

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Molecular identification of *Lodoicea maldivica* (coco de mer) seeds

Chinese Medicine 2011, **6**:34 doi:10.1186/1749-8546-6-34

Chun-yin Mak (cymak@govtlab.gov.hk)
Chuen-shing Mok (csmok@govtlab.gov.hk)

ISSN 1749-8546

Article type Research

Submission date 17 June 2011

Acceptance date 30 September 2011

Publication date 30 September 2011

Article URL <http://www.cmjournal.org/content/6/1/34>

This peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in *Chinese Medicine* are listed in PubMed and archived at PubMed Central.

For information about publishing your research in *Chinese Medicine* or any BioMed Central journal, go to

<http://www.cmjournal.org/authors/instructions/>

For information about other BioMed Central publications go to

<http://www.biomedcentral.com/>

Molecular identification of *Lodoicea maldivica* (coco de mer) seeds

Chun-yin Mak ^{*}, Chuen-shing Mok

Government Laboratory Hong Kong, 7/F Ho Man Tin Government Offices, 88 Chung Hau Street,
Kowloon, Hong Kong SAR, China

* Corresponding author:

Chun-yin Mak

Government Laboratory Hong Kong

7/F Ho Man Tin Government Offices

88 Chung Hau Street

Kowloon, Hong Kong SAR

China

Email addresses:

CYM: cymak@govtlab.gov.hk

CSM: csmok@govtlab.gov.hk

Abstract

Background The edible endosperm of *Lodoicea maldivica* with the common name of *coco de mer* is used in Chinese medicine for treating cough. Native to Seychelles, *Lodoicea maldivica* seeds have commanded high prices for centuries due to its scarcity. This study aims to develop a molecular identification method for the authentication of *Lodoicea maldivica* seeds. **Methods** DNA was extracted from the sample. Two polymerase chain reaction (PCR) systems were developed to amplify a region of the chloroplast DNA and the nuclear phosphoribulokinase (PRK) region specific to *Lodoicea maldivica* respectively. DNA sequence of a sample was determined and compared with that of the *Lodoicea maldivica* reference material. **Results** The PRK gene of *Lodoicea maldivica* was successfully amplified and sequenced for identification. **Conclusion** A new molecular method for the identification of *Lodoicea maldivica* seeds in fresh, frozen or dried forms was developed.

Background

Lodoicea maldivica (*coco de mer*), the sole member of the genus *Lodoicea*, is a fan-leaved palm native to Seychelles, bearing the largest and heaviest seed in the plant kingdom (weighing up to 30kg). The seed is enclosed in a hard shell resembling a pair of coconuts joined in the middle. The seed is sometimes also referred to as the sea coconut, bum seed, double coconut, coco fesse or Seychelles nut. *Lodoicea maldivica* palms take between 25-50 years to reach maturity and bear fruit. The fruit may take one to three years to germinate [1]. Only two natural populations of *Lodoicea maldivica* remain [1]. Individual plants are also cultivated in various botanical gardens around the world. The harvest of seeds has virtually stopped all natural regeneration of the plant. The populations are also threatened by fire and encroachment by invasive plants. In March 2010, *Lodoicea maldivica* was added to the appendices of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) in its 15th meeting of the Conference of the Parties [2]. Trade of *Lodoicea maldivica* is also governed by the Protection of Endangered Species of Animals and Plants Ordinance in Hong Kong (Cap. 586). According to the ordinance, unless exempted, the import, introduction from the sea, export, re-export or possession of *Lodoicea maldivica* requires a licence issued by the Hong Kong government. The edible endosperm (*ie* fruit) of the plant has long been used in Chinese medicine for treating cough and is used mainly in soup in Hong Kong, China [3]. As fake products are often found in the markets, an identification method is needed.

This study aims to develop a new method for the identification of *Lodoicea maldivica* with DNA sequencing analysis. DNA was extracted from the edible endosperm

sample. A polymerase chain reaction (PCR) system was designed to amplify a region of the chloroplast DNA to validate the quality of the extracted DNA [4-5]. A number of genes or spacer regions may be considered to identify unknown samples, for example, the internal transcribed spacer 1 [6], internal transcribed spacer 2 [6], nuclear ribosomal RNA genes 18S unit [6], large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase (*rbcL*) [7], intergenic region between the β subunit of ATP synthase (*atpB*) and *rbcL* [8] and the chloroplast trnL-trnF intergenic spacer region [9]. In this study, a PCR system targeting the nuclear phosphoribulokinase (PRK) region was designed. This region has been employed in many palm phylogenetics and shown ample characters to resolve relations at species level [10-13]. After DNA sequence analysis, the identity of the unknown was evaluated by pair-wise matching of its sequence with that of the certified *Lodoicea maldivica* reference species.

Methods

Materials

A *Lodoicea maldivica* seed certified by the Ministry of Environment, Natural Resources and Transport of the Republic of Seychelles was used as reference. The official certificate supports the authenticity of the species. The fresh reference seed is, thick, relatively hard and brown in colour with white flesh inside and about 10 kg in weight. The seed was assigned a voucher number of TD/TD/CRM/10001. The seed was kept frozen before analysis. Seven sliced commercial products with description of sea coconut or *coco de mer* (namely Sample A to Sample G) and one coconut sample (sample H) were purchased from the local market. Sample details are listed in Table 1

For the samples in the form of pre-packaged product mixed with other materials, the

suspected coconut portion was isolated from the background matrix before homogenisation. For fresh/frozen samples, the sample was homogenised by a blender, and about 1-5g of the sample was transferred into a flat bottom glass container. The sample was freeze-dried. The dried sample was ground into fine powder in liquid nitrogen. For dried samples, the sample was simply ground into fine powder in liquid nitrogen.

DNA extraction

DNA extraction was performed with the CTAB method [14]. A total of 1.0mL of pre-warmed (65°C) CTAB extraction buffer (2% w/v hexadecyl trimethylammonium bromide (CTAB), 0.1M Tris, 1.4M NaCl, 20mM EDTA) was added to 100mg of the test sample in a 2.0mL microcentrifuge tube. Alpha-amylase solution (10µL-20µL), 20µL of proteinase K solution (20mg/mL), 100µL of 10% w/v polyvinylpyrrolidone (PVP) and 5µL of RNase A (100mg/mL) were added. The mixture was incubated at 65°C overnight with constant agitation. After centrifugation (Eppendorf, Germany) for ten minutes at 16,000× g at room temperature, the supernatant was transferred to a new 2.0mL tube and an equal amount of chloroform was added. The sample was shaken vigorously for 30 seconds before centrifugation (Eppendorf, Germany) for ten minutes at 16,000× g at room temperature. The upper layer was transferred to a new 2.0mL tube and extracted with an equal amount of chloroform. The sample was centrifuged 16,000× g (Eppendorf, Germany) for five minutes. The upper layer was transferred to a new 2.0mL tube and two volumes of CTAB precipitation solution (0.5% w/v CTAB, 0.04M NaCl) were added. The sample was incubated at room temperature for 60 minutes without agitation. The supernatant was discarded after the sample was centrifuged (Eppendorf, Germany) for ten minutes

at 16,000× g. The precipitate was dissolved in 350µL of 1.2M NaCl solution and was centrifuged (Eppendorf, Germany) for ten minutes at 16,000× g. The upper layer was transferred to a new 1.5mL tube. Two volumes of cold absolute ethanol were added to precipitate the DNA. The sample was kept at -20°C for at least 20 minutes. The supernatant was discarded after centrifugation (Eppendorf, Germany) for ten minutes at 16,000× g at 4°C. A total of 500µL of 70% cold ethanol solution was added and the sample was centrifuged (Eppendorf, Germany) for ten minutes at 16,000× g at 4°C. The DNA pellet was dried in a 37°C oven for several minutes and was then re-dissolved in 100µL of DNase-free water.

PCR for chloroplast DNA

The chloroplast DNA region (partial) was amplified with 50ng DNA template, 1.0µL of 5mM dNTPs, 0.5µL of 25mM MgCl₂, 2.5µL of 10× PCR buffer plus MgCl₂ (Applied Biosystems, USA), 0.5µL of 10µM specific primers (Table 2) and 1U of AmpliTaq Gold polymerase (Applied Biosystems, USA). Water was added to make up to 25µL. The DNA template was denatured at 95°C for five minutes and then 35 cycles of 30 seconds at 94°C, 30 seconds at 55°C, 45 seconds at 72°C and final extension at 72°C for five minutes. The amplification products were electrophoresed on a 1% agarose gel, stained with ethidium bromide and observed under UV illumination.

PCR for PRK gene

The PRK gene was amplified with 50ng DNA template, 1.25µL of 5mM dNTPs, 2.0µL of 25mM MgCl₂, 2.5µL of 10× PCR buffer without MgCl₂ (Applied Biosystems, USA), 1.5µL of 10µM specific primers (Table 2), 1.25µL of 40 ng/µL

bovine serum albumin (BSA), 0.5 μ L of dimethyl sulfoxide (DMSO) and 1U of AmpliTaq Gold polymerase (Applied Biosystems, USA). Water was added to make up the final volume to 25 μ L. The DNA template was denatured at 95°C for four minutes, and then 35 cycles of one minute at 94°C, 30 seconds at 54°C, one minute at 72°C and final extension at 72°C for seven minutes. The amplification products were electrophoresed on a 1 % agarose gel, stained with ethidium bromide and observed under UV illumination.

DNA sequencing

The PCR products of PRK gene were purified with DNA Clean & ConcentratorTM -25 (Zymo Research, USA) according to the manufacturer's instructions. The BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) was used for the cycle sequencing reaction. The PCR product was first denatured at 96°C for one minute and then 25 cycles of thermal cycling was performed as follows: 96°C for ten seconds, 50°C for five seconds, 60°C for four minutes. The BigDye® XTerminatorTM Purification Kit (Applied Biosystems, USA) was used to purify the extension products according to the manufacturer's instructions and the products were then run on the ABI PRISM® 3130XL Genetic Analyzer (Applied Biosystems, USA).

Sequence Analysis

We aligned the sequence of the unknown sample in pair with that of the reference species using the bl2seq tool of the National Center for Biotechnology Information (NCBI) website [15]. Moreover, we performed multiple sequence alignment using the CLUSTAL2 tool of the European Bioinformatics Institute website [16].

Results and Discussion

DNA extraction

A study [17] reported that DNA from the solid endosperm of coconut was not a good starting material for molecular biology work because the matrix may hinder enzyme activity and because of the high lipid and polysaccharide content. The study suggested that DNA from young leaves of the plant should be used as the source of genomic DNA. Nevertheless, only the edible endosperm of *Lodoicea maldivica* is used in Chinese medicine and the commercial products thereof are made up of various forms of its edible endosperm. In order to develop a method for testing the authenticity of the claimed *Lodoicea maldivica* products, we think that PCR-ready DNA must be extracted from the endosperm of *Lodoicea maldivica*.

The chemical PVP is an important reagent in DNA extraction. Without the addition of PVP, the sample solution for the extracted DNA contained high levels of polysaccharide, as shown in the UV spectra. The contamination was greatly reduced by the addition of PVP during the initial incubation at 65°C in the extraction step. While the amount of DNA extracted from all samples in this study was very low (<10ng/μL), the quality of the extracted genomic DNA was sufficient for the PCR and sequencing.

PCR-chloroplast DNA

Plant 1 and Plant 2 primers are targeted for the amplification of a region of plant chloroplast DNA and were used to verify the quality of the template DNA [2]. A single PCR product of about 550bp was successfully amplified from all samples. The

positive PCR results indicated that DNA of sufficient quantity and quality were successfully extracted and that the PCR was not inhibited by other components of the analytical sample.

PCR PRK gene

The nuclear region of the PRK gene was amplified with the second PCR system. This region has been used successfully to resolve palm phylogenetic relations at the species level [10-13]. Positive results (*ie* a fluorescent DNA band of about 750bp) from this PCR system would indicate that the sample contained the PRK gene whereas negative results would indicate that the sample was not from *Lodoicea maldivica*.

Single-band PCR products of about 750bp were successfully amplified from *Lodoicea maldivica* reference material and two commercial products namely Sample A and Sample H. Other samples yielded either negative results or multiple PCR products in the amplification of PRK gene. *Lodoicea maldivica* yielded only one single PCR product in the analysis of PRK gene. The sample should be considered as not of *Lodoicea maldivica* in origin when no or more than one band was observed after agarose gel electrophoresis [4]. The identities of these unknown samples were found by using one of the markers [6-9] as mentioned in the background section.

Sequence analysis

All the amplified PRK genes were successfully sequenced. The nucleotide sequence of the *Lodoicea maldivica* reference material was submitted to the GenBank [GenBank: JF820816]. We compared the nucleotide sequence in pair between the reference material and each of the unknown samples using the bl2seq tool [15]. For

those sequences not matching the reference material, further search in the GenBank was performed. The PRK gene sequence of sample A was identical to that of the reference material (Figure 1). While the PRK gene sequence of sample H only matched that of the reference material with 80% homology (Figure 2), it matched the PRK sequence of species *Cocos nucifera* with 100% homology in GenBank database [GenBank: HQ265608] (Figure 3). Currently, there is only one single entry of the *Lodoicea maldivica* PRK gene sequence in NCBI GenBank [GenBank: AF453357]. The sequence was obtained from a *Lodoicea maldivica* voucher (voucher number: C.E. Lewis 98-020/BH) and reported by Lewis *et al.* in 2002 [12]. However, the exact source of the voucher was not reported. The BLAST [15] result of the *Lodoicea maldivica* reference material sequence obtained in this study showed a 99% homology to that of the [GenBank: AF453357] (Figure 1). The *Lodoicea maldivica* reference species used in this study was obtained from the country of origin of the species with official certificate. The present study provides an alternative source of nucleotide sequence data for GenBank for the identification of *Lodoicea maldivica*. The applicability and accuracy of the method have been demonstrated by the consistent results obtained from replicated analysis of the samples.

Conclusion

A new molecular method for the identification of *Lodoicea maldivica* seeds in fresh, frozen or dried forms was developed.

Abbreviations

CITES: Convention on International Trade in Endangered Species of Wild Fauna and Flora; CTAB: hexadecyl trimethylammonium bromide; DMSO: Dimethyl sulfoxide; EDTA: Ethylenediaminetetraacetic acid; NCBI: National Center for Biotechnology Information; PCR: polymerase chain reaction; PRK: phosphoribulokinase; PVP: polyvinylpyrrolidone

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CYM conceived and designed the study, collected the samples, performed the laboratory work, analyzed and interpreted the data, and drafted the manuscript. CSM supervised the project and finalized the manuscript. Both authors read and approved the final version of the manuscript.

Acknowledgements

The authors would like to thank Dr CM Lau, Government Chemist of the Government Laboratory of the HKSAR, China, for his support and encouragement during the study. We would also like to thank Dr Della WM Sin and Ms Silvia YK Tam for their discussion and comments.

Disclaimers

This article does not necessarily reflect the views of the Government of the HKSAR. The mentions of trade names or commercial products do not constitute any endorsement or recommendations.

References

1. **Palm and Cycad Societies of Australia**
[<http://www.pacsoa.org.au/palms/Lodoicea/index.html>]
2. **The Convention on International Trade in Endangered Species of Wild Fauna and Flora** [<http://www.cites.org/>]
3. Hu SY: **Food Plants of China**. The Chinese University Press; 2005.
4. Hirao T, Imai S, Sawada H, Shiomi N, Hachimura S, Kato H: **PCR Method for Detecting Trace Amounts of Buckwheat (*Fagopyrum* spp.) in Food**. *Biosci. Biotechnol. Biochem* 2005, **69**:724-731.
5. Taberlet P, Gielly L, Pautou G, Bouvet J: **Universal primers for amplification of three non-coding regions of chloroplast DNA**. *Plant Mol. Biol.* 1991, **17**:1105-1109.
6. Hills DM, Dixon MT: **Ribosomal DNA: molecular evolution and phylogenetic inference**. *Q. Rev. Biol.* 1991, **66**:411-453.
7. Soltis PS, Soltis DE, Smiley CJ: **An *rbcL* sequence from a Miocene *Taxodium* (bald cypress)**. *Proc Natl Acad Sci USA* 1992, **89**:449-451.
8. Savolainen V, Cuénoud P, Spichiger R, Martinez MDP, Crèvecoeur M, Manen J-F: **The use of herbarium specimens in DNA phylogenetics: evaluation and improvement**. *Plant Syst. Evol.* 1995, **197**: 87-98.
9. Baker WJ, Asmussen CB, Barrow SC, Dransfield J, Hedderson TA: **A phylogenetic study of the palm family (Palmae) based on chloroplast DNA sequences from the *trnL* - *trnF* region**. *Plant Syst. Evol.* 1999, **219**: 111-126.
10. Loo, AHB, Dransfield J, Chase MW, Baker WJ: **Low-copy nuclear DNA, phylogeny and the evolution of dichogamy in the betel nut palms and their relatives (Arecinae; Arecaceae)**. *Mol Phylogenet Evol* 2006,

39:598-618.

11. Thomas MM, Garwood NC, Baker WJ, Henderson SA, Russell SJ, Hodel DR, Bateman RM: **Molecular phylogeny of the palm genus *Chamaedorea*, based on the low-copy nuclear genes PRK and RPB2.** *Mol Phylogenet Evol* 2006, **38**:398-415.
12. Lewis CE, Doyle JJ: **A phylogenetic analysis of tribe Areceae (Arecaceae) using two low-copy nuclear genes.** *Plant Syst. Evol.* 2002, **236**: 1–17.
13. Cuenca A, Asmussen-Lange CB, Borchsenius F: **A dated phylogeny of the palm tribe Chamaedoreae supports Eocene dispersal between Africa, North and South America.** *Mol Phylogenet Evol* 2008, **46**:760-775.
14. **Foodstuffs - Methods of analysis for the detection of genetically modified organisms and derived products – Quantitative nucleic acid based methods.** ISO 21571:2005
15. **National Center for Biotechnology Information**
[<http://blast.ncbi.nlm.nih.gov/Blast.cgi>]
16. **European Bioinformatics Institute** [<http://www.ebi.ac.uk/Tools/>]
17. Angeles JGC, Laurena AC, Tecson-Mendoza EM: **Extraction of Genomic DNA From the Lipid-, Polysaccharide-, and Polyphenol- Rich Coconut (*Cocos nucifera* L.).** *Plant Mol. Biol. Rep.* 2005, **23**:297a-297i.

Table 1 Sample information and the results of their identities found

Sample code	Sample description (form)	Target region	Percentage identity	Identity found
Sample A	African seacoconut (dried)	PRK	99% [GenBank: (AF453357)]	<i>Lodoicea maldivica</i>
Sample H	Coconut (fresh)	PRK	100% [GenBank: (HQ265608)]	<i>Cocos nucifera</i>

Table 2 Nucleotide sequences of specific primers used for the amplification of the chloroplast DNA and PRK gene

Name	Oligonucleotide DNA sequence (5'-3')	Amplicon size	Specificity
Plant 1 (forward)	CGA AAT CGG TAG ACG CTA CG	550 bp	Chloroplast DNA
Plant 2 (reverse)	GGG GAT AGA GGG ACT TGA AC		
prk717f (forward)	GTG ATA TGG AAG AAC GTG G	750 bp	PRK gene
prk969r (reverse)	ATT CCA GGG TAT GAG CAG C		

Figure legends

Figure 1 - Sequences alignment for *Lodoicea maldivica* reference material (RM)

[GenBank: JF820816], Sample A and [GenBank:AF453357]

Figure 2 - Sequences alignment for *Lodoicea maldivica* reference material (RM)

[GenBank: JF820816] and Sample H

Figure 3 - Sequences alignment for Sample H and *Cocos nucifera* [GenBank:

HQ265608]

```

AF453357      GTCTTGACAGCATCAAAGCTAGCATTGAAGCTCGAAAGCCTGATTTTGAT 50
Sample_A      -----GCCTGATTTTGAT 13
Lodoicea_maldivica_RM -----

AF453357      GCTTATGTTGGTATGTCACCTTGCCCCAGCACCTAGTTGTTAGTACTTGA 100
Sample_A      GCTTATGTTGGTATGTCACCTTGCCCCAGCACCTAGTTGTTAGTACTTGA 63
Lodoicea_maldivica_RM -----

AF453357      AGAATTGGCAGTGCACCTTATTTGTTGGTGTAAATATAGAGCATTGCCATTA 150
Sample_A      AGAATTGGCAGTGCACCTTATTTGTTGGTGTAAATATAGAGCATTGCCATTA 113
Lodoicea_maldivica_RM -----GGCAGTGCACCTTATTTGTTGGTGTAAATATAGAGCATTGCCATTA 44
                        *****

AF453357      GTGGTGATTGATATATATAGTAATCTACTTCTATATATAGCAAATGCTTC 200
Sample_A      GTGGTGATTGATATATATAGTAATCTACTTCTATATATAGCAAATGCTTC 163
Lodoicea_maldivica_RM GTGGTGATTGATATATATAGTAATCTACTTCTATATATAGCAAATGCTTC 94
                        *****

AF453357      ATAAATAAACTAACCTAATAAAACAAAGGATCTGATCGATAA--GTCAAAT 248
Sample_A      ATAAATAAACTAACCTAATAAAACAAAGGATCTGATCGATAAAGGTCAAAT 213
Lodoicea_maldivica_RM ATAAATAAACTAACCTAATAAAACAAAGGATCTGATCGATAAAGGTCAAAT 144
                        *****

AF453357      AATGCACAAAACATACTGGACTCAAAATATACTCTTACTTCATGTGCAAT 298
Sample_A      AATGCACAAAACATACTGGACTCAAAATATACTCTTACTTCATGTGCAAT 263
Lodoicea_maldivica_RM AATGCACAAAACATACTGGACTCAAAATATACTCTTACTTCATGTGCAAT 194
                        *****

AF453357      TTTTGTGACACCTGTTTTTTTCAAAAAAAGAAAGGAGAGGGAGGGAGGG 348
Sample_A      TTTTGTGACACCTGTTTTTTTCAAAAAAAGAAAGGAGAGGGAGGGAGGG 313
Lodoicea_maldivica_RM TTTTGTGACACCTGTTTTTTTCAAAAAAAGAAAGGAGAGGGAGGGAGGG 244
                        *****

AF453357      TGAGAGTCACCTGATTCATTGAAATAATGAGACAAGTAGATATGTTGCAT 398
Sample_A      TGAGAGTCACCTGATTCATTGAAATAATGAGACAAGTAGATATGTTGCAT 363
Lodoicea_maldivica_RM TGAGAGTCACCTGATTCATTGAAATAATGAGACAAGTAGTTATGTTGCAT 294
                        *****

AF453357      CTCATGTTCTTCAGATATCTGAATATCCATTCTTTCTTTAGCACTCTTGG 448
Sample_A      CTCATGTTCTTCAGATATCTGAATATCCATTCTTTCTTTAGCACTCTTGG 413
Lodoicea_maldivica_RM CTCATGTTCTTCAGATATCTGAATATCCATTCTTTCTTTAGCACTCTTGG 344
                        *****

AF453357      ATTACTGACAGCTTGTAGGTTGCTGCCCTTGGCAGAGCCGCAGAAGCAAT 498
Sample_A      ATTACTGACAGCTTGTAGGTTGCTGCCCTTGGCAGAGCCGCAGAAGCAAT 463
Lodoicea_maldivica_RM ATTACTGACAGCTTGTAGGTTGCTGCCCTTGGCAGAGCCGCAGAAGCAAT 394
                        *****

AF453357      ATGCTGATGTCGTAATCGAAGTTTACCGACACAATTAATTCCTGGTGAC 548
Sample_A      ATGCTGATGTCGTAATCGAAGTTTACCGACACAATTAATTCCTGGTGAC 513
Lodoicea_maldivica_RM ATGCTGATGTCGTAATCGAAGTTTACCGACACAATTAATTCCTGGTGAC 444
                        *****

AF453357      AAAGAAAGGAAGGTGCTGAGAGTTTCGATTGGTGTGCAAGAAGGGGTGAA 598
Sample_A      AAAGAAAGGAAGGTGCTGAGAGTTTCGA----- 540
Lodoicea_maldivica_RM AAAGAAAGGAAG----- 456
                        *****

```

Figure 1

Lodoicea_maldivica_RM
Sample_H

CCCGGAAGCTCGACTTTGATGCTTATGTTGGTATGTCACCTTGCCAAAGC 50

Lodoicea_maldivica_RM
Sample_H
-----GGCAGTGCCTTATTTGTTGGTG 23
ACCTAGTTTGTGGTACTTGAAGAATTGGCAGAGCTATAATTTGTTGGTA 100
***** ** * *****

Lodoicea_maldivica_RM
Sample_H
TAATATAGAGCATTGCCATTAGTGGTGATTGATATATATAGTAATCTACT 73
TAATATAGAGCAGTATCATTAGTGTGATCGCGAATATAGTAATATACT 150
***** * ***** ** * *****

Lodoicea_maldivica_RM
Sample_H
TCTATATATAGCAAATGCTTCATAAAATAAACTAACCTAATAAACAAAGGA 123
TCAGTATGTAGCAAATGCTTCATAAAAAAACCACTAGCTAGCAAAGAA 200
** ** ***** ** * ***** *

Lodoicea_maldivica_RM
Sample_H
TCTGATCGATAAAGGTCAAATAATGCACAAAACATACTGGACTC-----A 168
TCTGATCGATAA--GTCAAATAATGCACAAGACCTACTGGACTGGACTAA 248
***** ***** ** ***** *

Lodoicea_maldivica_RM
Sample_H
AAATATACTCTTACTTCATGTGCAATT--TTGTTGACACCTGTTTTTTTC 217
AAATATAATGTTACTTCATGTGCAAACATTTTTTGGTACC--TTTTCTA 296
***** * ***** ** ** ** ***** *

Lodoicea_maldivica_RM
Sample_H
AAAAAAGAAAGGAGAGGAGGGGAGGGTGAGAGTCACCTGATTCATTGAA 267
AAAGAAAGAGAGGGTGAGGGGTAGGCGAACCCACCCGTTTCATTGAG 346
*** ***** * * * ***** ** * *****

Lodoicea_maldivica_RM
Sample_H
ATAATGAGACAAGTAGTTATGTTGCATCTCATGTTCTTCAGATATCTGAA 317
ATAATGAGACAATTAGATATGCTGAGTCTTAT-----ACATATCAGAA 389
***** ** ** * ** ** * ***** **

Lodoicea_maldivica_RM
Sample_H
TATCCATTCTTTCTTTAGCACTCTTGGATTACTGACAGCTTGTAGGTTGC 367
TATCCATTCATTTTCTAGCACGTTTGGATTATTGATGGCTTATAGATTGC 439
***** ** * ***** ***** ** ***** **

Lodoicea_maldivica_RM
Sample_H
TGCCCTTGGCAGAGCCGCAGAAGCAATATGCTGATGTCGTAATCGAAGTT 417
-GACCCTGGCAGACCCACAGAAGCAATATGCTGATGTCGTAATGAAGTT 488
* ** ***** ** *****

Lodoicea_maldivica_RM
Sample_H
TTACCGACACAATTAATTCCTGGTGACAAAGAAAGGAAG----- 456
TTACCGACACAATTAATTCCTGATGACAATGAAAGGAAGGTGCTGAGAGT 538
***** *****

Figure 2

HQ265608 TTTGTGATATGGGAGAACGTGGGCACAGTCTCGAAAAGCATCAAAGCTAGCATCGAAGCCC 60
Sample_H -----CCC 3

HQ265608 GGAAGCTCGACTTTGATGCTTATGTTGGTATGTCACCTTGCCAAAGCACCTAGTTTGTG 120
Sample_H GGAAGCTCGACTTTGATGCTTATGTTGGTATGTCACCTTGCCAAAGCACCTAGTTTGTG 63

HQ265608 GTACTTGAAGAATTGGCAGAGCTATAATTTGTTGGTATAATATAGAGCAGTATCATTAGT 180
Sample_H GTACTTGAAGAATTGGCAGAGCTATAATTTGTTGGTATAATATAGAGCAGTATCATTAGT 123

HQ265608 GTTGATCGCCGAATATAGTAATATACTTCAGTATGTAGCAAATGCTTCATAAAAAAACCA 240
Sample_H GTTGATCGCCGAATATAGTAATATACTTCAGTATGTAGCAAATGCTTCATAAAAAAACCA 183

HQ265608 ACCTAGCTAGCAAAGAATCTGATCGATAAGTCAAATAATGCACAAGACCTACTGGACTGG 300
Sample_H ACCTAGCTAGCAAAGAATCTGATCGATAAGTCAAATAATGCACAAGACCTACTGGACTGG 243

HQ265608 ACTAAAAATATAATGTTACTTCATGTGCAAACATTTTTTGGTACCTTTTTCTAAAAGAAA 360
Sample_H ACTAAAAATATAATGTTACTTCATGTGCAAACATTTTTTGGTACCTTTTTCTAAAAGAAA 303

HQ265608 GAAGAGGGTGAGGGGTAGGCGAACCCACCCGGTTCATTGAGATAATGAGACAATTAGA 420
Sample_H GAAGAGGGTGAGGGGTAGGCGAACCCACCCGGTTCATTGAGATAATGAGACAATTAGA 363

HQ265608 TATGCTGAGTCTTATACATATCAGAATATCCATTCATTTTCTAGCACGTTTGGATTATTG 480
Sample_H TATGCTGAGTCTTATACATATCAGAATATCCATTCATTTTCTAGCACGTTTGGATTATTG 423

HQ265608 ATGGCTTATAGATTGCGACCCTGGCAGACCCACAGAAGCAATATGCTGATGTCGTAATTG 540
Sample_H ATGGCTTATAGATTGCGACCCTGGCAGACCCACAGAAGCAATATGCTGATGTCGTAATTG 483

HQ265608 AAGTTTACCGACACAATTAATTCCTGATGACAATGAAAGGAAGGTGCTGAGAGTTCGAT 600
Sample_H AAGTTTACCGACACAATTAATTCCTGATGACAATGAAAGGAAGGTGCTGAGAGTTCGAT 543

HQ265608 TGGTGATGAAAGAAGGGGTGAAGTACTTCAATCCGTTTACCTCTTCGACGAAGGCTCCA 660
Sample_H TGGTGATGAAAGAAGGGGTGAAGTACTTCAATCCGTTTACCTCTTCGACGAAGGCTCCA 603

HQ265608 CTGTTTCATGGATACCCTGCGGNAGGAAACTAAGCTGCTCATCCCC 706
Sample_H CTGTTTCATGGATACCCTGCGGTAGGAAACTAAGCTGCTCAT---- 645

Figure 3