* for *Dryopteris*. MAFFT was used to align each gene. We inspected the alignments for diagnostic Single Nucleotide Polymorphisms (SNP). Using the program Primer3 , we designed primers needed to amplify each SNP. The program OTM was used to calculate optimal primer annealing temperatures (Fig. 2).

Each of the samples were amplified using our novel primers (Table 1) as well as *rbcL* D (Little, 2014) primers F52 and R193 (T_m = 52°C) in a 25 µL reaction with 100 ng of purified DNA, 0.5 µM of each primer, 0.025 µg/µL BSA, 0.05% (w/v) orange G, 10% (w/v) sucrose, 0.2 mM dNTPs, 1× Ex Taq PCR buffer, 0.5 u ExTaq (TaKaRa Bio). The reaction mixture was incubated at 95°C for 150 s, cycled 10 times (30 s at 95°C, 30 s at T_m, 30 s at 72°C), cycled 25 times (30 s at 88°C, 30 s at T_m, 30 s at 72°C), and then incubated for 10 minutes at 72°C. PCR reactions were run in a 2% agarose gel in 10 mM sodium borate (pH 8.0) at 10 V/cm and visualized with ethidium bromide.

Results

One SNP within *trnL–trnF* was identified for *B. dissectum* (Table 2): at alignment position 395, *B. dissectum* has a T while other *Botyrchium* species have an A. Unfortunately, *B. multifidum* may also have a T, thus alignment position 359 is also needed because *B. dissectum* and *B. multifidum* are consistently distinguishable (A versus G). No diagnostic SNPs were identified for *D. cristata,* but one SNP in *D. intermedia matK* was located (Table 3): at alignment position 651, *D. intermedia* is the only species that has a T while other *Dryopteris* species have a C. One diagnostic SNP was identified within *rbcL* for *E. fluviatile* and a combination of two SNPs were used to identify *E. hyemale* (Table 4): at alignment position 1219, *E. fluviatile* is distinguished by having a C while other *Equisetum* have a T. Alignment position 1212 distinguishes *E. hyemale* and *E. laevigatum* by having a C while other *Equisetum* have an A. Alignment position 1306 distinguishes *E. laevigatum* from *E. hyemale* (A versus G).

The *rbcL* D primer set successfully amplified the *Dryopteris* and *Equisetum* positive controls. The *Botrychium* positive control produced a faint band of the correct size in addition to a larger product (Fig. 3.1). The *trnL–trnF* primer set successfully amplified the *Botrychium positive control* (Figs. 1.2, 1.5). The *rbcL* primer set successfully amplified the *E. fluviatile* and *E. hyemale* positive controls (Fig 1.3). The *matK* primer set successfully amplified *D. intermedia* and *D. cristata* (Fig. 3.4). PCR amplification of eDNA was unsuccessful.

Discussion

The three novel sets of PCR primers were able to amplify diagnostic regions from positive control samples as designed. In addition, the *rbcL* D primer set worked well for *Dryopteris* and *Equisetum* species, but worked poorly for *Botrychium*. This underperformance was unexpected as the *rbcL* D primers were designed to be universal. The *Botrychium* positive control has two faint bands which may be a PCR artifact (Fig. 3.1). The use of a PCR additive, a change in PCR cycling temperatures, or slight change in primer sequence may be necessary to produce a single band.

The successful amplification of the positive controls indicate that our PCR technique was not the cause of failure (Fig. 3). It is most likely that the DNA extraction procedure and/or eDNA collection were the cause of failure. In order to determine whether the extraction protocol is at fault, additional extractions can be performed on eDNA samples collected with the extracted samples. If the collection procedure is at fault, we can collect additional samples from the forest.

References

Alexander et al. 2007. *Molecular Ecology Notes* 7: 5–9.
Atha et al. 2016. *Brittonia* 68: 245–277.
Banchi et al. 2020. *Science of The Total Environment* 738: 140249.
Banerjee et al. 2022. *AoB PLANTS* 14: plac031.
Katoh et al. 2013. *Molecular Biology and Evolution* 30: 772–780.
Koressaar et al. 2007. *Bioinformatics* 23: 1289–1291.
Little 2014. *Molecular Ecology Resources* 14: 437–446.