DNA barcodes for the confirmation of identity of five chosen plant materials

1 General Information: DNA extraction, PCR, Data Analysis 2
2 Cassia senna L. (plant part: leaves) 3
3 Epimedium sp. (plant part: leaves) 4
3.1 Commercial powder #1 “Named Epimedium sagittatum” 4
3.2 Commercial powder #2 “Named Epimedium sagittatum” 6
4 Trigonella foenum-graecum L. (plant part: seeds) 7
Marrubium vulgare L. (plant part: aerial parts) 9
4.2 Herbarium specimen: Field Museum accession number: 1663369 9
4.3 Commercial Marrubium vulgare 10
5 Pausynistalia johimbe (plant part: bark) 13

Simmler Charlotte

version 20181118

1 General Information: DNA extraction, PCR, Data Analysis

DNA extraction was performed with the Qiagen Plant DNAeasy minikit following the manufacturer instructions. The final DNA solution was diluted 50 times in TAE buffer (or nuclease free water, see associated protocol) for the commercial samples of *Epimedium* sp., *Cassia senna*, and *Marrubium vulgare* prior to PCR analysis.

The universal primers used were as published by S. Chen, H. Yao, J. Han, C. Liu, J. Song, L. Shi, Y. Zhu, Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species, PLoS One. 5 (2010) e8613. doi:10.1371/journal.pone.0008613.

- For ITS-2

S2F: ATGCGATACTTGGTGTGAAT
S3R: GACGCTTCTCCAGACTACAAT

- For TrnH-PsbA

PF: GTTATGCATGAACGTAATGCTC
TR: CGCGCATGGTGGATTCACAATCC

- For RbcL

1F: ATGTCACCACAAACAGAAAC
724R: TCGCATGTACCTGCAGTAGC

- For MatK (are not optimal for *Yohimbe, Trigonella foenum-graecum, Senna*)

390F: CGATCTATTCATTCAATATTTC
1326R: TCTAGCACACGAAAGTCGAAGT

PCR reactions were performed using the GoTaqGreen master mix. Approximately 3% DMSO was added in each PCR reaction mix (except for matK). The PCR conditions were as follows:

95 °C during 5 min
95 °C for 2 min
55 °C for 30 sec**
72 °C for 1 min
Repeat 40 times
72 °C for 7 min
Keep at 4 °C

**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.

In cases where the gel electrophoresis showed a weak or no amplification (low intensity of the DNA amplicon), a second PCR was performed (or repeated) using the first PCR solution as template.

PCR clean-up prior to DNA Sanger sequencing was performed using the ExoSap-IT enzyme following the manufacturer instructions.

Sequences were analyses in ApE and SerialCloner 2-6-1, alignments were performed using BLASTN 2.8.1

⇒ See also related protocol and power-point presentation

2 *Cassia senna L. (plant part: leaves)*

Botanical Synonyms: *Senna alexandrina* Mill. (commercial powder)

2.1.1 ITS-2 sequence: **399 pb**

TCTGCCTGGGTGTCACGCATCGTTGCCCCAAAACCTTGTCGTCCCTTCGGTCAACC
GGAGGCGATGAGGTGCTTGGGCGGAAGTTGGCCTCCCGTGAGCATTGCCTCGCG
GATGGTCGAAATTGGAGCCTGTGGGGAGCGACCGCCACGTTCCACGGTGGTTGA
GCAGATGCCTCGAGGCCGACCGTGCACGAGTTGTCCCCACGTCAAAGGCTGCGA
GACCCTTGCGAGCAAGAAAGTGCTCCCAACGCGACCCCAGGTCAGGCGGGGCCA
CCCGCTGAGTTTAAGCATATCAATAAGCGGAGGAAAAGAAACTAACAAGGATTCCC
CTAGTAACGGCGAGCGAACCGGGAAAAGCCCACCATGAGAATCGGTCGTCCTCG
GCGTCCGAATTGTAGTC

Significant alignment in Genbank ***Senna alexandrina*** isolate CIMAP-C039 clone Senna species 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence Sequence ID: KY492293.1,

Score	Expect	Identities	Gaps	Strand
737 bits(399)	0.0	399/399(100%)	0/399(0%)	Plus/Plus

And with ***Senna alexandrina*** internal transcribed spacer 2, complete sequence, Sequence ID: JQ301846.1

Score	Expect	Identities	Gaps	Strand
721 bits(390)	0.0	390/390(100%)	0/390(0%)	Plus/Plus

2.1.2 psbA-trnH sequence: **190 pb**

AAATTGTGGTCTTAATATATATGAGTTTTTGAACGTAAAGGAGCAATATCAAGAGGG
TTGATATTGCTCCTTTACTTTCTTTTTTAGTAGTCTTTTTCTTCATATTCATACAAATC
TTTTTTATTTACTTCAACATTCTTTAACATTATTTTAACATAAGAAAAAAATATGCGAG
TTTCATACTTTTTTT

Representing only a clean portion of the amplified DNA

Significant alignment in Genbank with ***Senna alexandrina*** isolate KMS0206A01 psbA-trnH intergenic spacer region, partial sequence; chloroplast Sequence ID: MF097036.1

Score	Expect	Identities	Gaps	Strand
350 bits(189)	1e-92	189/189(100%)	0/189(0%)	Plus/Plus

But also with ***Chamaecrista nigricans*** psbA-trnH intergenic spacer, partial sequence; chloroplast, Sequence ID: HQ161770.1

Score	Expect	Identities	Gaps	Strand
351 bits(190)	3e-93	190/190(100%)	0/190(0%)	Plus/Plus

2.1.3 rbcL sequence**: 719 pb**

GTGAGGTTTGGGTCAAGCTGGTGTTAAGATTATAAATTGACTTATTATACTCCTGAA
TATGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCTGG
AGTTCCGCCTGAAGAAGCAGGTGCCGCGGTAGCTGCTGAATCTTCTACTGGTACA
TGGACAACTGTGTGGACCGATGGGCTTACCAGTCTTGATCGTTACAAAGGACGAT
GCTACCACATCGAGCCCGTTGCTGGAGAAGAAAATCAATATATTGCTTATGTAGCT
TATCCCTTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCCATTGTG
GGTAATGTATTTGGATTCAAGGCCCTGCGCGCTCTACGTCTGGAGGATTTGCGAAT
CCCTACTTCTTATATTAAAACTTTCCAAGGTCCGCCTCACGGCATCCAAGTTGAGA
GAGATAAATTGAACAAGTACGGCCGTCCCCTATTGGGATGTACTATTAAACCTAAA
TTGGGGTTATCCGCTAAGAATTACGGTAGAGCAGTTTATGAATGTCTCCGCGGTGG
ACTTGATTTTACCAAAGATGATGAGAATGTGAATTCCCAACCATTTATGCGTTGGAG

AGACCGTTTCTTATTTTGTGCCGAAGCAATTTTTAAAGCACAGGCCGAAACTGGTG
AAATCAAAGGGCATTACTTGAATGCTACTGCAGGTAACATGCGAAA

Significant alignment in Genbank with ***Cassia didymobotrya*** **chloroplast rbcL gene Sequence ID: Z70154.1,**

Score	Expect	Identities	Gaps	Strand
1271 bits(688)	0.0	706/714(99%)	4/714(0%)	Plus/Plus

also with ***Cassia senna*** chloroplast rbcL gene, Sequence ID: Z70155.1

Score	Expect	Identities	Gaps	Strand
1260 bits(682)	0.0	704/714(99%)	4/714(0%)	Plus/Plus

Conclusion: Collectively, all the amplified DNA sequences confirm that this commercial sample contains *Cassia senna* L. (*Senna alexandrina* Mill.), with possibly other closely related species. The overall chemical analysis corroborates the identity of *Cassia senna* with the presence of chemical markers sennosides A and B in the crude extract.

3 Epimedium sp. (plant part: leaves)

3.1 Commercial powder #1 "Named Epimedium sagittatum"

3.1.1 ITS-2 sequence: 485 pb

AGGCTGGGACTCGAGTCTTTGACGCAAGTTGCGCCCAAGGCCATTAGGTCGAGG
GCACGTCTGCCTGGGCGTCACGCACAGCGTCGCTCCCACCATTATGCCTTTGTTC
TCTTATCGGGCAACTGCAACGTGGCTTGGGAAGCGGATATTGGCCCCCCGTACCT
TTGTAGGCGCGGCCGGCCTAAAATTCGGCCCTCGGCGACGAGCGTCACGATCAG
TGGTGGTTGAATAACCCCTTTGTCATAGACCGGTATCGTGTTGTTTCGTCGTCTATT
TGGGCCACATGGACCCTTGCGTGTCGTATAAACGACATTCACTCTGCGACCCCAG
GTCAGGCGGGACTACCCGCTGAGTTTAAGCATATCAATAAGCGGAGGAGAAGAAA
CTTACAAGGATTCCCTTAGTAACGGCGAGCGAACCGGGATCAGCCCAGCTTGGGA
ATCGGGCGACTTCGTTGTCCGAATTGTAGTCTGGAAAAAGCGTCA

Significant alignment in Genbank with: ***Epimedium wushanense*** isolate horny_goat_weed 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence. Sequence ID: KU724206.1

Score	Expect	Identities	Gaps	Strand
846 bits(458)	0.0	461/462(99%)	1/462(0%)	Plus/Plus

But also with ***Epimedium simplicifolium*** isolate Y1 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene, complete sequence; and internal transcribed spacer 2, partial sequence, Sequence ID: KC494651.1

Score	Expect	Identities	Gaps	Strand
791 bits(428)	0.0	460/475(97%)	3/475(0%)	Plus/Plus

Significant alignment with ***Epimedium koreanum*** genes for 18S rRNA, ITS1, 5.8S rRNA, ITS2, 28S rRNA, partial and complete sequence, Sequence ID: AB758657.1

Score	Expect	Identities	Gaps	Strand
652 bits(353)	0.0	356/357(99%)	1/357(0%)	Plus/Plus

3.1.2 psbA-trnH sequence: 371 pb

AATATTTTTTGCTTCAAACTCACATAGGCATCTTTTTGCTTTTTCAGGAAAAGAAAAA
AAAGCATTACTATAACTACCGGATTCATTCTAGATTATCCTAATGTTGTATATCTTCG
TCAAAATTGTTTCTAATAATTTTTTTATCAACTTAAAGTGAAAGTTCACTAATTCGTAT
TAGTGTATTGTTTTCGAATGAATAGTCTGTCAATTTTTCTTTACTTTGCAAAAAAGTA
TAAAGAAAAACGAAATGAAAAAAGTCATGTAATACTAAGTAGAGTGGTAGAGTGGA

AATTTCATTCTTTTGGGTTAAGAAGTTGAGAAATTAAAGGGGCGGATGTAGCCAAG
TGGATTAAGGCAGTGGATTGTGAATC

Significant alignment in Genbank with ***Epimedium koreanum*** chloroplast, complete genome, Sequence ID: KM207675.1

Score	Expect	Identities	Gaps	Strand
641 bits(347)	4e-180	369/378(98%)	8/378(2%)	Plus/Minus

But also with ***Epimedium acuminatum*** chloroplast, complete genome, Sequence ID: KU522469.1

Score	Expect	Identities	Gaps	Strand
641 bits(347)	4e-180	364/372(98%)	1/372(0%)	Plus/Minus

And with ***Epimedium sagittatum*** chloroplast, complete genome, Sequence ID: KU204899.1

Score	Expect	Identities	Gaps	Strand
641 bits(347)	4e-180	364/372(98%)	1/372(0%)	Plus/Minus

3.1.3 rbcL sequence: 718 pb

GGCGTTTGCATCAGCGGGTGTTAAGATTACAAATTGACTTATTATACTCCTGACTAT
GTAACGAAGGATACTGATATTTTGGCAGCATTCCGCGTCACTCCTCAACCTGGAGT
TCCACCTGAAGAAGCAGGGGCCGCTGTAGCTGCCGAATCTTCTACAGGTACATGG
ACAACCGTGTGGACCGATGGACTTACCAGTCTTGATCGTTACAAAGGACGGTGCT
ACCACATTGAGCCTGTTGCTGGAGAAGACAATCAATATATTTGTTACGTAGCCTATC
CTTTAGACCTTTTTGAAGAGGGTTCTGTTACTAACATGTTTACTTCTATTGTGGGTA
ATGTATTTGGGTTCAAAGCGCTGCGCGCTCTACGTCTGGAGGATCTGCGAATTCCT
GTTGCTTATGTTAAAACTTTCCAAGGCCCGCCTCATGGTATCCAAGTTGAGAGAGA
TAAATTGAACAAGTATGGTCGTCCTCTATTAGGATGTACTATTAAACCAAAATTGGG
ATTATCCGCTAAGAACTATGGTAGAGCGGTTTATGAATGTCTCCGCGGTGGGCTTG
ATTTTACCAAGGATGATGAGAACGTGAACTCCCAGCCATTTATGCGTTGGAGAGAT
CGTTTCCTATTTTGTGCCGAAGCTATTTATAAATCACAGGCGGAAACAGGTGAAAT
CAAAGGACATTACTTGAATGCTACTGCAGGTAACATGGCGAAAT

Significant alignment in Genbank with ***Epimedium koreanum*** chloroplast, complete genome Sequence ID: KM207675.1

Score	Expect	Identities	Gaps	Strand
1277 bits(691)	0.0	694/695(99%)	1/695(0%)	Plus/Plus

But also with ***Epimedium dolichostemon*** chloroplast, complete genome, Sequence ID: KU522470.1

Score	Expect	Identities	Gaps	Strand
1266 bits(685)	0.0	692/695(99%)	1/695(0%)	Plus/Plus

And with ***Epimedium sagittatum*** chloroplast, complete genome, Sequence ID: KU204899.1

Score	Expect	Identities	Gaps	Strand
1260 bits(682)	0.0	691/695(99%)	1/695(0%)	Plus/Plus

…and other closely related ***Epimedium*** species.

3.1.4 matK sequence: 849 pb

CTACCCCATCCATTTTGAACTCTTGATTCAAACCCTTCGCTATTGGATACAGGATAC
CCCCGCTTTGCATTTATTACGATTCTTTCTCTACGAGTCTCAGAATTCGAATAATCT
GATTACTCAAAAAAAAAGAGATATTTCGCATTTTTCAAATCAGAATCAAAGATTTTTC
TTGTTCCTATATAATATTCATATATATGAATGCGAATCCATATTCGTTTTTCTCCGTA
AACAATCTGTTCATTTACGATCAAGATCGTATAGAGCCCTTCTTGAGCGAACACATT
TTTATCGAAAAATAGACAAGTTTTTCTTCATTTTTCATAAAAATTTTCAGACCACCTTA
TGGTTGTTCAAGGATCCTTTCATGAATTATGTCAGATATCAAGGAAAAGCCATTCTG
GCTTCAAAAGGAACACCTCTTCTGATAAAAAAATGGAAGTATTACCTTGTCAATTTT

TGTCAAGGTTATTTTGACTTGTGGCCTCAACCAGATAGAATTCAAATAAACCAATTC
TCCAAGCACTCCCTCGATTTTCTGGGCCATCTTTCAAGTTTACGGCTAAAGCCTTGT
GTGGTAAGGAGTCAAATGTTAGAAAATTCATTTATTATAGATGTTTCTATTAATAAGT
TTGATACTATAGTCCCCACAATTCCTTTGATAGGAGCATTGGCTAAAGCGAAATTTT
GTAACGTATCGGGGCATCCTATTAGTAAGCCGGCTTGGACAGATTCTGCAGATTCT
GATATTATCGATAGATTTGCGCGTATATGCAGAAATCTTTTTCATTTTTTTAGTGGAT
CCGCAGAAAAAAATACTTTGTTTCGAGTAAAGTACATACTTCGACT

Significant alignment in Genbank with ***Epimedium koreanum*** chloroplast, complete genome Sequence ID: KM207675.1,

Score	Expect	Identities	Gaps	Strand
1557 bits(843)	0.0	847/849(99%)	0/849(0%)	Plus/Minus

….But also with ***Epimedium grandiflorum*** tRNA-Lys (trnK) gene, partial sequence; and maturase K (matK) gene, complete cds; chloroplast, Sequence ID: JN010335.1

Score	Expect	Identities	Gaps	Strand
1552 bits(840)	0.0	846/849(99%)	0/849(0%)	Plus/Plus

…And with ***Epimedium sagittatum*** chloroplast, complete genome, Sequence ID: KU204899.1

Score	Expect	Identities	Gaps	Strand
1546 bits(837)	0.0	845/849(99%)	0/849(0%)	Plus/Minus

Conclusion: The commercial sample contains *Epimedium* species. There was slightly lower alignment scores and sequence homologies with *E. sagittatum*. The chemical analysis indicated the presence of characteristic compounds such as icariin.

3.2 Commercial powder #2 "Named Epimedium sagittatum"

3.2.1 psbA- trnH: **259 pb**

TTTTTCTATTCAAGTAATTGAATAATTAAACTTGCATAAGAACGACCCGACCTTTTGT
AATTGTAAAGAGGTTGGGTCGTTCTTGTGCATTGAATATTTTTTGCTTCAAACTCAC
ATAGGCATCTTTTTTCTTTTTCAGGAAAAGAAAAAAAGCATTACTATAACTACCGGA
ATCATTCCAAATTATCCTAATTCCTAATGTTGTATATCTTCGTCAAAATTGTTTCTTAT
ATTTTTTTTATCAACTTAAAGTGAAAGT

Representing only a clean portion of the amplified DNA

Significant alignment in Genbank with ***Epimedium koreanum*** chloroplast, complete genome, Sequence ID: KM207675.1

Score	Expect	Identities	Gaps	Strand
451 bits(244)	4e-123	254/259(98%)	0/259(0%)	Plus/Minus

And Sequence ID: KU522471.1

Score	Expect	Identities	Gaps	Strand
451 bits(244)	4e-123	254/259(98%)	0/259(0%)	Plus/Minus

The homologies for ***Epimedium sagittatum*** species was less significant:

Epimedium sagittatum voucher PS1502MT03 psbA-trnH intergenic spacer region, partial sequence; chloroplast, Sequence ID: GQ435416.1

Score	Expect	Identities	Gaps	Strand
374 bits(202)	1e-99	242/260(93%)	8/260(3%)	Plus/Plus

3.2.2 rbcL sequence: **722 pb**

CAGAAAGGGTTGGGGAATCAAGGCGGGTGTTAAGATTACAATTGACTTATTATACT
CCTGACTATGTAACGAAGGATACTGATATTTTGGCAGCATTCCGCGTCACTCCTCA
ACCTGGAGTTCCACCTGAAGAAGCAGGGGCCGCTGTAGCTGCCGAATCTTCTACA

GGTACATGGACAACCGTGTGGACCGATGGACTTACCAGTCTTGATCGTTACAAAG
GACGGTGCTACCACATTGAGCCTGTTGCTGGAGAAGACAATCAATATATTTGTTAC
GTAGCCTATCCTTTAGACCTTTTTGAAGAGGGTTCTGTTACTAACATGTTTACTTCT
ATTGTGGGTAATGTATTTGGGTTCAAAGCGCTGCGCGCTCTACGTCTGGAGGATCT
GCGAATTCCTGTTGCTTATGTTAAAACTTTCCAAGGCCCGCCTCATGGTATCCAAG
TTGAGAGAGATAAATTGAACAAGTATGGTCGTCCTCTATTAGGATGTACTATTAAAC
CAAAATTGGGATTATCCGCTAAGAACTATGGTAGAGCGGTTTATGAATGTCTCCGC
GGTGGGCTTGATTTTACCAAGGATGATGAGAACGTGAACTCCCAGCCATTTATGCG
TTGGAGAGATCGTTTCCTATTTTGTGCCGAAGCTATTTATAAATCACAGGCGGAAA
CAGGTGAAATCAAAGGACATTACTTGAATGCTACTGCAGGTACATGCGAA

Significant alignment in Genbank with ***Epimedium koreanum*** chloroplast, complete genome Sequence ID: KM207675.1

Score	Expect	Identities	Gaps	Strand
1284 bits(695)	0.0	706/711(99%)	2/711(0%)	Plus/Plus

But also with ***Epimedium pinnatum*** ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; plastid, Sequence ID: KM360766.1

Score	Expect	Identities	Gaps	Strand
1284 bits(695)	0.0	706/711(99%)	2/711(0%)	Plus/Plus

And also with ***Epimedium sagittatum*** chloroplast, complete genome, Sequence ID: KU204899.1

Score	Expect	Identities	Gaps	Strand
1267 bits(686)	0.0	703/711(99%)	2/711(0%)	Plus/Plus

Conclusion: The commercial sample contains *Epimedium* species. The sequences obtained for *psbA*-trnH and *matK* were very dirty showing overlap of bases, suggesting a mixture of DNA amplicons. Overall the best sequence was obtained for rbcL with a significant alignment for several DNA sequences of *Epimedium species including E. koreanum, and to a lesser extent E. sagittatum*. The chemical analysis indicated the presence of characteristic compounds such as icariin.

4 *Trigonella foenum-graecum L. (plant part: seeds)*

Commercial powder

4.1.1 ITS-2 sequence: **398 pb**

CACGTCTGCCTGGGTGTCACATATCGAAGCCTCATGCCAATTTCCTTTTTTAGTAG
GTATTGTGCATGCTGGTGAATGTTGGCCTCCCGTGAGCTCTATTGTCTCATGGTTG
GTTGAAAATCGAGACCTTGGTAGGGTGTGCCATGATAGATGGTGGTTGTGTGACC
CACGAGAACCAAGATCATGTGCTTCCCTATTCAATTTGGCCTCTTTTACCCATATGC
GTTTTGTGAACGCTCGTGATGAGACCTCAGGTCAGGCGGGGCTACCCGCTGAATT
TAAGCATATCAATAAGCGGAGGAAAAGAAACTAACAAGGATTCCCTTAGTAACGGC
GAGCGAACCGGGATAAGCCCACCATGAAAATCGGTCGCCTTCGGCGTTCGAATTG
TAGTCTGG

Significant alignment in Genbank with *Trigonella foenum-graecum* voucher 65230283 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence, Sequence ID: KF454107.1,

Score	Expect	Identities	Gaps	Strand
713 bits(386)	0.0	386/386(100%)	0/386(0%)	Plus/Plus

as well as voucher 200505082 5.8S ribosomal RNA gene, partial sequence; internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence, Sequence ID: MH688883.1

Score	Expect	Identities	Gaps	Strand
710 bits(384)	0.0	385/386(99%)	0/386(0%)	Plus/Plus

4.1.2 psbA-trnH sequence: **358 pb**

TAGGGGTACTTCAACGACATGGCTCTCCGCCCTATACTATGTCTAAATTACAGAAC
TTTTTATACCTTTTGATTATCATTTATTTTCTTCTCTTCTAAGTTCTTACGAGTCTTCT
AAAATACTATTCTAAAATACTAGAAAATAAAATAAGAAAACATTTATTTCTGTATATGT
ATACATACTAACAAAAAGTAAAGGGGGGAGTAATATCAACAATGTTGATATTACTCC
TTTACCCCTTTACTTTCAAAAAGTCGTATCCTTCTTTAAAACAAAAATATTATCCATTT
ATAGATGGAGCCTCGACCGCAGCTAGGTCTAGAGGGAAATTATGAGCATTACGTT
CATGCATAACAAA

Significant alignment *Trigonella foenum-graecum* PsbA (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His (trnH) gene, partial sequence; chloroplast, Sequence ID: MG947135.1

Score	Expect	Identities	Gaps	Strand
604 bits(327)	4e-169	327/327(100%)	0/327(0%)	Plus/Minus

4.1.3 rbcL sequence: **719 pb**

CGAGAGGTGGGGTTCAAGCTGGTGTTAAGATTATAAATTGACTTATTATACTCCTG
ACTATGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAAGTCCTCAACCT
GGAGTTCCTGCTGAAGAAGCAGGTGCAGCGGTAGCTGCCGAATCTTCCACTGGGA
CATGGACAACTGTGTGGACCGATGGACTTACCAGTCTTGATCGTTATAAAGGACGC
TGCTACCACATCGAGCCTGTTGCTGGCGAAGATAATCAATTTATTGCTTATGTAGCT
TATCCCTTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACCTCCATTGTA
GGTAATGTATTTGGGTTCAAGGCCTTGCGCGCTCTACGTCTGGAAGATTTGCGAAT
CCCAGTTGCTTATGTTAAAACTTTCCAAGGCCCTCCTCACGGAATCCAAGTTGAGA
GAGATAAATTGAACAAGTATGGACGTCCCTTATTGGGATGTACTATTAAACCTAAAT
TGGGGTTATCCGCTAAGAATTACGGTAGAGCAGTTTATGAATGTCTACGCGGTGGA
CTTGATTTTACCAAAGATGATGAAAATGTGAACTCCCAACCATTTATGCGTTGGAGA
GACCGTTTCTTATTTTGTGCCGAAGCTATTTATAAATCACAGGCCGAAACAGGTGA
AATCAAAGGACATTATTTGAATGCTACTGCAGGTACATGCGAAA

Significant alignment in Genbank with *Trigonella foenum-graecum* ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit gene, partial cds; chloroplast, Sequence ID: MG946901.1

Score	Expect	Identities	Gaps	Strand
1293 bits(700)	0.0	711/716(99%)	2/716(0%)	Plus/Plus

Conclusion: All the DNA barcode sequences together with the chemical analyses confirm that the sample contains *T. foenum graceum* (seeds).

Marrubium vulgare L. (plant part: aerial parts)

4.2 Herbarium specimen: Field Museum accession number: 1663369

Specimen collected in 1966, in Peru

Serving as reference Sequences (DNA extracted from the one leaf)

4.2.1 ITS-2 sequence: **390 pb**

CACGTCTGCCTGGGCGTCACGCATCGCGTCGCCCCCCTCCCCCGCGGGGTGGGG
GGGCGGAGATTGGCCCCCCGCGCGCGCACGTCCGCGCGCGGCCGGCCCAAATG
CCAATCCGCCGTCGACTCACGTCGCGACCAGTGGTGGTTGAACTATCAACTCGCG
TGCTGTCGCGCTCCACGGCGTCGTCGGTCCGGAAACAGCAACGCAACCCAACGG
CGCGAGCACGCATCGTGCCCACGACCGCGACCCCAGGTCAGGCGGGATCACCCG
CTGAGTTTAAGCATATCAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAG
TAACGGCGAGCGAACCGGGAATAGCCCAACTTGAGAATCGGGCGGCCACGCCGT
CCGAATTGTA

Lack of *M. vulgare* ITS/ ITS-2 DNA sequences in Genbank!

Sequence alignments (97% identities) were obtained with *Phlomis mongolica* voucher PS1732MT01 internal transcribed spacer 2 and 28S ribosomal RNA gene, partial sequence, Sequence ID: FJ546872.1

Score	Expect	Identities	Gaps	Strand
499 bits(270)	2e-137	289/298(97%)	2/298(0%)	Plus/Plus

4.2.2 psbA- trnH sequence: **260 pb**

TGGTCTTAGTGTATAGGAGTTTTTGAAAATAGAATAGATAAATATAAGGAGCAATAA
CCCCTCTTGATAAAACAAGAAAGAGTTTATTAGCTCCTTAATTTTCTTTTCAATTACT
TTTTTCCTTTCCATTAAAGGATTCAGAAAATGAAAGAAGAAAAAAAACGATTGAAATT
AAAAGTAAATTGAATTTCATTATACTATTTCATTATACTAATAGTTGAGGGGCGGAT
GTAGCCAAGTGGATCAAGGCAGTGGATTGT

Sequence alignment in Genbank with *Marrubium vulgare* voucher SA1664 psbA-trnH intergenic spacer, partial sequence; chloroplast, Sequence ID: HQ902823.1

Score	Expect	Identities	Gaps	Strand
451 bits(244)	2e-129	251/254(99%)	2/254(0%)	Plus/Plus

4.2.3 rbcL sequence: **659 pb**

CCCCTGAATACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCT
CAACCTGGAGTTCCGCCCGAAGAAGCAGGGGCTGCGGTAGCTGCCGAATCTTCG
ACTGGTACATGGACAACTGTGTGGACCGATGGACTTACCAGCCTTGATCGTTACAA
AGGGCGATGCTACCACATCGAGCCTGTTCTTGGAGAAAAAGATCAATATATCTGTT
ATGTAGCTTACCCTTTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTC
CATTGTAGGAAATGTATTTGGATTCAAAGCCCTACGTGCTCTACGTCTGGAAGATC
TGCGAATCCCTCCTGCTTATGTTAAAACTTTCCAAGGCCCACCTCATGGGATCCAA
GTTGAGAGAGATAAATTGAACAAGTATGGTCGTCCTCTGTTGGGATGTACTATTAA
ACCGAAATTGGGGTTATCTGCTAAAAACTATGGTAGAGCAGTTTATGAATGTCTTC
GCGGTGGACTTGATTTTACCAAAGATGATGAAAACGTGAACTCCCAGCCATTTATG
CGTTGGAGAGATCGCTTCTTGTTTTGTGCCGAAGCAATTTATAAATCACAGGCTGA
AACAGGTGAAATCAAAGGGCATTATTTGAATGCTACTGCAGGT

Significant alignment in Genbank with *M. vulgare* (I) chloroplast rbcL gene for rubisco (large subunit) (partial), Sequence ID: Z37411.1

Score	Expect	Identities	Gaps	Strand
1206 bits(653)	0.0	657/659(99%)	0/659(0%)	Plus/Plus

4.2.4 matK sequence: **850 pb**

ATGGTAATACCTCGCCCTGTTCATGCGGAAATCTTGGTTCAAATCCTTCGCTATTG
GGTAAAAGATGCTTCCTCCTTGCATTTATTACGAGTCTTTCTCAACGAATATTGTAG
TTGGAATAGTCTTCTTATTCCAAAGAAAGCCAGTTCCCCGTCTTTAAAAAAAAATCA
AAGATTATTCTTATTCTTATATAATTCTCATGTATGCGAATATGAATCCATTTTCGTCT
TTCTACGTAACCAATCTTTTCATTTACGATCAACATCTTCTGGAGTTTTTCTTGAACG
AATATATTTCTATATAAAAATAGAACGTCTTGTGAACGTCTTTGTTAAGATTACGGAT
TTGGGGGCAAACCTGCGGTTGGTCAAGGAACCTTTCATGCATTATATTAGGTATCA
AAAAAGATCCATTCTGGCTTCAAAAGGGACATTTCTTTTCATGAAGAAATGGAAATT
TTACCTTGTCACCTTTTGGCAATGGCATTTTTTGGTGTGGTTTCATTCAAGAAGCAT
TTATATAAACCAATTATCCAAGCATTCCCTTGAGTTTTTGGGCTATCTTTCAAGCGT
GCAAATGAACCCTTCCGTAGTACGCAGTCAAATTATAGAAAATTCATTTCTAATCAA
TAATGCTATTAAGAAGTTTGAAACTCTTGTTCCAATTATTCCTCTGATTGCGTCATTG
GCTAAAGCAAAATTTTGTAACGTATTGGGGCATCCCGTTAGTAAGCCGATTCGGAC
TGATTTATCAGATTCTAATATTATTGACCGATTTGGGCGTATATGCAGAAATATTTCT
CGTTATCATAGTGGATCTTCAAAAAAAAAGAGTTTGTCTCGAATAAAGT

Significant alignment in Genbank with *Marrubium vulgare* isolate NMW529 maturase K (matK) gene, partial cds; chloroplast, Sequence ID: JN895787.1

Score	Expect	Identities	Gaps	Strand
1469 bits(795)	0.0	795/795(100%)	0/795(0%)	Plus/Plus

4.3 Commercial Marrubium vulgare

4.3.1 ITS-2: **290 pb**

GCTCGCCGTTACTAGGGGAATCCTTGTAAGTTTCTTTTCCTCCGCTTATTGATATGC
TTAAACTCAGCGGGTGATCCCGCCTGACCTGGGGTCGCGGTCGTGGGCACGATG
CGTGCTCGCGCCGTTGGGTTGCGTTGCTGTTTCCGGACCGACGACGCCGTGGAG
CGCGACAGCACGCGAGTTGATGGTTCAACCACCACTGGTCGCGACGCGGGTCGA
CGGCGGATTCGCATTTGGGCCGGCCGCGCGCGGGCGTGCGCGCACGGGGGGCC
AATCTCCGCCCCCCACCC

Representing only a clean portion of the DNA amplicon (really dirty sequence).

Significant sequence alignment in GenBank with *Marrubium supinum* 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence, Sequence ID: AF335642.1

Score	Expect	Identities	Gaps	Strand
379 bits(205)	2e-101	217/223(97%)	0/223(0%)	Plus/Plus

Other significant sequence alignment with *Phlomis mongolica* voucher PS1732MT01 internal transcribed spacer 2 and 28S ribosomal RNA gene, partial sequence, Sequence ID: FJ546872.1

Score	Expect	Identities	Gaps	Strand
412 bits(223)	2e-111	242/251(96%)	2/251(0%)	Plus/Minus

As there is a lack of ITS-2 sequences for *M. vulgare* in GenBank, the sequence was also compared to the reference Herbarium (#1663369) ITS-2 sequence:

```
Alignment of Sequence_1:  [Cleaned ITS2 sequence from Commercial sample/reverse complement, 290
pB] with  Sequence_2: [Herbarium_MvulgareITS2_forward.xdna, 291 pB] with SerialCloner 2-6-1

Similarity : 284/291 (97.59 %)

Seq_1  290   GGGT-GGGGGGCGGAGATTGGCCCCCCGTGCGCGCACGCCCGCGCGCGGCCGGCCCAAAT  232
             ||||#||||||||||||||||||||#|||||||||#||||||||||||||||||||||
Seq_2  1     GGGTGGGGGGGCGGAGATTGGCCCCCCGCGCGCGCACGTCCGCGCGCGGCCGGCCCAAAT  60
```

```
Seq_1  231   GCGAATCCGCCGTCGACCCGCGTCGCGACCAGTGGTGGTTGAACCATCAACTCGCGTGCT  172
              ||#|||||||||||||||#|#|||||||||||||||||||||||#|||||||||||||||
Seq_2  61    GCCAATCCGCCGTCGACTCACGTCGCGACCAGTGGTGGTTGAACTATCAACTCGCGTGCT  120


Seq_1  171   GTCGCGCTCCACGGCGTCGTCGGTCCGGAAACAGCAACGCAACCCAACGGCGCGAGCACG  112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  121   GTCGCGCTCCACGGCGTCGTCGGTCCGGAAACAGCAACGCAACCCAACGGCGCGAGCACG  180


Seq_1  111   CATCGTGCCCACGACCGCGACCCCAGGTCAGGCGGGATCACCCGCTGAGTTTAAGCATAT  52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  181   CATCGTGCCCACGACCGCGACCCCAGGTCAGGCGGGATCACCCGCTGAGTTTAAGCATAT  240


Seq_1  51    CAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGAGC  1
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  241   CAATAAGCGGAGGAAAAGAAACTTACAAGGATTCCCCTAGTAACGGCGAGC  291
```

4.3.2 psbA-trnH: **173 pb**

GGATAAGACTTGGTCTTAGTGTATAGGAGTTTTTGAAAATAGAATAGATAAATATAA
GGAGCAATAACCCCTCTTGATAAAACAAGAAAGAGTTTATTAGCTCCTTAATTTTCT
TTTCAATTACTTTTTTCCTTTCCATTAAAGGATTCAGAAAATGAAAGAAGAAAAAAAA
C

Representing only a clean portion of the DNA amplicon.
Significant alignment in Genbank with *Marrubium crassidens* voucher SKU:outgroup maturase K (matK) gene, partial cds; chloroplast, Sequence ID: KP993196.1

Score	Expect	Identities	Gaps	Strand
313 bits(169)	1e-81	172/173(99%)	1/173(0%)	Plus/Plus

Alignment with Marrubium peregrinum voucher dimou1024TAU PsbA (psbA) gene, partial sequence; psbA-trnH intergenic spacer, complete sequence; and tRNA-His (trnH) gene, partial sequence; chloroplast Sequence ID: EU627584.1

Score	Expect	Identities	Gaps	Strand
307 bits(166)	6e-80	171/173(99%)	1/173(0%)	Plus/Plus

and with *Marrubium vulgare* voucher SA1664 psbA-trnH intergenic spacer, partial sequence; chloroplast, Sequence ID: HQ902823.1

Score	Expect	Identities	Gaps	Strand
281 bits(152)	3e-72	157/159(99%)	1/159(0%)	Plus/Plus

The sequence was also compared to the reference Herbarium (#1663369) psbA-trnH sequence:

```
Alignment of Sequence_1:  [CommercialMvulgare_TP_forward173pb.xdna] with  Sequence_2:
[Herbarium_Mvulgare_TP trimmed to match the number of pB.xdna]

Similarity : 173/173 (100.00 %)

Seq_1  1     GGATAAGACTTGGTCTTAGTGTATAGGAGTTTTTGAAAATAGAATAGATAAATATAAGGA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  1     GGATAAGACTTGGTCTTAGTGTATAGGAGTTTTTGAAAATAGAATAGATAAATATAAGGA  60


Seq_1  61    GCAATAACCCCTCTTGATAAAACAAGAAAGAGTTTATTAGCTCCTTAATTTTCTTTTCAA  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  61    GCAATAACCCCTCTTGATAAAACAAGAAAGAGTTTATTAGCTCCTTAATTTTCTTTTCAA  120


Seq_1  121   TTACTTTTTTCCTTTCCATTAAAGGATTCAGAAAATGAAAGAAGAAAAAAAAC  173
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  121   TTACTTTTTTCCTTTCCATTAAAGGATTCAGAAAATGAAAGAAGAAAAAAAAC  173
```

There was no gap detected with our herbarium sequence.

4.3.3 rbcL sequence: **649 pb**

TATACCCCTGAATACGAAACCACAATGATACTGATATCTTGGCAGCATTCCGAGTA
ACTCCTCAACCTGGAGTTCCGCCCGAAGAAGCAGGGGCCGCGGTAGCTGCCGAA
TCTTCGACTGGTACATGGACAACTGTGTGGACCGATGGACTTACCAGCCTTGATCG
TTACAAAGGGCGATGCTACCACATCGAGCCTGTTCTTGGAGAAAAAGATCAATATA
TCTGTTATGTAGCTTACCCTTTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTT
TACTTCCATTGTAGGAAATGTATTTGGATTCAAAGCCCTACGTGCTCTACGTCTGGA
AGATCTGCGAATCCCTCCTGCTTATGTTAAAACTTTCCAAGGCCCACCTCATGGGA
TCCAAGTTGAGAGAGATAAATTGAACAAGTATGGTCGTCCTCTGTTGGGATGTACT
ATTAAACCGAAATTGGGGTTATCTGCTAAAAACTATGGTAGAGCAGTTTATGAATGT
CTTCGCGGTGGACTTGATTTTACCAAAGATGATGAAAACGTGAACTCCCAGCCATT
TATGCGTTGGAGAGATCGCTTCTTGTTTTGTGCCGAAGCAATTTATAAATCACAGG
CTGAAACAGGTGAAATCAAAGGGCATTATTT

Significant alignment in Genbank with *M. vulgare* (I) chloroplast rbcL gene for rubisco (large subunit) (partial), Sequence ID: Z37411.1

Score	Expect	Identities	Gaps	Strand
1186 bits(642)	0.0	647/649(99%)	2/649(0%)	Plus/Plus

4.3.4 matK sequence: 774 pb

TGGTTCAAATCCTTCGCTATTGGGTAAAAGATGCTTCCTCCTTGCATTTATTACGAG
TCTTTCTCAACGAATATTGTAGTTGGAATAGTCTTCTTATTCCAAAGAAAGCCAGTT
CCCCGTCTTTAAAAAAAAATCAAAGATTATTCTTATTCTTATATAATTCTCATGTATG
CGAATATGAATCCATTTTCGTCTTTCTACGTAACCAATCTTTTCATTTACGATCAACA
TCTTCTGGAGTTTTTCTTGAACGAATATATTTCTATATAAAAATAGAACGTCTTGTGA
ACGTCTTTGTTAAGATTACGGGTTTGGGGGCAAACCTGCGGTTGGTCAAGGAACCT
TTCATGCATTATATTAGGTATCAAAAAAGATCCATTCTGGCTTCAAAAGGGACATTT
CTTTTCATGAAGAAATGGAAATTTTACCTTGTCACCTTTTGGCAATGGCATTTTTTG
GTGTGGTTTCATTCAAGAAGCATTTATATAAACCAATTATCCAAGCATTCCCTTGAG
TTTTTGGGCTATCTTTCAAGCGTGCAAATGAACCCTTCCGTAGTACGCAGTCAAATT
ATAGAAAATTCATTTCTAATCAATAATGCTATTAAGAAGTTTGAAACTCTTGTTCCAA
TTATTCCTCTGATTGCGTCATTGGCTAAAGCAAAATTTTGTAACGTATTGGGGCATC
CCGTTAGTAAGCCGATTCGGACTGATTTATCAGATTCTAATATTATTGACCGATTTG
GGCGTATATGCAGAAATATTTCTCGTTATC

Significant alignment in Genbank with *Marrubium crassidens* voucher SKU:outgroup maturase K (matK) gene, partial cds; chloroplast, Sequence ID: KP993196.1

Score	Expect	Identities	Gaps	Strand
1419 bits(768)	0.0	772/774(99%)	0/774(0%)	Plus/Plus

…and with *Marrubium vulgare* maturase K (matK) gene, partial cds; chloroplast, Sequence ID: HM850794.1

Score	Expect	Identities	Gaps	Strand
1413 bits(765)	0.0	771/774(99%)	0/774(0%)	Plus/Plus

Conclusions: the ITS-2 and psbA-trnH sequences for the commercial samples were dirty, possibly indicating a mixture of closely related *Lamiaceae* species. Comparing the acquired DNA sequences with those obtained from the Herbarium specimen together with the significant alignment obtained in Blastn for rbcL and matK sequences led to the confirmation that this commercial sample contains *M. vulgare*. However, from a chemical point of view, no marrubiin was detectable (see associated ppt) in the crude extract prepared with the commercial sample. This botanical material may contain closely related species in addition to *M. vulgare*.

5 *Pausynistalia johimbe (plant part: bark)*

Basionym: *Corynanthe johimbe* K. Schum.

Commercial sample: powder

5.1.1 rbcL sequence: **658 pb**

TTGACTTATTATACTCCCGAATATCAACCCCAGGATACCGATATCTTGGCAGCATTC
CGAGTAACCCCTCAACCTGGAGTTCCGTCAGAAGAAGCAGGAGCCGCAGTAGCTG
CCGAATCTTCTACTGGTACATGGACAACTGTATGGACCGACGGACTTACCAGTCTT
GATCGTTACAAAGGACGATGCTACCACATCGATGCCGTTCCTGGAGAAGACAATCA
ATATATTTGTTATGTAGCTTACCCATTAGACCTTTTTGAAGAAGGTTCTGTTACTAAT
ATGTTTACTTCCATTGTGGGTAATGTATTTGGGTTCAAGGCCCTGCGTGCTCTACG
TTTGGAGGATTTGCGAATCCCTGTTGCTTATATAAAAACTTTCCAAGGCCCGCCTCA
CGGTATCCAAGTTGAGAGAGATAAATTGAACAAGTATGGCCGTCCTCTACTGGGAT
GCACTATTAAGCCGAAATTGGGGTTATCCGCTAAAAACTATGGTCGAGCAGTTTAT
GAATGTCTTCGCGGTGGACTTGATTTTACCAAAGATGACGAAAACGTGAACTCCCA
GCCATTTATGCGTTGGAGAGACCGTTTCTTATTTTGTGCCGAAGCAATTTATAAAGC
ACAGGCCGAAACAGGTGAAATCAAAGGGCATTACTTG

Significant sequence alignment with *Opuntia maxima* (!) ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast, Sequence ID: HM850212.1

Score	Expect	Identities	Gaps	Strand
1214 bits(657)	0.0	657/657(100%)	0/657(0%)	Plus/Plus

This alignment does not make sense in light of the powder (macro)/microscopic characteristics and its chemistry.

The alignment with *Pausynistalia* species gave less homologies (identity percentage 91%):
Alignment with Pausinystalia lane-poolei subsp. ituriense chloroplast partial rbcL gene for ribulose-1, 5-bisphosphate carboxylase, specimen voucher Hart 1350 (BR), Sequence ID: AJ346999.1

Score	Expect	Identities	Gaps	Strand
893 bits(483)	0.0	600/658(91%)	2/658(0%)	Plus/Plus

Alignment with Pausinystalia johimbe chloroplast partial rbcL gene for ribulose-1, 5-bisphosphate carboxylase, Sequence ID: AJ346998.1

Score	Expect	Identities	Gaps	Strand
893 bits(483)	0.0	600/658(91%)	2/658(0%)	Plus/Plus

Conclusion: The PCR products revealed weak amplification after each PCR analysis. Besides the extracted DNA may contain a few PCR inhibitors (polymeric tannins). In most of the case, the PCR reactions were performed twice successively with 1 µL DNA template (+ amplicons) of the previous PCR reaction mix. The sequences obtained for ITS-2 were dirty possibly suggesting a mixture of related species. No amplicon was observed for trnH-psbA with the selected primers. The rbcL sequence displayed homologies around 90% with sequences of *Pausinystalia* species published in GenBank. However, the phytochemical analysis confirmed the presence of species specific chemical markers, such as yohimbine, in the crude extract.