Using DNA Barcoding to Identify Species Within a Flower Pit

Author: Alsten, Erik, Neo, Queena & Mentor: Xinyue Wang

DNALC Asia

Introduction

DNA Bar coding is a technology that can sequence certain types of DNA and identify the genus and the species that the DNA is taken from. The types of DNA that DNA bar coding can work with includes COI , which exists in an animal's mitochondrion, or rbcL and matK, which exists in a plant's chloroplast.

This technology can be used to quickly identify species, such advantage can be seen while identifying biodiversity and testing for ingredients in substances.

In our case, we are using the Rapid DNA Isolation method to collect DNA and identify the species of samples that we collected from flower bed in front of the DNA learning center. This isolation method is fast, inexpensive, easy to perform, and suitable for chloroplast DNA collection. After the project is done we can identify around 18 species that grows in the flower bed.

Materials & Methods

The Samples collected include many types of leaves, grass, ferns, and twigs. In front of the DNA learning center there is a flower pit with a stone ring circulating it's perimeter, the samples are picked and taken from there. Materials we used include many test tubes, paper discs, micro pipettes, many pipette tips, pens, gel, PCR solution (which includes primer F, Primer R, loading dye, Taq polymerase, dd H2O, Tris HCl, mM KCl, MgCl2, dNTP, and rbcL primer), and electrophoresis kit.

Rapid DNA Isolation is a method suitable for collecting plant DNA and gives results rather quickly, so we decided to use this method for our extraction. In this extraction we break down hard medium of cells through grinding the sample in lysis solution, then DNA is transferred and washed through dropping a paper disc in to lysis solution, then move the disc to wash buffer, and air dry on the edge of the test tube. At that point the DNA is on the paper disc, so the disc is then transferred to a new tube with TE buffer, and then the buffer is sent for PCR, which the 2ul of the buffer will be added to PCR solution (compositions mentioned above), and the PCR solution containing the DNA will be put in to PCR machine. Within the machine, the DNA samples will undergo a process of 15 second denaturation in temperature of 94C, 15 seconds annealing at temperature of 54C, and 30 seconds extension at temperature of 72C. After all that is done, the sample is loaded in to the gel for electrophoresis, the unused samples are stored at low temperature.

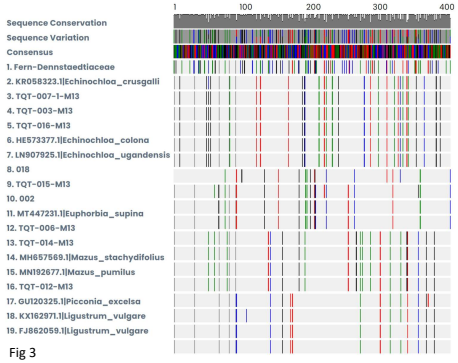

Fig 3

Figure 3 : This MUSCLE allignments displays DNA sequence alignments of samples (labeled as TQT - 001 to TQT - 018) against reference sequences. Each row represents a sample, with color - coded nucleotides indicating base matches (e.g., green for A, red for T) to visualize sequence similarity. The "Consensus" column lists dominant species/sequence groups (e.g., *Echinochloa crusgalli, Euphorbia supina*). Bands at ~600 - 800 bp (consistent with gel electrophoresis results) correlate to amplified rbcL gene fragments, supporting species identification.

Results: Species Identification

For the muscle multiple Sequence Comparison to fern by Log-Expectation, the result shows that our project TQT-007, 003, 016, 014 and GST-001 from other group have the same DNA sequences, which similar to the species Echinochloa colona or Echinochloa crusgalli .Final analyzation proves that GST-001 ,TQT-007 and 016 are the after ,while 003 is the other one species which can be very close to this species .006 as well as 014 belong to the same species. For samples TQT-018 ,015 as well as 002 ,they all own the similar sequence compared to Euphorbia supina. So for the forward analyzation ,TQT-002 and 018 are the same species euphorbia maculata ,while 015 is ephorbia serpens. Sample TQT-012 is a more unique one, which belongs to ligstrum ovalifolium.

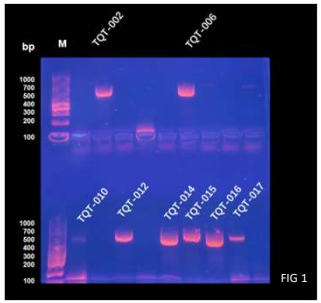

FIG 1

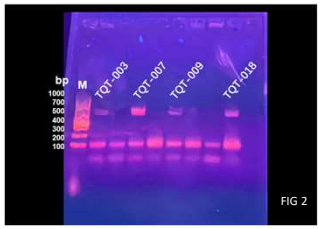

FIG 2

Table 1

Processed samples	Preliminary identification	Traditional taxonomy identification	DNA Barcoding identification
TQT-001	*Purple nut sedge*	_	_
TQT-002	*Roundleaf Toothcup*	*Milk purslane*	*Euphorbia maculata*
TQT-003	*Wood*	*Jungle rice*	*Echinochloa colona*
TQT-004	*Purple nut sedge*	_	_
TQT-005	*Marchantia polymorpha L.*	_	_
TQT-006	*Unknown flowers*	*Japanese mazus*	*us pumilus*
TQT-007	*Unknown leafs*	*Barnyard grass*	*hinochloa crusgalli*
TQT-008	*Unknown leafs*		_
TQT-009	*Purple nut sedge*	_	_
TQT-010	*Purple nut sedge*	_	_
TQT-011	*Unknown trees branches*	_	_
TQT-012	*Chinese Littleleaf Box*	*Common privet*	*Ligstrum ovalifolium*
TQT-013	*Common liverwort*	_	_
TQT-014	*Roundleaf Toothcup*	*Japanese mazus*	*Echinochloa colona*
TQT-015	*Roundleaf Toothcup*	*Matted sandmat*	*Euphorbia Serpens*
TQT-016	*purple nut sedge*	*Barnyard grass*	*Echinochloa crusgalli*
TQT-017	*Roundleaf Toothcup*	_	_
TQT-018	*Cyperus rotundus L*	*Milk purslane*	*Euphorbia maculata*

Total:18 samples / 9 have resluts-5 types of plant

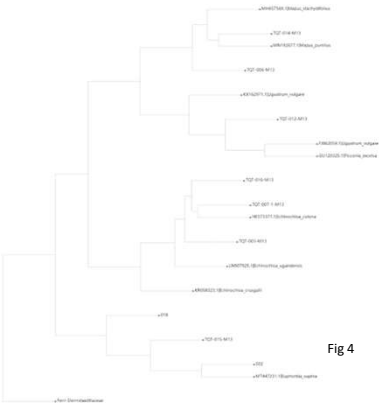

Fig 4

Results: Gel Electrophoresis

Figure 1: The result of samples 1 through 18 after running through gel electrophoresis. Samples TQT-002, 006,012,014, 015, 016, 017 are shown to be successful. Sample TQT-010 shows a dim band.

Figure 2: Agarose gel electrophoresis analysis of PCR-amplified products of the rbcL gene. This gel image shows the PCR-amplified products of the rbcL gene from different samples (labeled as 5, 5, 7, 8, 9, 11, 13, 18) and a sample marked as TQT. The DNA fragments were separated by agarose gel electrophoresis using the DL1000 marker (as shown in Figure 2) for molecular weight calibration. The 600-800 bp of the bands indicate the sizes of the amplified DNA fragments, with clear bands observed in samples 5, 7, 9, and 13, suggesting successful amplification, while the remaining samples showed weak or no bands, indicating low amplification efficiency.

Results: Species Identification

Table 1: Detailed Identification Results
Column A (Processed Samples): Sample IDs (e.g., TQT-001) corresponding to Table 1. Column B (Preliminary Identification): Common name provided during initial submission. Column C (Traditional Taxonomy Identification): Identification results based on morphological analysis. "_" indicates no result.
Column D (DNA Barcoding Identification): Final scientific name confirmed through DNA analysis. "_" indicates no result.
Summary:
Total Samples Processed: 18 Samples with DNA Barcoding Results: 9 (representing 5 distinct plant types). Discrepancies: Some samples show mismatches between preliminary, traditional, and DNA-based identifications (e.g., TQT-002, TQT-016).

Figure 4: This phylogeny tree is built by muscle multiple sequence comparison and demonstrates the relationship between samples (TQT-002/003/006/007/012/014/015/016/018,GST-001) and some common kits (Echinochloa colona/crusgalli/supina/euphorbia maculata/serpens/ligstrum ovalifolium).The fern-Dennstaedtiaceae acts as primer which builds the whole relationship with both samples and common kits, supporting the reference for the similarity of samples.

Discussion

Dry leaves have relatively high concentration of polysaccharide and polyphenol, polysaccharides disturbs the silica resin from binding to the DNA, polyphenol further disturbs the polymerase chain reaction. The DNase can still maintain active during the process of getting dry, after the dna is broken down into smaller fragments, the efficiency of the silica binding to the dna is lower, that might be a cause of the failure of extracting the dna. Phenolics, such as terpenoids and tannins, undergo rapid oxidation upon their release from leaf tissue and irreversibly bind to the phosphate backbone of DNA.Moss contain a lot of pectin which surrounds the dna from binding to the silica. It is also been seen on articles that the importance of grinding the sample fully is emphasized.

Although we attempted to analyze the dominant plant within the stone circle (suspected Paper nut sedge), we faced challenges extracting DNA or amplifying the rbcL barcode. Dry leaves of this species likely contain high polysaccharides and polyphenols (consistent with earlier discussion on dry tissue impacts), inhibiting silica binding and PCR. Future efforts could prioritize fresh samples or modify grinding protocols (e.g., longer lyophilization) to improve lysis, ensuring we capture this key species in biodiversity analysis.

Biodiversity is essential for the processes that support all life on Earth, including humans. Without a wide range of animals, plants and microorganisms, we cannot have the healthy ecosystems that we rely on to provide us with the air we breathe and the food we eat. And people also value nature of itself. This is one of the reason why we wanted to study on this project.

For failed samples (e.g., TQT - 010, TQT - 16), primer mismatches or unsuitable barcode regions likely contributed. The rbcL primers used may not align with sequence variations in these samples, especially if they belong to species with divergent chloroplast genomes (e.g., mosses, unique angiosperms). Future work could test alternative primers (e.g., matK for broader taxonomic coverage) or target multiple barcode regions (e.g., ITS for fungi/angiosperms) to improve amplification success, especially for complex samples like moss (high pectin content).

Abstract

DNA Bar coding is a technology that can be used to quickly identify species, such advantage can be seen while identifying biodiversity and testing for ingredients in substances. The Samples collected include many types of leaves, grass, ferns, and twigs. In front of the DNA learning center there is a flower pit with a stone ring circulating it's perimeter. Rapid DNA Isolation is the method we used, suitable for collecting plant DNA and gives results rather quickly, so we decided to use this method for our extraction. Final analyzation proves that GST-001 ,TQT-007 and 016 are the after ,while 003 is the other one species which can be very close to this species .006 as well as 014 belong to the same species. For samples TQT-018 ,015 as well as 002 ,they all own the similar sequence compared to Euphorbia supina. For the forward analyzation ,TQT-002 and 018 are the same species euphorbia maculata ,while 015 is ephorbia serpens. Sample TQT-012 is a more unique one, which belongs to ligstrum ovalifolium.

References

https://www.mbari.org/wp-content/uploads/2015/11/CTAB-DNA-Extract_GenomeSeq-quality.pdf
Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant speciesAcknowledgements

Acknoledgement

Thanks to Ms. Xinyue Wang, and Dr.Cris and The matierials and machines provided by DNALC Asia