Using rbcL, matK and ITS primers to identify species

Zixiu Liu, Guanghua Cambridge international school
Yiran Lu, Ulink college of Suzhou industrial park
Yupeng Tao, Wuxi No.1 Highschool
Mentor: John Mark Olson. Xinyue Wang

BARCODE SUZHOU
冷泉港亚洲DNA学习中心
Cold Spring Harbor Asia DNA Learning Center

• Abstract

In the experiment, primers rbcL, matK and ITS were used to identify the species of plant. According to the experiment, we can distinguish different clovers. And when we find the species of plants on the introduction sign was different from the species that was determined from plant identification application, we are able to know the real species after these experiments. Finally, we revealed the result that the introduction sign gave us a wrong species of plant.

• Introduction

DNA barcode technology is a molecular identification technology that uses one or several gene fragments to identify species. Nowadays, the research and application of DNA barcodes in plant groups are still in its infancy. The ideal barcode sequences should be accurate and efficient in identifying a lot of closely related species within a single family and genus. In our experiments, we tried to know which primer is most suitable to identify clover samples which belong to same genus but different species. And after this experiment, we can identify different clovers clearly. During the experiments, we used three primers rbcL, matK and ITS2. Because RbcL is universal. It is also easy to amplify. MatK evolves rapidly, but it is difficult to amplify and sequence between different clade groups. ITS is suggested to used as the core DNA barcode in seed plants. When we were collecting plant samples, we detected that some species of plants on the sign which used to introduce the plant were different with the species which was shown on the plant identification application "xingse". As a result, we also collected these samples and used different primers to identify the real species of them.

• Method

1.1 DNA extraction

We use silica DNA Isolation method to extract FAW001 to 008 plants sample. The samples (FAW009 to 012) have already been extracted for us; 009 has been extracted by rapid method, the others are by silica method.

1.2 PCR

We use 2 ul DNA sample to run PCR with the primers respectively.

1.3 electrophoresis send for sequencing

2% Agarose gel + gel red was used to justify if the sample have been amplified or not. The result with clear, bright amplification band will be sent for sequencing.

1.4 analysis

Use DNA subway to trim, pair and consensus edit the sequenced DNA. Use BLASTN to choose species with similar DNA sequence. Use MUSCLE to form DNA barcoding and PHILIP NJ and ML find the percentage of similarity they shared.

Primers	Profile
Plant rbcL primer set: (rbcLaF / rbcLa rev)	Initial step: 94°C 1 minute 35 cycles of the following profile: Denaturing step: 94°C 15 seconds Annealing step: 54°C 15 seconds Extending step: 72°C 30 seconds One final step to preserve the sample: 4°C *od infinitum*
Plant *matK* primer set: (matK-3F / matK-1R)	Initial step: 94° C 3 minutes 41 cycles of the following profile: Denaturing step: 94° C 30 seconds Annealing step: 48° C 40 seconds Extending step: 72° C 1 minute Additional extending step: 72° C 10 minutes One final step to preserve the sample: 10°C *od infinitum*
Plant Plant-ITS primer set: (nrITS2-S2F / nrITS2-S3R)	Initial step: 95° C 2 minutes and 30 seconds 35 cycles of the following profile: Denaturing step: 95° C 30 seconds Annealing step: 56° C 30 seconds Extending step: 72° C 30 seconds Additional extending step: 72° C 10 minutes One final step to preserve the sample: 10°C *od infinitum*

Figure 1 Steps of PCR for different primers

• Materials

FAW001	Canna indica
FAW002	Rosa chinensis
FAW003	Hydrangea macrophylla
FAW004	Cyccas revoluta
FAW005	Cerasus yedoenis
FAW006	Cinnamomum
FAW007	Rosa chinensis
FAW008	Rosa rugosa

- FAW009-013 are different clover samples. All the samples' species are identified by identification app "xingse"

• Result

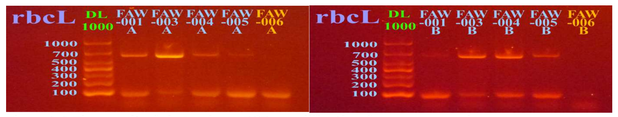

Figure1.1 The result of rbcL primer PCR amplification for some samples

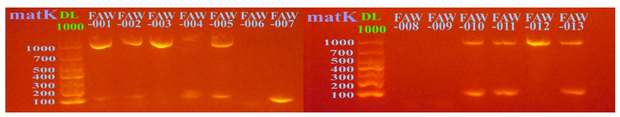

Figure1.2 The result of matK primer PCR amplification for all samples

Figure1.3 The result of ITS primer PCR amplification for all samples

Figure1.4 The result of rerun matK primer PCR amplification

Figure1.5 The result of rerun plant-ITS primer PCR amplification

Table1 The results of getting DNA sequence from different primers

	rbcL	matK	ITS
FAW001	O/+	O/+	O/+
FAW002		O	
FAW003	O/+	O/+	O/+
FAW004	O/+	O/+	O/+
FAW005	O/+	O/+	O/+
FAW006			
FAW007			
FAW008			
FAW009	+		O/+
FAW010	+	O	O/+
FAW011	+	O	O/+
FAW012	+	O/+	O/+
FAW013	+	O	O/+
	100%	55.6%	88.9%

*Run electrophoresis and sent:O

*The sequence have been obtained: +

*The table shaded in yellow have already got the sequence.

• Discussion

Through the DNA sequencing, we managed to get DNA sequences for the two target plant which was thought to be mislabeled and analyzed them in DNA subway, through the blast hits, we managed to determine the genus of the plant samples. BLAST hits of the rbcL and plant-ITS (the online sequence viewer have no matK function, so for those sequences only MUSCLE is run) suggests that sample FAW-001 highly likely belongs to *Canna* genus, and there's high possibility that sample FAW-004 is a species under *Cycas* genus. Since the scientific names of the two samples labeled on the signs are respectively *Hosta Plantaginea* and *Brainea insignis*， the assumption that these to plant samples are mislabeled can be verified. For further verification, DNA sequences for the species labeled on the sign and shown by identification apps is obtained from GenBank and run the MUSCLE process with the blast hits and original sample.

1. KC704826|Hosta_plantaginea
2. FAW001-M13
3. JX887615|Canna_indica
4. JX903257.1|canna_indica

Download this alignment
Launch in Jalview

	C	1	2	3	4
C	-	96.02	98.39	100.00	99.82
1	96.02	-	93.62	95.65	95.24
2	98.39	93.62	-	99.28	99.82
3	100.00	95.65	99.28	-	100.00
4	99.82	95.24	99.82	100.00	-

Table 2.1 MUSCLE result for sample FAW001

1. AB040576|Brainea_insignis
2. AY056556|Cycas_revoluta
3. FAW004B-M13
4. MH069537.1|cycas_tropophylla
5. MH069538.1|cycas_brachycantha

Download this alignment
Launch in Jalview

	C	1	2	3	4	5
C	-	89.39	95.44	98.71	99.82	100.00
1	89.39	-	84.43	82.90	83.42	83.60
2	95.44	84.43	-	97.97	99.64	99.82
3	98.71	82.90	97.97	-	99.64	99.82
4	99.82	83.42	99.64	99.64	-	99.82
5	100.00	83.60	99.82	99.82	99.82	-

Table 2.2 MUSCLE result for sample FAW004

General MUSCLE result for sample FAW-001 suggests that the sample have only 93.26% gene similarity with the species labeled by the signs ,while the gene consistency between the sample, results suggest by identification app and blast hits are all higher than 99%.

For sample FAW-004, the sample have only around 84 percent of gene similarity with the specie suggested on the labeling, while the specie given by the identification app have nearly 98 percent of similarity. This number is a little bit low, probably because of the variation of species within the ***Cycas*** genus.

According to the data above, it can be verified that the signs in the park have mislabeled the respective plants.

Compare clovers with three primers

The ML trees of three primers are shown, the primer rbcL can be sequenced for all the samples and can basically identify all the samples. The successful sequenced sample of matK primer is rather fewer; only two samples have been successfully sequenced. But we can see larger differences between these samples on the table. As the number of sequences are limited, we cannot do further judgment. For primer ITS, we can barely find some reference DNA from NBCI, but most of the DNA samples have been successfully sequenced(FAW009 are not been sequenced for both matK and ITS, we guess that it may due to that it was extracted by rapid method.). The ITS primer does not show the differences between samples quite clearly and significantly.

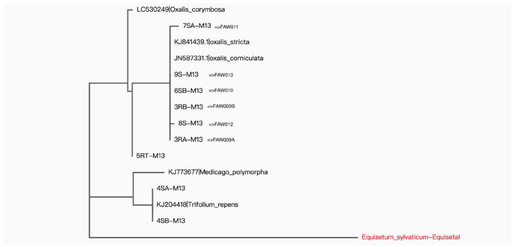

Figure 2.1 ML diagram of rbcL

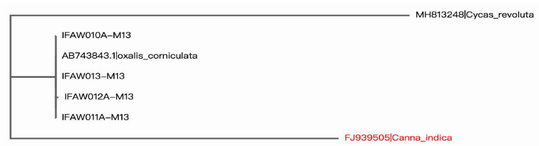

Figure 2.2 ML diagram of ITS

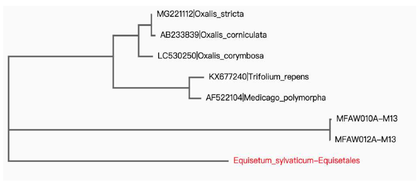

Figure 2.3 ML diagram of matK

Sequence Conservation	C	1	2	3	4	5	6	7
Sequence Variation	C	-	69.18	66.96	86.41	82.23	92.22	91.72	93.95
Consensus	1	69.18	-	99.63	45.90	54.49	57.90	56.25	56.52
1. MFAW010A-M13	2	66.96	99.63	-	52.80	53.20	56.43	54.44	56.33
2. MFAW012A-M13	3	86.41	45.90	52.80	-	94.21	77.16	77.26	77.86
3. AF522104|Medicago_polymorpha	4	82.23	54.49	53.20	94.21	-	75.49	75.90	76.30
4. KX677240|Trifolium_repens	5	92.22	57.90	56.43	77.16	75.49	-	93.06	94.03
5. LC530250|Oxalis_corymbosa	6	91.72	56.25	54.44	77.26	75.90	93.06	-	98.96
6. AB233839|Oxalis_corniculata	7	93.95	56.52	56.33	77.86	76.30	94.03	98.96	-
7. MG221112|Oxalis_stricta									

Table 3.1 matK

Sequence Conservation	C	1	2	3	4	5	6
Sequence Variation	C	-	95.65	99.04	99.81	100.00	100.00	100.00
Consensus	1	95.65	-	94.62	95.65	95.65	95.65	95.56
1. MN601879|Oxalis_stricta	2	99.04	94.62	-	99.02	99.02	99.02	99.49
2. IFAW012A-M13	3	99.81	95.65	99.02	-	99.81	100.00	100.00
3. IFAW011A-M13	4	100.00	95.65	99.02	99.81	-	100.00	100.00
4. IFAW013-M13	5	100.00	95.65	99.02	100.00	100.00	-	100.00
5. IFAW010A-M13	6	100.00	95.56	99.49	100.00	100.00	100.00	-
6. AB743843.1|oxalis_corniculata								

Table 3.2 ITS

Sequence Conservation	C	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Sequence Variation	C	-	95.98	92.44	92.36	91.99	99.37	99.84	99.53	100.00	100.00	100.00	100.00	100.00	97.01	96.21
Consensus	1	95.98	-	95.98	96.83	97.25	89.60	90.33	90.28	90.33	90.33	90.33	91.25	91.51	91.79	92.76
1. KJ773677|Medicago_polymorpha	2	92.44	95.98	-	99.84	100.00	91.81	92.28	92.10	92.44	92.44	92.44	91.62	92.70	92.44	92.06
2. 4SA-M13	3	92.36	96.83	99.84	-	100.00	92.20	92.20	92.36	92.36	92.36	92.36	91.62	92.70	92.36	92.06
3. 4SB-M13	4	91.99	97.25	100.00	100.00	-	91.80	91.99	91.99	91.99	91.99	91.99	91.71	92.41	92.35	92.35
4. KJ204418|Trifolium_repens	5	99.37	89.60	91.81	92.20	91.80	-	99.21	99.05	99.37	99.37	99.37	99.81	99.81	96.38	96.03
5. 7SA-M13	6	99.84	90.33	92.28	92.20	91.99	99.21	-	99.37	99.84	99.84	99.84	100.00	100.00	96.86	96.21
6. 3RA-M13	7	99.53	90.28	92.10	92.36	91.99	99.05	99.37	-	99.53	99.53	99.53	100.00	100.00	96.52	96.21
7. 8S-M13	8	100.00	90.33	92.44	92.36	91.99	99.37	99.84	99.53	-	100.00	100.00	100.00	100.00	97.01	96.21
8. 3RB-M13	9	100.00	90.33	92.44	92.36	91.99	99.37	99.84	99.53	100.00	-	100.00	100.00	100.00	97.01	96.21
9. 6SB-M13	10	100.00	90.33	92.44	92.36	91.99	99.37	99.84	99.53	100.00	100.00	-	100.00	100.00	97.01	96.21
10. 9S-M13	11	100.00	91.25	91.62	91.62	91.71	99.81	100.00	100.00	100.00	100.00	100.00	-	100.00	97.02	96.37
11. KJ841439.1|oxalis_stricta	12	100.00	91.51	92.70	92.70	92.41	99.81	100.00	100.00	100.00	100.00	100.00	100.00	-	96.82	95.92
12. JN587331.1|oxalis_corniculata	13	97.01	91.79	92.44	92.36	92.35	96.38	96.86	96.52	97.01	97.01	97.01	97.02	96.82	-	99.28
13. 5RT-M13	14	96.21	92.76	92.06	92.06	92.35	96.03	96.21	96.21	96.21	96.21	96.21	96.37	95.92	99.28	-
14. LC530249|Oxalis_corymbosa																

Table 3.3 rbcL

• Reference

张娜,乾义柯,贾凤勤,焦子伟,张祥林. 野生樱桃李的DNA条形码基因筛选与评价[J]. 植物检疫,2018,32(03):34-42.
<< Using DNA Barcodes to Identify and Classify living Things>>