DNA BARCODING OF PLANTS AT SHAW NATURE RESERVE USING *matK* AND *rbcL* GENES

LIVINGSTONE NGANGA.

Missouri Botanical Garden.

What is DNA barcoding?

- Barcoding is the use of short DNA sequences to identify and differentiate species.
- The sequences are intraspecifically conserved but vary interspecifically.

USES OF DNA BARCODES

- Easy identification of species.
- Can be used to detect illegal trade of restricted animal parts or plants.
- Testing the authenticity of plant products.
- Complement taxonomy in difficult groups.

FEATURES OF A GOOD DNA BARCODE

- DNA barcodes should,

a) Be short for easy PCR amplification.

b) Be easily sequenced in both the forward and reverse direction.

c) Have universal primers.

d) Be conserved within species but variable among different species.

CURRENT STATUS

- Plastid genes preferred
- In animals, the mitochondrial gene *CO1* is used
- Several genes proposed for plants with the most prominent being *atpF-atpH, matK, nhdJ, psbK-psbI, rbcL* and *trnH-psbK* (all chloroplast genes) and *ITS* (a nuclear gene)
- Plastids are clonal and can give info such as hybridization events
- A combination of *rbcL* and *matK* was proposed as the plant barcode by CBOL

rbcL and *matK* comparison

- ***rbcL***

1. Entire gene ~1400bp long
2. ~ 650 bp used for barcode
3. Can discriminate genera but not species
4. Easily amplified
5. Over 20,000 species sequences publicly available

- ***matK***

1. Entire gene ~1550bp long
2. ~ 880 bp used for barcode
3. High species discrimination
4. Low amplification in some groups
5. ~22,000 species sequences publicly available

OBJECTIVE

- Test the effectiveness of the *rbcL*+*matK* as a barcode on plants at Shaw Nature Reserve.

A glade in Shaw Nature Reserve. Photo credit: Livingstone Nganga.

Shaw Nature Reserve

- About 30 miles from St. Louis
- Vegetation types include upland and bottomland forest, glades, tall grass praire, wetland, old graze lands and cultivated fields
- Includes 1053 plant species in 503 genera and 151 families

Diversity

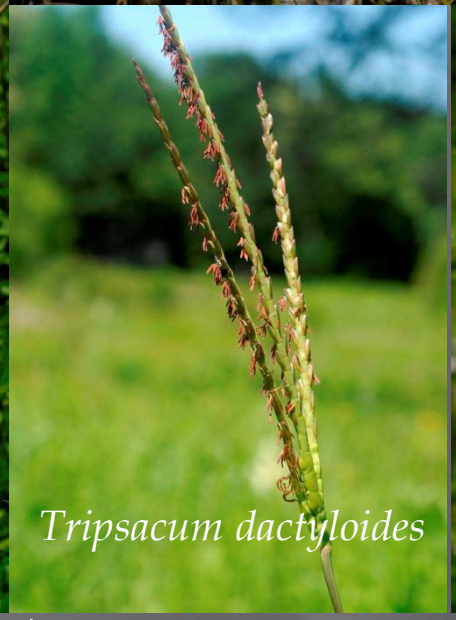

Echinacea paradoxa


Oenothera macrocarpa


Hydrangea arborescens.


Matelea decipiens


Melanthium virginicum


Tripsacum dactyloides


Rosa setigera

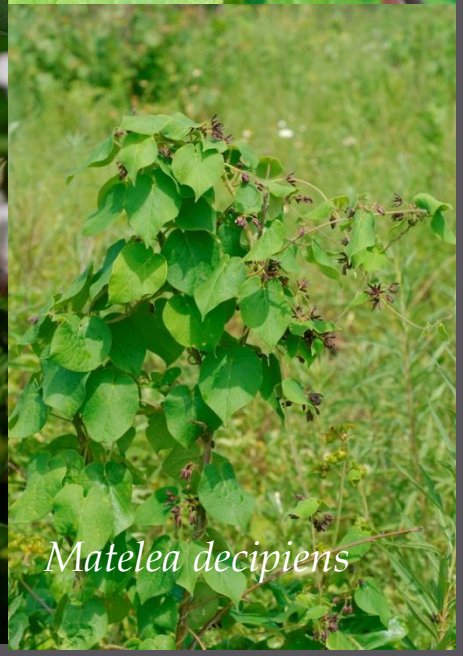
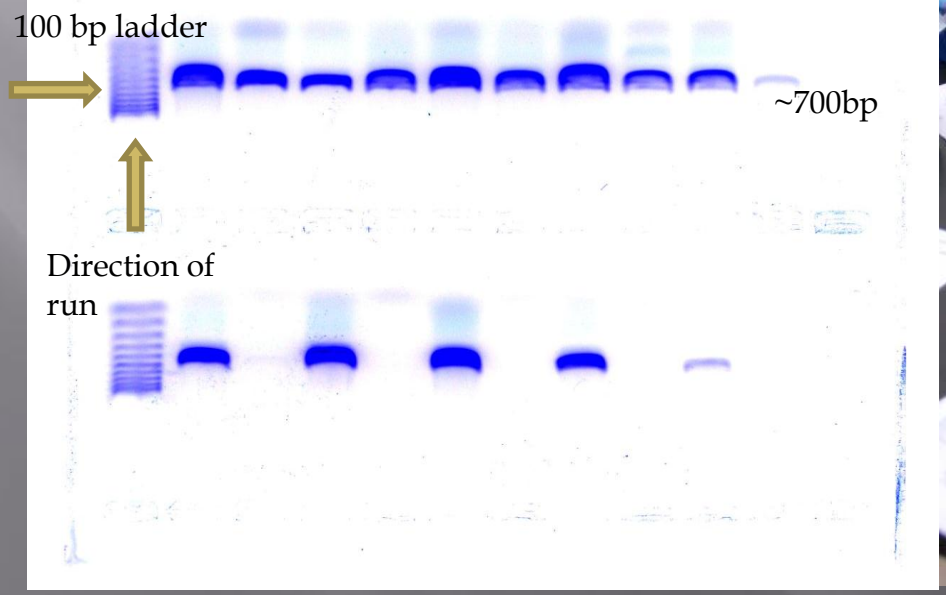

Matelea decipiens

Photo credit. David Bogler.

Materials and methods

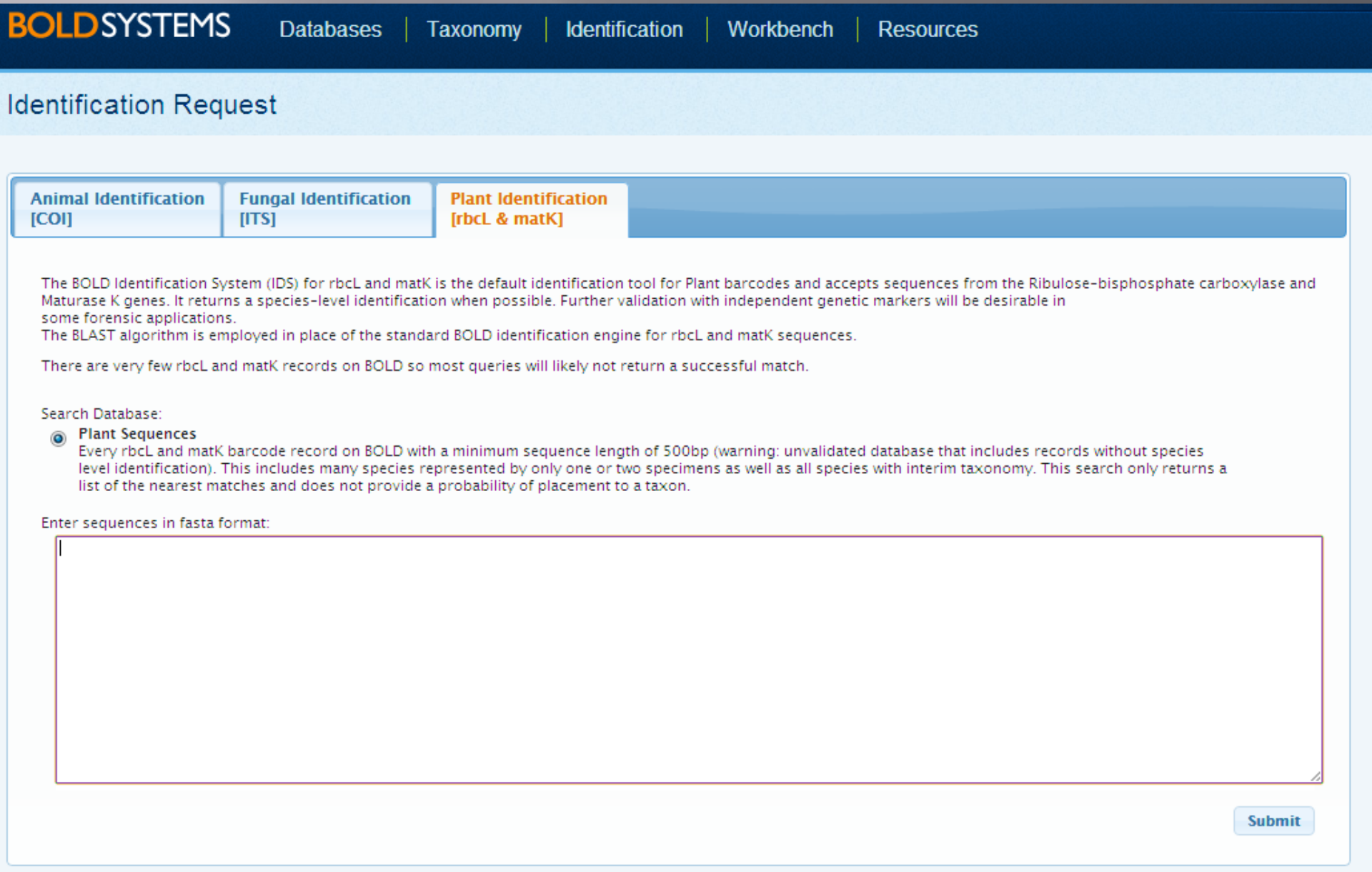
- 62 plants species collected and a few leaves cut and dried in silica
- DNA extracted from dried leaves
- *rbcL* (primers *rbcLa_F /rbcLa_R*) and *matK* (primers *KIM_F/KIM_R*) genes amplified
- PCR product run in 1% agarose gel electrophoresis at 72V
- Successful amplicons sequenced at Yale University DNA sequencing facility
- Sequences over 500 bp long were considered successful and were used to identify plant in the BOLD database (www.barcodinglife.com)


Pressing plants at SNR


100 bp ladder
~700bp
Direction of run


DNA extraction in the lab.

Identification using the BOLD



RESULTS

- 45 species amplified for *rbcL* and 38 species for *matK* and were sequenced
- 20 species had high sequence success (sequences 500 bp or higher) for both *rbcL* and *matK*
- 6 species identified with 100% match. The rest could only identify the genus with 98-99% match
- 10 identified the genus of the species although the species themselves were not on the database

RESULTS

Feature	rbcL	matK	
PCR success	72.6% (45/62)	61.3% (38/62)	
Seq > 500 bp	46.7% (21/45)	68.4% (26/38)	
Seq >400bp	73.3% (33/45)	76.3% (29/38)	

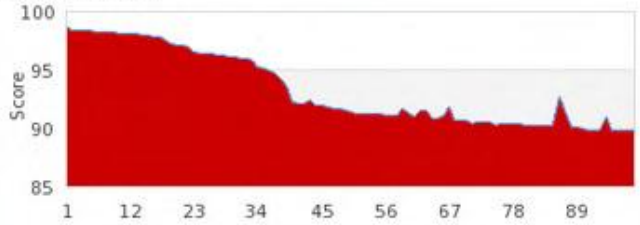
Specimen Identification Request

Query: dalea_purpurea matK ATAAAACTCTTTTTTTTTGAGGATCCATTGTAATAATGAGAAATTTTTCTGCGGATCCGCAAAAATCGGT

Search Request:
Type:

Search Result:

Score Summary:



Scores indicate the degree of similarity between the query sequence and hits. Higher is better.

E-Value Summary:

E-Values are an indicator of the likelihood that a given match was generated randomly. Lower is better.

Processid	Phylum	Class	Order	Family	Genus	Species	Score	Similarity	E-Value
GBVN4037-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*purpurea*	717	98.65	0
GBVN4029-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*melantha*	713	98.38	0
GBVN4028-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*lumholtzii*	713	98.38	0
GBVN4036-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*pulchra*	713	98.38	0
GBVN4035-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*pulchra*	713	98.38	0
GBVN4018-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*albiflora*	711	98.24	0
GBVN4040-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*versicolor*	711	98.24	0
GBVN4039-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*scandens*	711	98.24	0
GBVN4038-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*purpusii*	711	98.24	0
GBVN4027-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*leporina*	709	98.11	0
GBVN4041-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*wrightii*	709	98.11	0
GBVN4026-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*lanata*	709	98.11	0
GBVN4019-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*bicolor*	709	98.11	0
GBVN4020-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*brachystachya*	707	97.97	0
GBVN4024-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*hospes*	707	97.97	0
GBVN4033-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*pogonathera*	705	97.84	0
GBVN4021-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*cliffortiana*	705	97.84	0
GBVN4022-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*exigua*	701	97.57	0
GBVN4030-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*mollis*	697	97.3	0
GBVN4025-11	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Dalea	*lachnostachys*	695	97.16	0

CONCLUSION

- *rbcL+matK* barcode works well although cannot distinguish some of the plants species.
- Sequences for some plants are missing in the BOLD database. Sequences will have to be uploaded for barcoding to be possible.

ACKNOWLEDGEMENTS

I would like to thank :

- My mentor, David Bogler, Ph. D.
- Sandra Arango-Caro, Ph.D.
- George Yatskievych, Ph.D.
- Sharon Carter.
- NSF.
- Missouri Botanical gardens.

SELECTED SOURCES

1. CBOL plant working group. 2009. A DNA barcode for land plants. PNAS. 106. 12794-12797
2. Hollingsworth, P. M., S. W. Graham, and D. P. Little.2011. Choosing and using a plant DNA barcode. PlosOne journal. 6. Issue 5. E19254
3. Ford, D.S, K.L Ayres, N. Toomey, N. Haider *et al.* 2009. Selection of candidate coding DNA barcoding regions for use on land plants. Botanical Journal of the Linnean Society. 159, 1-11
4. Haider, N. Chloroplast-specific universal primers and their uses in plants studies. 2011. Biologia Plantarum 55 (2). 225-236.
5. Vijayan, K. and C.H Tsou. DNA barcoding in plants: taxonomy in a new perspective. Current Science. 99. Issue 11. 1530-1541.