Proposed DNA Barcoding Protocol: Starter & Familiarization for Phytochemists

CENAPT

https://cenapt.pharm.uic.edu/

This protocol is meant to help initiating DNA barcoding for identification of botanical samples in a phytochemical laboratory.

Charlotte Simmer,
Center for Natural Product Technologies, UIC

Table of Contents

CONSUMABLES **2**
EQUIPMENTS AND STOCK SOLUTIONS **2**
PRIMERS **3**
STEP 1: DNA EXTRACTION AND PURIFICATION **4**
STEP 2: POLYMERASE CHAIN REACTION **6**
A. ADAPTED PROTOCOL MIXTURE / REACTION TUBE 6
B. PCR REACTION 6
C. EVALUATION OF THE AMPLIFICATION SUCCESS AND CLEANLINESS: GEL ELECTROPHORESIS 7
STEP 3: PCR CLEANUP BEFORE SEQUENCING **8**
STEP 4: PREPARE AND SEND SAMPLE FOR SEQUENCING **9**
STEP 5: ANALYZE YOUR SEQUENCING RESULTS **9**
EXAMPLE: ANALYSIS OF *TRIGONELLA FEONUM GRAECUM* SEEDS ITS2 SEQUENCE **10**

Consumables

GoTaq® Green Master Mix is a premixed, ready-to-use solution containing bacterially derived *Taq* DNA polymerase, dNTPs, $MgCl_2$ and reaction buffers at optimal concentrations for efficient amplification of DNA templates by PCR

	Provider	Part #	qty	cost	
DNeasy Plant Mini Kit (50)	Quiagen	69104	1	224	50 isolations
ExoSAP-IT™ PCR Product Cleanup Reagent	ThermoFisher	78200.200.UL	1	108	100 reactions
Promega™ Go Taq™ Master Mixes	Fischer scientific	Promega™ M7123	1	437.5	1000 reactions
Triple Pure Zirconium beads	Sigma Aldrich	Z763802-50EA	1	143	50 vials

⇒ Upon receipt of the Go Taq Master Mixes, prepare several aliquots of 500 µL each and keep these aliquots in the freezer at -20 °C.

⇒ Keep your Exosap IT at -20 °C as well.

Equipments and Stock Solutions

- Mini bead beater (Beadbug)
- Nanodrop 1000 spetrophotometer Thermoscientific
- Centrifuge 5417R Eppendorf
- Electrophoresis kit BioRad power pac 200
- Agarose: SEAKEM LE Agarose (Lonza), powder
- Ethidium Bromide (Promega cat# H5041 10mg/mL)
- TEA solution 1X (stock solution 50X ThermoScientific # B49)
- DNA 1kB GeneRuler SM0313 ready to use (keep at -20 °C for long storage)
- Microwave for the preparation of agarose gel
- MasterCycler gradient Eppendorf (PCR reaction)
- Eppendorf and PCR tubes
- Pipettes P100-P200-P10 µL

Primers

Order primers at IDT: Integrated DNA Technologies

for UIC use an ilab account through the DNA Resource Center (at UIC): (http://idtdna.com/uic/)

Universal primers selected for Plant DNA barcoding

CODE	sequences (5'-3')	DNA region	DNA length	% GC & tm
MTR-F	GCTCCAACAAATGGATAAGAC	psbA-trnH_Marrubium species	Expected DNA length : 275 pb	42.9 (tm 57.4)
MTR-R	ACTGCCTTGATCCACTTGG	psbA-trnH		52.6 (tm 57.3)
S2F	ATGCGATACTTGGTGTGAAT	ITS_universal	variable according to species	40 (tm 54.3)
S3R	GACGCTTCTCCAGACTACAAT	ITS-universal	< 500 pb	47.6 (tm 59.4)
PF	GTTATGCATGAACGTAATGCTC	psbA-trnH_universal	variable according to species	40.9 (tm 58.4)
TR	CGCGCATGGTGGATTCACAATCC	psbA-trnH	< 400 pb	56.5 (tm 66.4)
390F	CGATCTATTCATTCAATATTTC	MatK_universal	variable according to species	26.3 (tm 44.9)
1326R	TCTAGCACACGAAAGTCGAAGT		Expected ~700 pb	45.5 (tm 56)
1f	ATGTCACCACAAACAGAAAC	RBCL_universal	variable according to species	40 (tm 51.1)
724r	TCGCATGTACCTGCAGTAGC		Expected ~800 pb	55 (tm 57.3)

Concentration: 25nM without laboratory formulation

- Upon primer reception, re-dilute in 250 µL distilled water or buffer solution from DNA extraction
- Final primer concentration: 100 µM
    - utilize 0.1 µL/PCR reaction

Other option:

- Re-dilute the 100 µM stock solution to 10 µM final concentration
    - utilize 1 µL/PCR reaction

References articles for the chosen primers:

1. *Chen S. et al. "Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species" PlosOne 2010, 5: e8613*
2. *Pawar RS et al. "Assessment of the Authenticity of Herbal Dietary Supplements: Comparison of Chemical and DNA Barcoding Methods." Planta Medica, 2017, 83 :921-936*

Note: all the primers used except the MTR-F/R (designed) were found in the supporting information of the publication by Chen S. et al. [1]

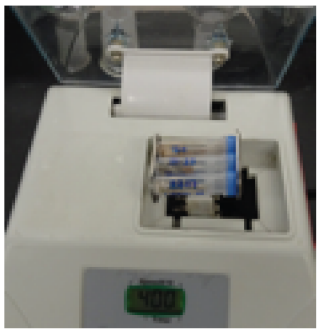
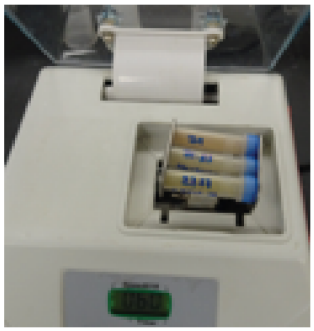
In some cases, the DNA sequences obtained with such "universal primers" can only confirm the botanical genus. Hence, for further identification down to the level of a species, specific primers (specific to the target species) would have to be used or designed. The users can refer to websites cited at the end of the document and data from the literature.

Step 1: DNA extraction and purification

According to the Plant DNAeasy minikit instructions with some modifications.

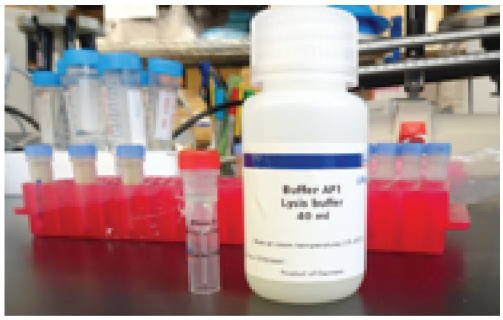
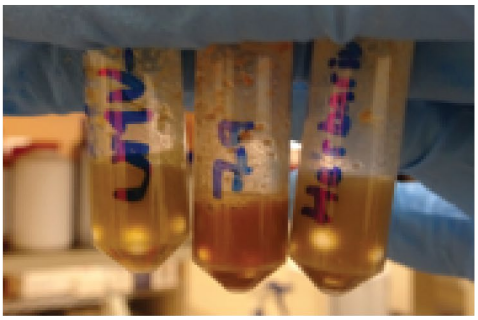
Total duration of extraction for one sample:

1. Disrupt the plant tissue using the mini bead beater (Beadbug) and the vials containing the Zirconium beads
    a. Qty of plant powder: < 100 mg (25-75 mg)
    b. Note exactly the weight of plant powder as this will define the volume of AP1 solution
    c. Use the beadbug microtube homogenizer at speed 400 and time 60 sec

mini BeadBeater Before & After vortex (60 sec, speed 400)

2. Add **400 µL AP1 solution** (for max 25 mg dry powder) and **4 µL RNAase**

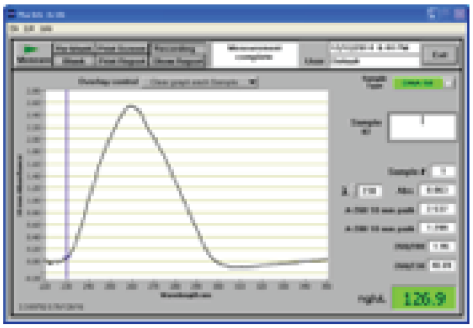
Adding the AP1 solution to lyse the cells in the BeadBug microtubes

    a. mix well (vortex) no clamps should be seen
    b. this solution AP1 is utilized to lyse the cells, and contains acids
    c. the RNAse is added to destroy any remaining RNA and avoid a wrong estimation of your total extracted DNA
3. Incubate the tube at 65 C during 10 min ideally under slow agitation, if not manually invert the tubes 3 times during the 10 min incubation.
4. Add 130 µL min of P3 buffer (for 400 µL AP1) and put your solution on ice for 5 min.

    This P3 buffer neutralize the solution is should help precipitating the proteins, detergents and possible polysaccharides present in your plants.

5. Centrifuge the vials during 5 min at 14 000 RPM
6. Collect the supernatant on the Quiashredder Column (**violet**) and centrifuge 2 min, 14 000 RPM
7. Collect the flow-through solution without disturbing any potential cell debris at the bottom of the vial (not always present) and place the liquid in a new 2 mL tube
8. Add 1.5 volume of AW1 buffer (1.5 X volume of collected flow through solution) to the 2 mL vial and mix by pipetting several time, which should potential lead to the formation of precipitate
9. Take 650 µL of this mixture and place it on a DNAeasy mini-spin column, then centrifuge at 9500 RPM for 1 min. At this step you clean/ trap the extracted DNA on the column.
10. Repeat step 9 with the remaining liquid from step 8
11. Place the mini spin column in a new 2 mL tube and add 500 µL of AW2 buffer, centrifuge 9500 RPM 1 min, repeat twice. This step cleans the column/DNA from other plant constituents.
12. Dry the column with a final step of 500 µL AW2 buffer and a centrifugation at 14,000 RPM during 2 min

    Make sure the column is dry and does not touch the flow through Ethanol solution as this will affect the measurement of the DNA concentration
13. Transfer the column in to a clean 1.5 mL centrifuge tube and elute the DNA with the Buffer AE 100uL, 5 min incubation at RT, centrifugation at 9500 RPM, 1 min
14. Collect the flow through solution (containing your purified DNA) into a new Eppendorf tube and label it properly for storage and future use, This solution contained your DNA template on which your future PCR will be performed
15. Your DNA solution should be colorless/translucid.
16. Measure the DNA concentration using the Nanodrop 1000 instrument: target ratio (A260/280)= 1.7-2, **with ideally a symmetric absorbance around 260 nm**.



Ideal DNA purification result

Remark: If the absorbance does not look like that, it does not mean you have no DNA or that you can't run any PCR. It is still worth trying to run a few PCR and see what you can get.

⇒ **Preparing your blank/ PCR negative control**: Ideally, you want additionally to produce here your negative control for the PCR reaction by performing all the steps of the extraction but without adding your actual sample. This blank should not contain any DNA and thus will not give any PCR amplicons.

Other kits to try for DNA extraction that offer removal of PCR inhibitors: DNeasy PowerMax Soil Kit and QIAAmp DNA stool Mini Kit.

Step 2: Polymerase Chain Reaction

According to the Go Taq Green protocol

For a 25µl reaction volume:

Component	Volume	Final Conc.
GoTaq® Green Master Mix, 2X	12.5µl	1X
upstream primer, 10µM	0.25–2.5µl	0.1–1.0µM
downstream primer, 10µM	0.25–2.5µl	0.1–1.0µM
DNA template	1–5µl	<250ng
Nuclease-Free Water to	25µl	N.A.

A. Adapted protocol mixture / reaction tube

1. 12.5 µL Master Mix
2. 1 µL primers (10 uM) /reaction (tube) = 2uL
3. 12.5 µL Nuclease free water
4. 1 µL of (un)diluted DNA template
5. 0.5 to 1 µL of DMSO (for amplification of GC rich DNA template only)

Keep a **PCR negative control** for each set of primers with no added DNA (add nuclease free water if you did not prepare your blank extraction/control as said above). Ideally, add a **positive control** (DNA template of a reference botanical which DNA will be amplified with your primer mix) to your reaction.

⇒ Dilute your DNA template solution (~concentration 50-20 ng/uL)

⇒ No DMSO in the PCR mix for MatK amplification (the amplified sequence is not rich in GC)

Example of sample preparation for PCR reaction:

With **3 botanical DNA templates**, including one DNA collected from an herbarium specimen (=positive control) and **2 sets of primers**.

- Total number of reactions = 3 x 2 + 2 negative controls = **8 reactions**

Prepare **your total reaction solution for 10 reactions** (consider dead volume and easy sample preparation), divide the mixture in 2 sets of Eppendorf tubes to add your primers. Once your primers are added, distribute your solution in each of your PCR tubes for final addition of DNA template.

- Prepare 10 x 12.5 µl mastermix + 10 x 12.5 µl nuclease free water + 10 x 1µL DMSO (= 260 µL total)
- Distribute 130 µL in 2 Eppendorf tubes and add 5 µL of each primers (= 140 µL total)
- Distribute 28 uL in each PCR reaction tubes
- Add 1 µL of DNA template in the dedicated tube or 1µL of water in the negative control.

B. PCR reaction

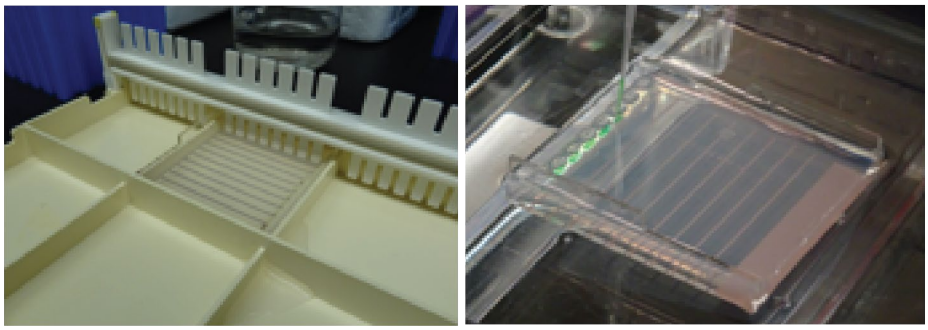
- 94 °C for 5 min
- 40 cycles of
  - 94 °C for 1min (max can be run faster)
  - **55 °C for 30-40 sec****
  - 72°C for 1 min ----- The rule is 30 sec elongation time (at 72°C) for 500 pb.
- 72 °C for 7 min (extension)
- 4°C hold

**A touch-down method (decrease of 0.5 °C every cycle from 65 to 45°C during the annihilation time) was set up for the PCR reaction involving the MatK primers as their Tm show > 10 °C difference.

C. Evaluation of the amplification success AND cleanliness: Gel electrophoresis

a. Preparation of 1% agarose gel with Ethidium Bromide:

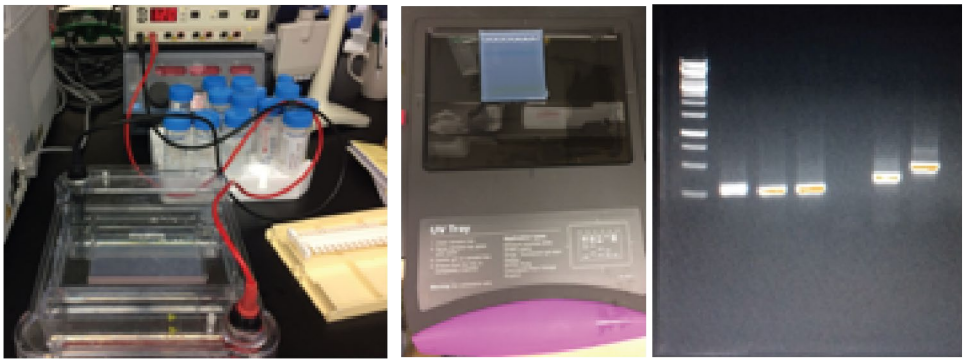
- Dissolve 1g of agarose (SEAKEM LE Agarose) in 100 mL of 1X Tris-Acetate EDTA (TAE) buffer
- Microwave for ~ 1 min or until a total dissolution of your agarose is reached
- Add 5µL of Ethidium Bromide (Promega cat# H5041 10mg/mL)
- Cast your still liquid and hot gel in a rack with a comb defining the number of samples to run
- Be careful and pour your gel slowly so as to avoid making and trapping tiny bubbles in your gel as this will affect the quality of your results.
- Let it dry for ~ 30 min



Preparing and loading your electrophoresis gel

b. Run the electrophoresis

- Put the cast in the electrophoresis chamber and cover it with TAE 1X
- Load 5 µL of the PCR products directly in each well
- Load the DNA molecular weight ladder 1kb (extreme left side, see reference in the first page)
- Run the gel at 120V for ~20 min (run to red = positive as DNA is negatively charged)
- Observe your gel under UV
- Determine the approximate size of your PCR product



"Run" towards red UV tray for visualizing your results

NOTE: If you did not have any PCR product, it does not mean there is no DNA template or the specific species you are looking for is not present in your sample. Either your primers are not adapted to amplify the DNA of your sample, or some PCR inhibitors (enzyme inhibitors) are present in the solution of your DNA template preventing any amplification. In that case, a possible way to obtain PCR amplicon is to dilute your DNA template solution (e.g. 50 times, thereby diluting as well your PCR inhibitors such as tannins or other polyphenols) and use this diluted solution to re-run your PCR.

Step 3: PCR cleanup before sequencing

Information on Exosap-IT: destroys primers before sequencing. The enzyme should be kept on ice when preparing the reaction mixture for PCR cleaning.

A. Preparation of reaction tube:

a. 1 µL Exosap-IT
b. 5 µL sample
c. 4 µL nuclease free water

Remark: it is also possible to use 1 µL of Exosap-IT directly in 5 µL of sample (i.e., no dilution) if you want to reduce possible background noise during Sanger sequencing.

B. Reaction:

a. 15 min at 37 °C
b. Inactivation: 80°C for 15 min

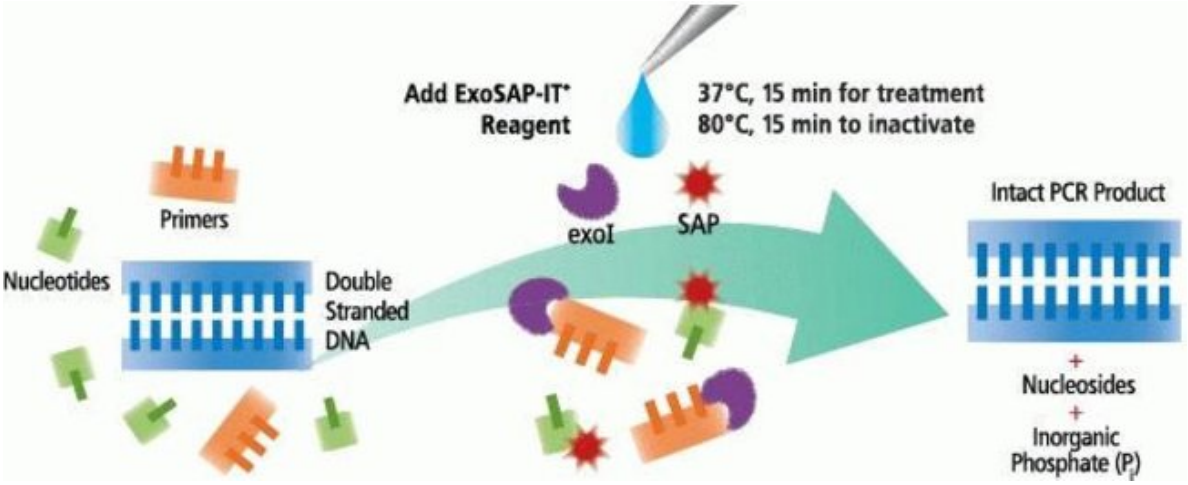


Illustration from Exosap-IT provider

Step 4: Prepare and send sample for sequencing

For UIC User only: request your sequencing at https://my.ilabsolutions.com/account/login#

Ask for sequencing in both direction to improve the accuracy of identification and quality of sequencing results. Therefore, 2 reactions are performed per DNA amplicon= 2 Excel lines per sample as shown below.

Reaction Number	SAMPLE NAME (Plasmid - 100 ng/uL; PCR - 10 ng/uL)	INDICATE IF IT IS Plasmid or PCR	TOTAL SIZE (KB)	PRIMER NAME (5 uM)
1	UES_U1	PCR	250	S2F
2	UES_U1	PCR	250	S3R
3	UMV-U1	PCR	250	S2F
4	UMV-U1	PCR	250	S3R
5	UPJ_U1	PCR	250	S2F
6	UPJ_U1	PCR	250	S3R
7	UTF_U1	PCR	250	S2F
8	UTF_U1	PCR	250	S3R
9	USA_M1	PCR	250	S2F
10	USA_M1	PCR	250	S3R
11	USA_U1	PCR	250	S2F
12	USA_U1	PCR	250	S3R
13	UTF_U2	PCR	250	S2F
14	UTF_U2	PCR	250	S3R

Additionally, you should prepare max 5 µL primer / reaction. So here you will need max ~ 30 µL of each primer for 5 to 10 µL of sample.

For UIC: The samples should be prepared in 1.5 mL centrifuge tubes and labelled with black, freezer safe ink

Step 5: Analyze your sequencing results

- Free Software utilized: **Serial Cloner 2-6-1** and/or **ApE** (observe your chromatogram and /or edit your sequence)
- Websites:
- **GenBank BLASTn** (https://blast.ncbi.nlm.nih.gov/Blast.cgi), where you can paste your sequence and check the proposed identity and percentage of homology between sequences.
- **T COFFEE** (http://tcoffee.crg.cat) for DNA sequence alignment and comparison using text file.

Simple analysis:

- Take your sequence, open your chromatogram and determine the cleanest region to select
- Copy and paste it into BLASTn, click on Blast
- Observe the results (% identity of the query sequence with those available in GenBank) check the source of the GenBank sequence.

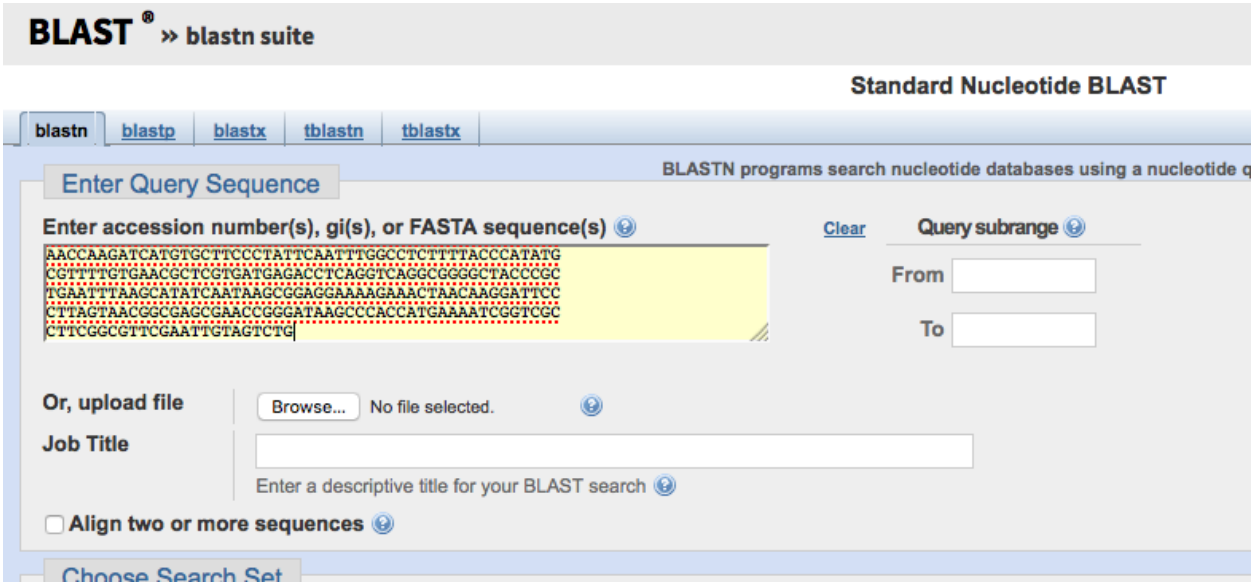
Example: Analysis of *Trigonella feonum graecum* seeds ITS2 sequence

ITS2 amplicon: 424 pb

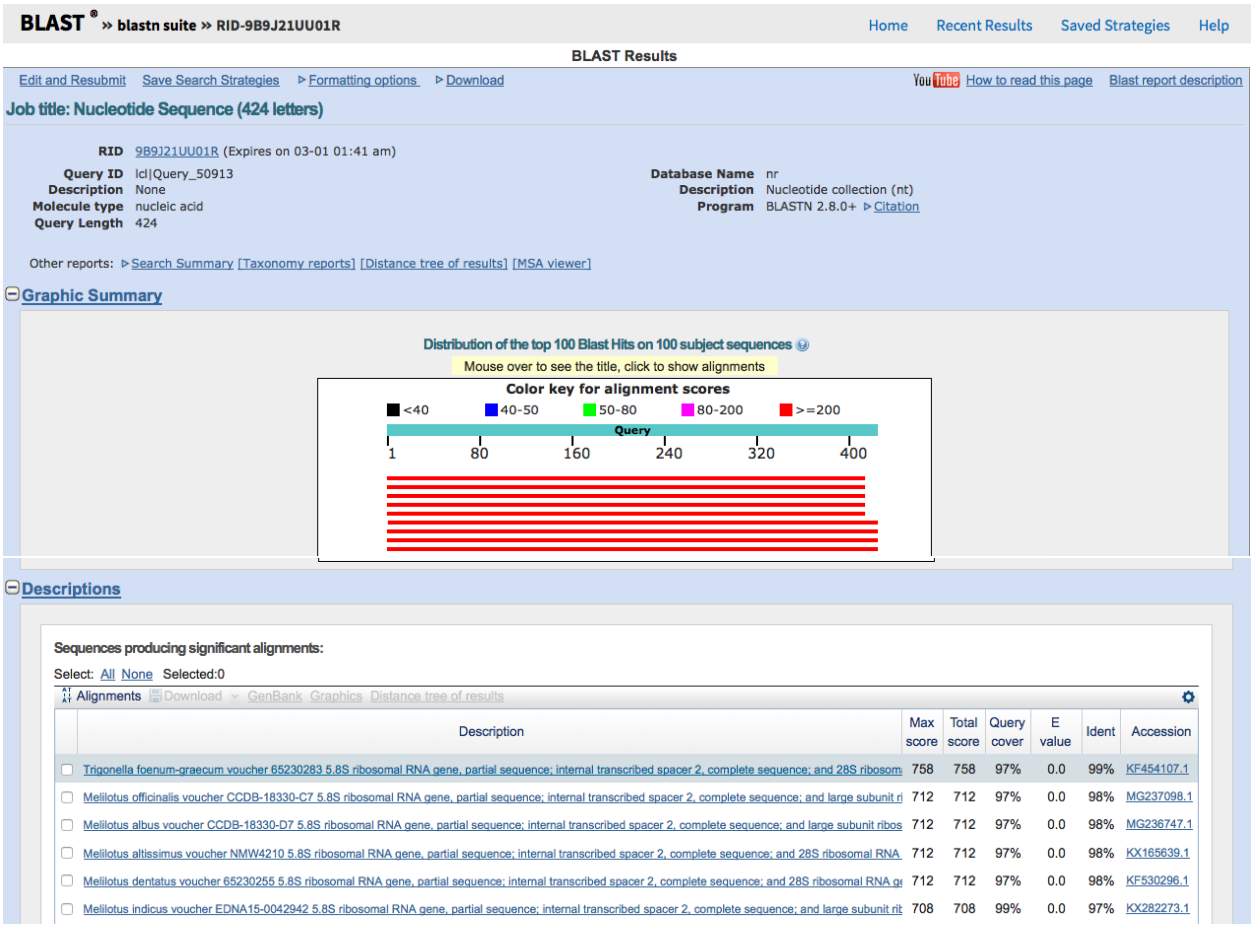
>R40_13.UTF_U2-S2F = name of the sample

TTGCGCCCGATGCCATTAGGTTGAGGGCACGTCTGCCTGGGTGTCACATATCGAAGCCTCATGCCAATTTCCTTTTT
TAGTAGGTATTGTGCATGCTGGTGAATGTTGGCCTCCCGTGAGCTCTATTGTCTCATGGTTGGTTGAAAATCGAGA
CCTTGGTAGGGTGTGCCATGATAGACGGTGGTTGTGTGACCCACGAGAACCAAGATCATGTGCTTCCCTATTCAAT
TTGGCCTCTTTTACCCATATGCGTTTTGTGAACGCTCGTGATGAGACCTCAGGTCAGGCGGGGCTACCCGCTGAAT
TTAAGCATATCAATAAGCGGAGGAAAAGAAACTAACAAGGATTCCCTTAGTAACGGCGAGCGAACCGGGATAAG
CCCACCATGAAAATCGGTCGCCTTCGGCGTTCGAATTGTAGTCTG

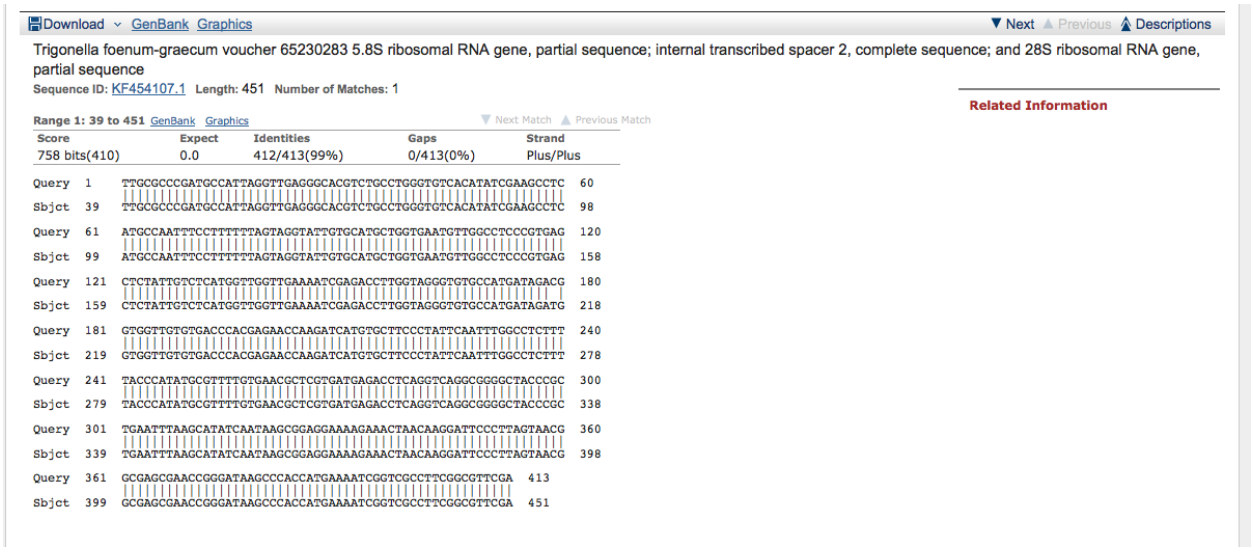
Copy and paste into BLASTn:



Blast and check the results:



Fortunately in this case, the species has been well identified:



You can also check your chromatogram with ApE:

Very IMPORTANT: Understand the limitations of DNA barcoding and possible alternatives according to your results! Put your results into proper perspective.

Some useful references and websites

1. *Raclariu AC et al.* "Benefits and Limitations of DNA Barcoding and Metabarcoding in Herbal Product Authentication." Phytochem. Anal. **2018** Mar; 29(2): 123-128
2. https://www.dnabarcoding101.org/resources/
3. www.boldsystems.org
4. http://www.ibol.org/phase1/about-us/what-is-dna-barcoding/
5. http://botany.si.edu/projects/dnabarcode/proj_db.htm
6. https://www.kew.org/blogs/kew-science/what-plant-is-this-scan-the-genomic-barcode