American Journal of Plant Sciences, 2015, 6, 3045-3074
Published Online December 2015 in SciRes. http://www.scirp.org/journal/ajps
http://dx.doi.org/10.4236/ajps.2015.619299

Exploitation of Concatenated Olive Plastome DNA Markers for Reliable Varietal Identification for On-Farm Genetic Resource Conservation

Muhammad Noman[1,2*], Wajya Ajmal[1*], Muhammad Ramzan Khan[1,2#], Armghan Shahzad[1,2], Ghulam Muhammad Ali[1,2#]

[1]National Institute for Genomics and Advanced Biotechnology, National Agricultural Research Centre (NARC), Islamabad, Pakistan
[2]PARC Institute of Advanced Studies in Agriculture, National Agricultural Research Centre (NARC), Islamabad, Pakistan
Email: [#]drmrkhan_nigab@yahoo.com, [#]drgmali@yahoo.com

Received 20 September 2015; accepted 5 December 2015; published 8 December 2015

Abstract

Rapid and reliable identification of olive plants using DNA markers has been attempted in the past but the selection of polymorphic regions for discrimination at varietal level remained obscure. Recent sequencing of plastid genome of the olive flaunts high resolution Cp markers for olive DNA fingerprinting. Using this information, we designed a combination of chloroplast markers to amplify genes recruited in photosynthesis, ribosomal and NADH energy metabolism for varietal identification of olive plants. Concatenated DNA sequences of more than 100 unknown and 10 reference plants samples were analyzed using various bioinformatics and phylogenetic tools. Conserved blocks of nucleotide sequences were detected in multiple alignments. Phylogenetic reconstruction differentiated the unknown plants into various clusters with known varieties. Further narrowing down of the samples through UPGMA tree clearly separated the plants into Arbosana, Frantoio and Koroneiki as the major varieties. Multiple alignments of these clusters revealed important variety specific SNPs including G and T nucleotides at specific positions. Sequence identifying at intra cultivar level was more than 98.79% while it dropped to 97%, and even to 96% at inter varietal level. Furthermore, a neighbor net network analysis separated these three clusters, thus validating the results of UPGMA tree. Over all, out of 100 plants samples, 49 plants were identified that fall into 10 varieties including Arbosana, Carolea, Chetoui, Coratina, Domat, Frantoio, Gemlik, Koroneiki,

[*]Contributed equally.
[#]Corresponding authors.

How to cite this paper: Noman, M., Ajmal, W., Khan, M.R., Shahzad, A. and Ali, G.M. (2015) Exploitation of Concatenated Olive Plastome DNA Markers for Reliable Varietal Identification for On-Farm Genetic Resource Conservation. *American Journal of Plant Sciences*, **6**, 3045-3074. http://dx.doi.org/10.4236/ajps.2015.619299

Leccino and Moraiolo. The maximum number of known plants belongs to Frantoio and Gemlik (8 each). The least number of samples was identified from Carolea, Domat and Moraiolo with 2 samples each. However, 51 plants could not be identified, as plants were not clustered with any of reference control. Our results have implications in on-farm conservation of olive germplasm and provision of genuine material for multiplication of authentic varieties. This strategy can be extended to varietal identification of other plant species.

Keywords

DNA Fingerprinting, Olive, Marker, Chloroplast, Genome, Identification, Phylogenetic

1. Introduction

One of the characteristic fruit trees of the Mediterranean area is the evergreen and long-lived olive (*Olea europaea* L.). It is diploid with 46 chromosomes from the Oleaceae family [1]-[3]. Olive can be as older as 500 years but over 2000 years older trees are also in record. It is a medium sized tree with grey-green leaves arranged opposite to one another. The olive comes from the genus Olea that has 3 subgenera Paniculatea, Olea and Tetrapilus [4]. *Olea europaea* L. is the only single species that bears edible fruit [5] [6]. The origin of olive is still unclear, but the main hypothesis suggests that it originated from the Eastern shores of the Mediterranean [7].

The fruit and oil of olive are of prime importance worldwide. Although 90% of world olive production is used for oil extraction [8], the consumption of table olives is also growing globally. Today, the olive tree is grown commercially within latitudes 30˚ and 45˚ in both the Northern and Southern hemispheres, where climatic conditions are similar to the Mediterranean basin, with mild winters and warm, dry summers [9] [10]. Pakistan lies in the belt between 30˚ - 45˚ North and South of the equator, hence it is a potential area for olive cultivation. The suitable areas include Pothwar, Khyber Pakhtunkhwa, Swat, Dir, Malakand, Loralai, Khuzdar and Quetta districts etc. Edible oil is the biggest food import item of Pakistan. Pakistan imports olive oil and fruit every year and huge funds are consumed on their cost. Self-sufficiency in edible can be attained by cultivating olive orchards in the marginal lands (more than 3 million acres; 30% of total land) of Pakistan. Under different projects, the total olive tree cover is more than 800 ha comprising of 106,048 trees. These plants are at fruiting stage and some of these plants are giving very good yield. But the biggest problem which is restricting their large scale propagation is that these olive varieties/plants are unidentified and there is no record which variety/cultivar they are. Therefore, oil extracted from these plants is mixed and does not get its premium price in the market. Unavailability of known high yielding and quality oil producing varieties/plants is the biggest hurdle for large scale propagation of olive in the poor lands of Pakistan. Furthermore, the unavailability of true to type olive nurseries is also impeding the olive propagation in the potential regions.

The olive's ancient origin, easy propagation and popularity have resulted in the presence of its numerous cultivars across the world. Several cultivars may have the same name (homonyms), or the same cultivar may be called by different names (synonyms) in different areas [11]-[13]. Many areas in botany depend on the efficiency to discriminate plant genotypes and calculate the amount of diversity and similarity in a group of genotypes. This has been done traditionally through morphological and biochemical markers and presently through DNA markers or DNA fingerprinting [14]. Molecular markers are preferred because they have several advantages over their alternatives. Like, they are co-dominantly inherited and highly polymorphic. They can be easily visualized and are spread over the whole genome evenly. They are stable, quick, inexpensive and simple to use. They require small amount of DNA and do not require any pre-info about the genome [15]. The olive gene pools have also been characterized utilizing the high resolving capacity of the molecular markers. Many researchers have traced the origin of olive germplasm using different molecular makers like RAPD [16].

An advanced genome screening technique is that of the plastome sequencing or screening the chloroplast DNA through specific markers. The chloroplasts are inherited maternally in the cultivated olives [17]. The plastidial variability is low in the cultivated olives in contrast to that detected at the subspecies level. The mitochondria and chloroplasts both pass through recurrent mutations but the level of mutations is low as demonstrated by [18]. Taking advantage of the highly conserved nature of cpDNA, universal primers for the cpDNA introns have been developed for numerous plant species [19]. Besnard and his colleagues detected 14 polymorphisms in the 3 chloroplast regions (trnT-L, trnQ-R and matK) in the *Olea europaea* complex [20].

For this study, the last approach of sequencing of the entire chloroplast genome of the *Olea europaea* subsp. *europaea* cv. *Frantoio* to identify the polymorphic regions was employed. The resulting availability of the entire plastome map allowed to evaluate the sequence arrangement of the plastid genome in *Olea europaea* and to identify new organellar polymorphisms that could discriminate between cultivated olive varieties [21]. In order to propagate only the better and high yielding cultivars, there is dire need to screen the cultivated olive plants in Pakistan to identify variety/cultivar. We can also graft our desired varieties onto the wild plants. This can enhance the olive fruit and oil yield in Pakistan. The work of this nature has not been done in Pakistan to date. Olive growers can name accurately some cultivars with distinguished phenotypic traits. But they confuse while differentiating the cultivars having similar morphological characters. Due to this problem, certified and good quality material for the establishment of new olive orchards is not available. Hence rapid and reliable identification of unknown olive plants growing at various olive farms through DNA marker is essential. Therefore, the objective of the present study was to screen unknown and known plants through specific cholorplast DNA markers for identification of polymorphic regions, identification of unknown cultivars of olive growing at different orchards in Pakistan using DNA markers, and to infer their evolutionary relationship through phylogenetic reconstruction. The results demonstrate that olive genome harbours some very advantageous polymorphic sites which can be employed for the reliable screening of unknown olive varieties through cultivar specific SNPs. The evolutionary relationship explored by phylogenetic investigation also helped in identifying the plants. Finally the neighbor-net network analysis validated the clustering of plants into specific variety.

2. Plant Materials and Methods

2.1. Selection of Materials and Sampling Plan

Information about the olive plants growing at different locations in Pakistan was obtained from National Director of the Olive Project, Pakistan Agricultural Research Council, Islamabad, Pakistan. Different areas of Khyber Pakhtunkhwa province were selected for plant sampling. Each plant was labeled using olive farm name, orchard number, row number and plant number. After plant labeling, fresh leaf tissue was harvested from the plants. The samples were stored at −80˚C until DNA extraction was performed.

2.2. Sequence Retrieval and Primer Designing

The chloroplast genomes of 8 olive cultivars including Frantoio were retrieved from NCBI database. The genome sequences were aligned and scanned using MacVector7.2 software [22] and polymorphic markers were selected. In this case, three chloroplast markers Oe-psbK-psbI-trnS-trnG-trnS-GCU-trnGUCC, Oe-rps8-rpl14-rps8-rpl14 exon, Oe-ndhF exon-rpl32-rpl32-trnL-1-trnL-UAG were selected. Three pairs of primers were designed for all the selected genes and regions.

2.3. DNA Extraction, PCR Amplification and Sequencing

A total of 110 plant leaf samples were used for DNA extraction using CTAB method [23]. For quality assessment DNA was run on 0.8% agarose gel. The diluted DNA samples were used as a template for PCR amplification with three primer pairs. The primer pairs used were named as CP3, CP4 and CP5. The sequence of the CP5 forward primer was 5’-CTGACAATTCATTTCTATTTCTAGA-3’ and reverse primer was 5’-CATTATTTATCTATAATTCGTTGGA-3’. Their position in cpDNA is 8986 to 9705 and they amplify a fragment of 720 bp length.

Each PCR reaction (50 µl) contained 10 ng DNA template, 10× reaction buffer, 5 µL $MgCl_2$, 1 µL dNTPs, 1 µL of each primer, and 0.5 µL of Taq DNA polymerase (Promega, Madison, WI, USA). The reaction mixtures were incubated in a thermocycler (Applied Biosystems Inc) for 5 min at 95˚C, followed by 36 cycles of 1 min at 94˚C (denaturing), 1 minute at the annealing temperature 58˚C, and 1 min at 68˚C (extension). PCR products were run on 1.2% agarose gel to view the amplification success. The PCR product was sent to Macrogen (Korea) for sequencing.

2.4. Sequence Analysis and Multiple Alignments

The sequence files obtained were edited and analyzed with MacVector7.2 program [22]. Blastn was done for target identification in NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The BioEdit software [24] was used to trim the sequences to remove the mismatched/flanking regions from both the ends. The ClustalW multiple alignment of the sequences was done using BioEdit and MEGA6 software [24] [25]. The mutations were detected, recorded and matched with previously available known data of different olive cultivars. Furthermore sequence identity at intra and inter varietal level was calculated through pairwise alignments. In this way, different olive cultivars were discriminated based on sequences similarities. A dataset was prepared that comprised 100 unknown and 10 known plants marker region sequences to be analyzed with bioinformatics software.

2.5. Phylogenetic Reconstruction

In order to infer the evolutionary relationship among different cultivars, phylogenetic reconstruction using UPGMA algorithm was done in MEGA6. The data generated was also helpful in cultivar identification.

It is well demonstrated that phylogenetic network could better reveal the evolutionary history including hybridization, recombination and homoplasmy etc. than a tree like structure. Therefore, a neighbor-net network reconstruction analysis was implemented in SplitTree4 package with default parameters using an uncorrected P distance method [26].

2.6. Unknown Plant Identification

The results from cultivar specific mutations *i.e.* SNPs, multiple alignments and phylogenetic reconstruction were combined and analyzed for plant identification. The identified plants were tabulated and shown graphically in results section.

3. Results

3.1. The Selected Marker Genes in Olive Plastome Are Polymorphic

Mariottiand his colleagues sequenced entire chloroplast genome of Frantoio cultivar and reported a number of polymorphic markers [21]. Using this information we set out to find the most variable regions with high resolving power that can be used to identify the olive plants at variety level. Scanning of the olive chloroplast genome revealed three polymorphic regions (**Supplementary Figure S1**). The region 1 coding for the photosystem thylakoid membrane (psb-A) and transfer RNA (trnL) gene is located in the start from 8986 bp to 9705 bp. This region spans a length of 720 bp. It is the most polymorphic region as it harbors six different types of mutations including two SNPs, two indels and two SSRs. The details about these regions are given in **Table 1**. Similarly, the region 2 is located between 83112 bp to 83852 bp with a stretch of 740 bp. This region was also quite polymorphic and encodes ribosomal protein S (rps). Region 3 is located in the extreme distal portion. This region could amplify a size of 1334 bp between 101263 bp to 102599 bp. Ribosomal protein S (rpsT) and NADH dehydrogenase (ndhF) are encoded by these markers genes. Based on this information, three primer pairs CP5, CP4 and CP3 were designed for the amplification of selected regions 1, 2, and 3, respectively using "primer tool" in MacVector 7.2 software (**Supplementary Table S1**).

Initially, PCR amplification followed by sequencing analysis for five known cultivars, Carolea, Gemlik, Domat, Leccino and Moraiolo grown at NARC revealed that CP5 gave the best amplification and sequencing results in comparison with CP3 and CP4 primers. There were fewer polymorphic sites detected in regions ampli-

Table 1. Mutations detected in the selected polymorphic region of olive plastome. Type and position of mutation are also mentioned.

Sr. No.	Polymorphism type	Motif	Position
1	SSR	T10-11	9072
2	SNP	C/T	9463
3	SNP/Indel	A/T/-	9535
4	Indel	TTAGATA/-	9536
5	Indel	A4(G)A5/-	9574
6	SSR	A11-14	9579

fied using CP3 and CP4 primers. Furthermore the sizes of their products were also longer in comparison with CP5 (Data not shown). On the other hand, CP5 revealed a number of polymorphic sites. Hence CP5 primer pair was selected for the amplification of olive samples. Moreover the product size with CP5 was smaller (less than 720 bp) that could be easily amplified which reduced the sequencing cost as well. At least three PCR products were sequenced for each sample. The sequences were edited using BioEdit program and trimmed in order to eliminate the errors induced by sequencing procedure and to get the reliable sequence for analysis.

To explore the variability in the upstream regions of chloroplast genes, five reference plants sampled from NARC were compared with Frantoio sequence of NCBI database. For this purpose a multiple alignment was generated in BioEdit program. The alignment in **Figure 1** shows that the selected region is quite polymorphic. In a short span of 600 bp, 14 mutations can be identified. These mutations included SNPs and deletion/insertions. There are two deletions located at 445 bp and 514 bp position, where A is deleted. The most frequent substitutions present are A and G nucleotides. There are specific SNPs in the NARC Carolea including A at position 46, 86, 294 and 296. Similarly another SNP of the nucleotide G is present only in NCBI Frantoio at 238th position. These mutations seem to be cultivars specific.

```
NARC_Carolea    AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Domat      AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Gemlik     AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Leccino    AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Moraiolo   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NCBI_Frantoio   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA

NARC_Carolea    TACCGAAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
NARC_Domat      TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
NARC_Gemlik     TACCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
NARC_Leccino    TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
NARC_Moraiolo   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
NCBI_Frantoio   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG

NARC_Carolea    AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
NARC_Domat      AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
NARC_Gemlik     AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
NARC_Leccino    AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
NARC_Moraiolo   AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
NCBI_Frantoio   AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT

NARC_Carolea    TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAAAAAGGG ATTCAAACCC TCGGTACGAA
NARC_Domat      TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
NARC_Gemlik     TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
NARC_Leccino    TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
NARC_Moraiolo   TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
NCBI_Frantoio   TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA

NARC_Carolea    TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Domat      TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Gemlik     TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Leccino    TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Moraiolo   TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NCBI_Frantoio   TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT

NARC_Carolea    TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
NARC_Domat      TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
NARC_Gemlik     TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGAT-TACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
NARC_Leccino    TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
NARC_Moraiolo   TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
NCBI_Frantoio   TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA

NARC_Carolea    GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAAACTT CTTG
NARC_Domat      GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTT
NARC_Gemlik     GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAA-TAAAG AAAAAAAAAA AAAAAAACTT CTTT
NARC_Leccino    GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACAT CTTT
NARC_Moraiolo   GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTT
NCBI_Frantoio   GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT
```

Figure 1. Multiple alignments of the marker region sequences of 5 olive varieties collected from NARC and one sequence of Frantoio retrieved from NCBI database, using BioEdit software. The shaded regions show the conserved sequences in the marker region of the chloroplast DNA of these different varieties. The regions that are not shaded exhibit the sites of mutations. These are SNPs and indels. SNPs are substitutions of single nucleotides. The gaps are the indels.

The above results allow us to infer that upstream region of the olive plastome is highly polymorphic with cultivar specific SNPs. Thus, this region *i.e.* CP5 primer specific can be used to identify plants at the variety level.

3.2. Phylogenetic Reconstruction Clustered the Unknown with Known Varieties of Olive

After sampling, the leaf material was immediately processed for DNA extraction using CTAB method [23]. A total of 110 samples were run on agarose gel for quantification. Chloroplast DNA was also present in this genomic DNA. These DNA samples were labelled and stored at −20˚C. As CP5 primer pair was found to be the most polymorphic that could amplify a very short region of 720 bp containing 6 different mutations; therefore this primer pair was used to amplify Oe-psbK-psbI and Oe-trnS-trnG-1-4 regions of the plastome DNA of olive. It was possible to amplify the entire plate of 96 samples in a single PCR reaction. The amplified products were resolved on agarose gel against 1 kb ladder (**Figure 2**). The quality and quantity of PCR product was good enough for sequencing.

Sequencing of all the 110 samples was carried out using the services of MACROGEN Korea. Targets samples were selected using BLAST search. The sequences were edited using BioEdit software [24]. The sequences were trimmed and aligned. This region contains all the SNPs, indels and SSRs showing polymorphism in different samples.

Based on sequence data, three types of *in silico* approaches were adopted to identify the unknown olive samples/sequences. Firstly, comparison of unknown sequences with known sequences through multiple alignments Secondly, identification of variety specific SNPs, indels and SSRs in unknown plant samples. Thirdly, phylogenetic reconstruction of unknown plants with known plants using UPGMA and neighbor-net network analysis. In order to get the final results about the plant samples identification, these three approaches were combined.

Multiple alignments of all the samples were generated (**Supplementary Figure S2**). The sequences for all the samples were highly conserved but different groups of plants with specific mutations were detectable. SNPs, indels and deletions were found throughout the aligned regions. The conserved region was shaded while the sites of mutations were not as shown in the **Supplementary Figure S2**. Though chloroplast like mitochondria is inherited from the mother parent only, this is exempted from genetic recombination during meiosis. Even, the major portion of the CpDNA is conserved, but the sequencing of the whole plastome of olive revealed that mutations such as SNPs, indels and SSRs are present. Some of the mutations are variety specific and this level of polymorphism is suitable to be used for cultivar identification.

In order to differentiate the unknown plants, phylogenetic reconstruction was carried out for all the samples including 100 unknown plants samples along with 10 known plants. A circular phylogenetic tree (**Figure 3**) demonstrates 17 clusters and 21 branches. Of them, 49 unknown plants clustered with 10 varieties of olive plants. These clusters include Frantoio and Gemlik (8 plants each), Coratina (5 plants), Arbosana and Chetoui (6 plants each), Carolea, Domat and Moraiolo (2 plants each), Leccino (3 plants) and Koroneiki (7 plants). The

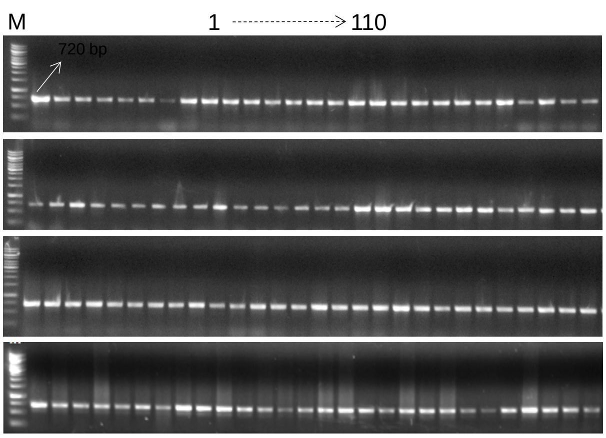

Figure 2. PCR product amplified with CP5 primer visualized on agarose gel. Each fragment is about 720 bp in length. 1→110 indicates samples and control PCR products These include 100 unknown samples and 10 reference known samples. "M" denotes marker (1 kb).

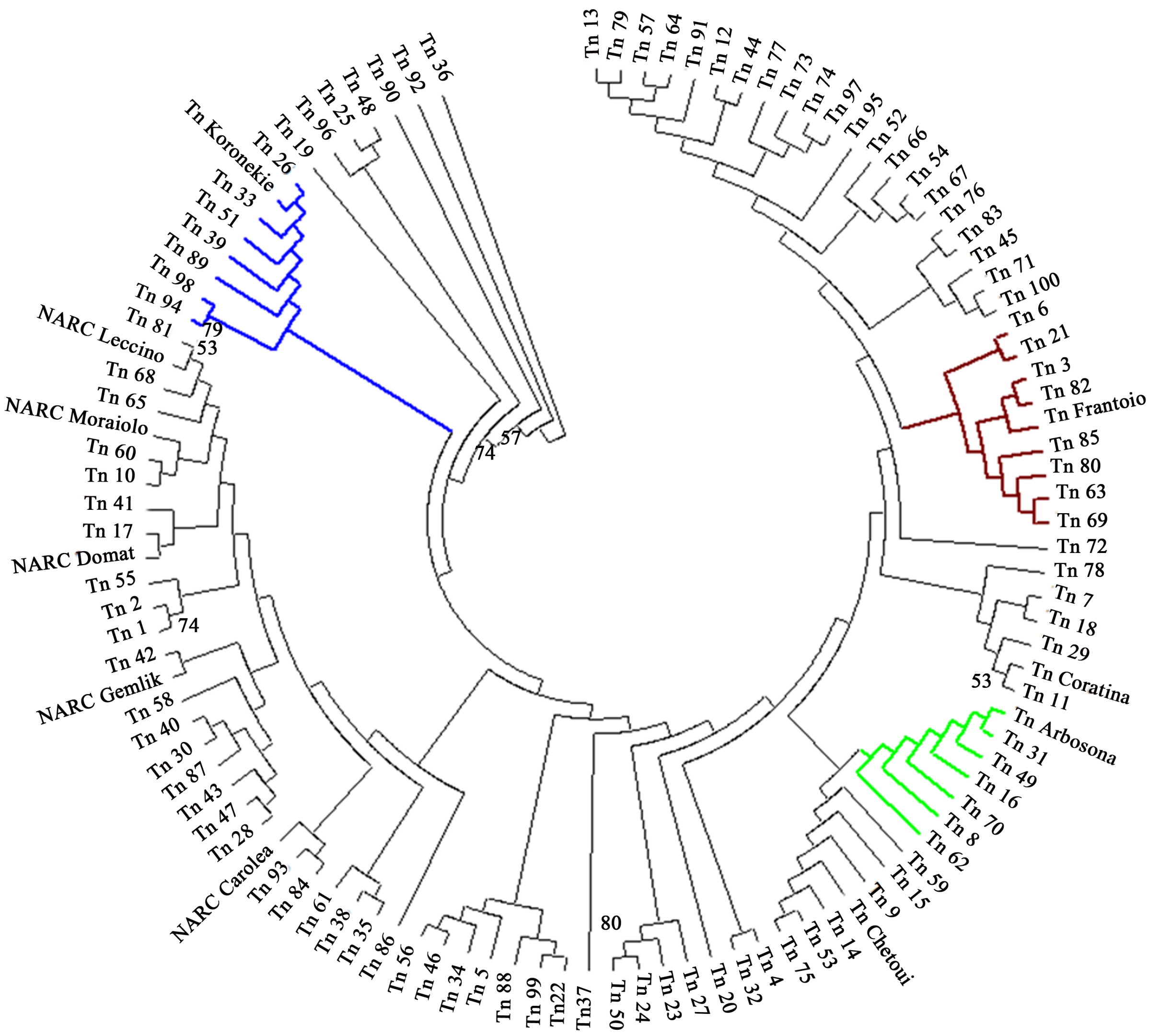

Figure 3. Phylogenetic circular tree of all the 110 olive plants samples. The evolutionary history was inferred using the UPGMA method. The optimal tree with the sum of branch length = 0.77561405 is shown. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 110 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 523 positions in the final dataset. Evolutionary analyses were conducted in MEGA6. The clusters with coloured branches were selected for further validation in two other phylogenetic reconstructions.

reference plants getting the maximum matches of 8 plants were Frantoio and Gemlik (8 each) while the olive varieties with minimum matches were Carolea, Domat and Moraiolo (2 plants each) (**Table 2**). The Koroneiki is found at the basal position while Frantoio is the most recent variety. The rest of samples did not cluster with any of the reference samples. They clustered together, separately from the known varieties and remained unknown. They constitute majority of the samples (51).

3.3. Variety Specific SNPs, Indels, SSRs Can Be Detected in Amplified Regions

For zooming in the data were fragmented into smaller sets. For example the 1st set contains the sequences of only Arbosana, Frantoio and Koroneiki and of the unknown plants in their clusters. A smaller phylogenetic UPGMA tree was constructed in MEGA6. **Figure 4** demonstrate that all the three clades retained their integrity by the clustering of the same unknown plants to their reference plants as in the circular tree thus validating the results obtained from the circular tree. The neighbour-net network better reveals recombination, homoplasmy and evo-

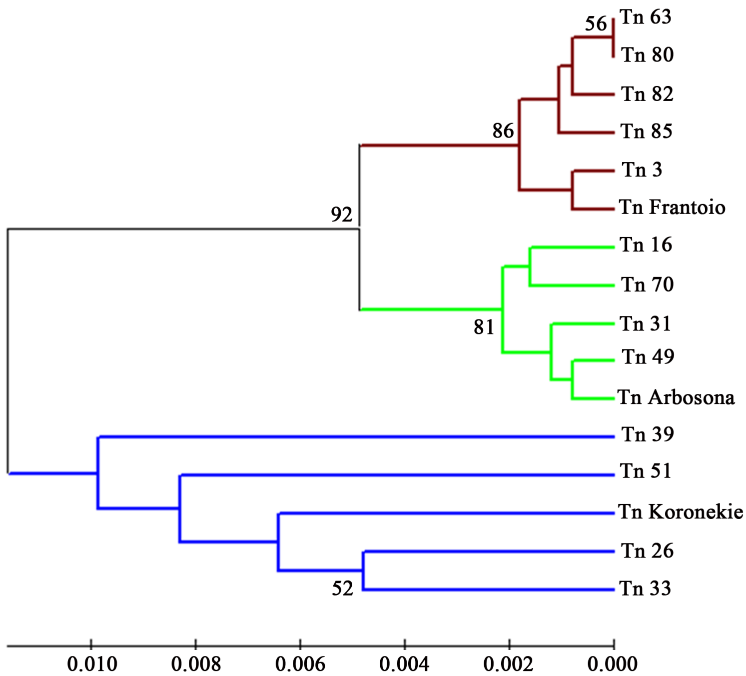

Figure 4. UPGMA phylogenetic tree showing unknown plants along with their reference plants. Tree was constructed using MEGA6 software. The topology of the tree is as that of the corresponding clusters in the circular tree. The number on the nodes indicates bootstrap values for 1000 replicates.

Table 2. List of plants identified using multiple alignments and phylogenetic reconstruction.

Sr. No.	Known Cultivars	No. of identified Plants
1	NARC_Carolea	2
2	NARC_Domat	2
3	NARC_Gemlik	8
4	NARC_Leccino	3
5	NARC_Moraiolo	2
6	Tn_Arbosana	6
7	Tn_Chetoui	6
8	Tn_Coratina	5
9	Tn_Frantoio	8
10	Tn_Koroneiki	7

lutionary relationship than a tree like structure. To further validate our results, the neighbour-net network of the sequences of three clusters was constructed in SpitsTree4 software (**Figure 5**). The resulting phylogenetic tree exhibited the same clusters of reference plants and unknown plants. The tree is clearly differentiated into three clusters. Though branches are scattered and are at distance in Koroneiki but it is the same cluster. Furthermore the tree retained the topology as UPGMA tree. So it can be concluded from all the three phylogenetic trees, that the mutations in the marker regions are variety specific. This marker region is reliable for the identification of olive varieties.

The multiple alignments of the sequences of marker regions of 16 plants showed a number of different SNPs at specific positions (**Figure 6**). Variety specific SNPs are present specifically in the marker region sequences of Frantoio and its clustered plants at positions 82, 258, 275 and 357 collected from Ternab. Similarly, Koroneiki

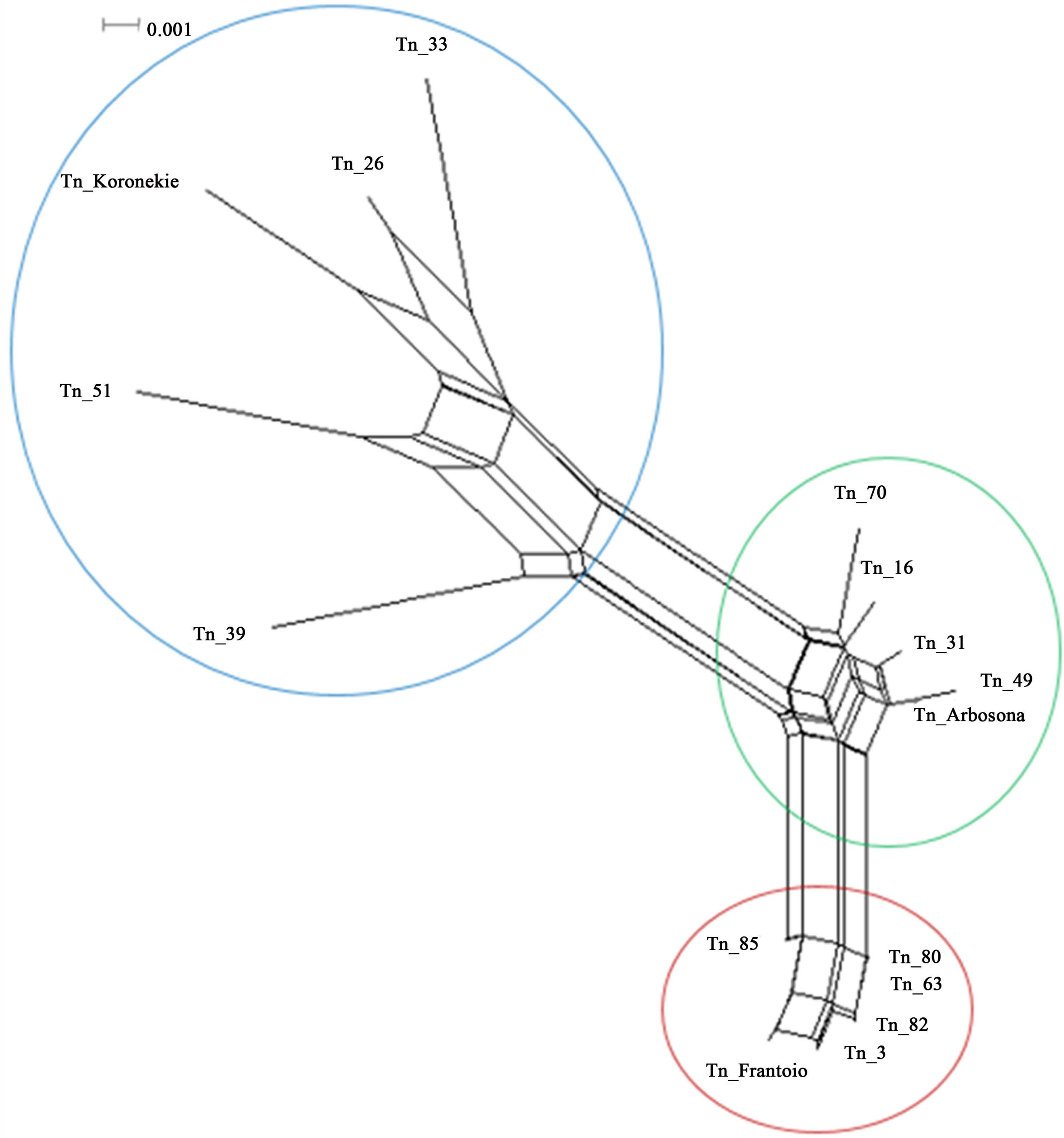

Figure 5. Neighbour-net network constructed with SplitsTree 4. The clusters retained their integrity thus further validating the corresponding clusters of the circular tree.

and the unknown plants in its cluster taken from Ternab have SNPs at 147, 163 and 221 positions. Similarly, Arbosana and the unknown plants in this cluster from Ternab have a common SNP at position 163 where T has substituted. Two deletions are also found at positions 532 and 536.

The unknown plants that have mutations corresponding to their reference plants and on this basis they have clustered together with a unique reference plant. These can be considered to be that variety sharing similarities in the chloroplast DNA sequence. This small dataset validated our results.

Over all data show that there are 49 plants differentiated into 10 varieties given as Arbosana, Carolea, Chetoui, Coratina, Domat, Frantoio, Gemlik, Koroneiki, Leccino and Moraiolo (**Figure 7**). A total of 188 mutations are present including SNPs and indels in 110 plants in the region amplified with CP5 marker shown in **Supplementary Figure S2**.

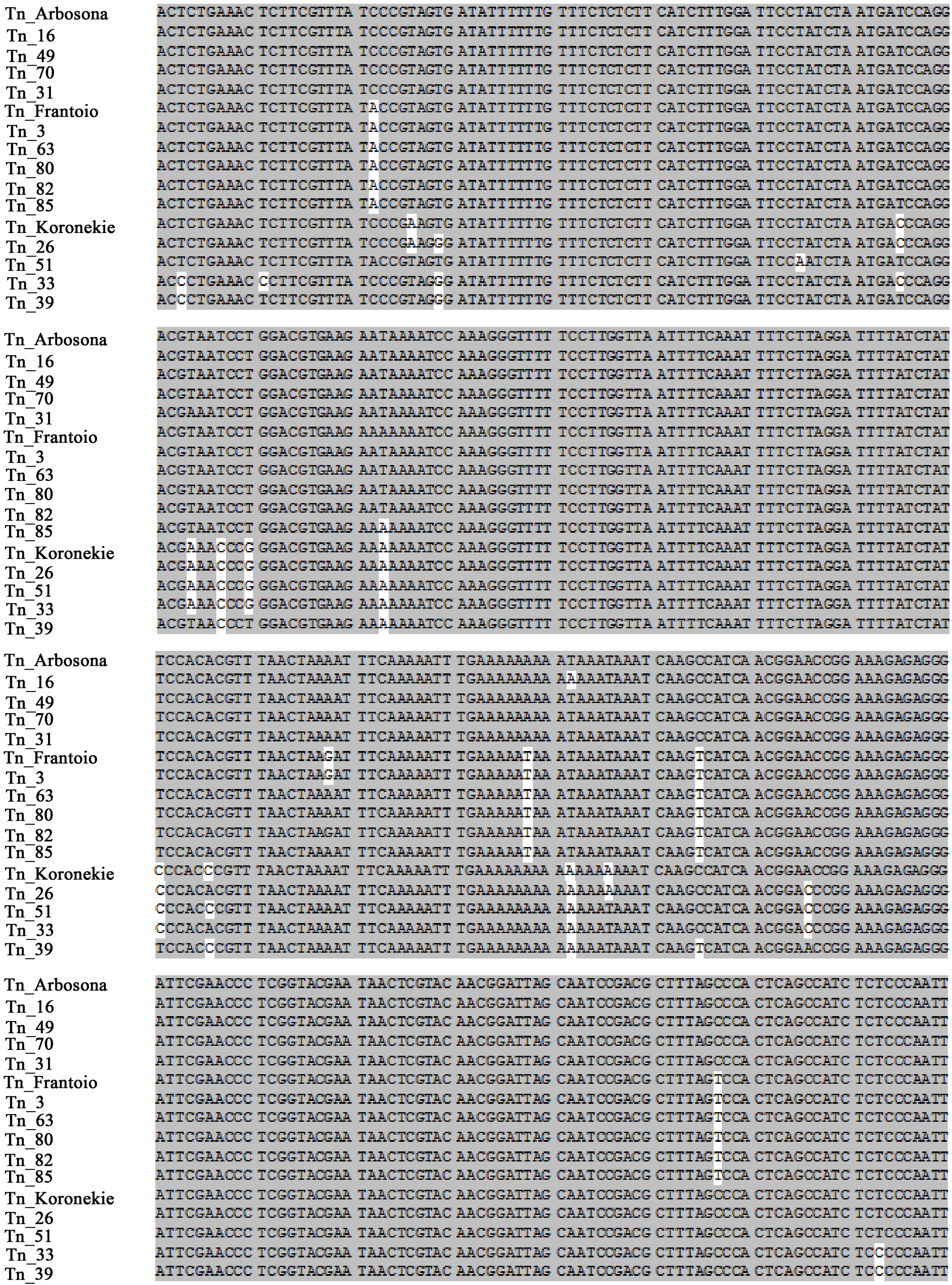

Figure 6. Multiple alignment of the sequences of the marker regions generated in BioEdit software. These are the sequences of the CP5 amplified marker regions of three reference plants (Arbosana, Frantoio and Koroneiki) and the 13 unknown plants (Tn_16, Tn_49….). Variety specific SNPs can be seen in the unknown and reference plant. The regions that are not shaded are the sites of SNPs. The shaded regions are the conserved sequences in this region of cpDNA of the olive plants shown here.

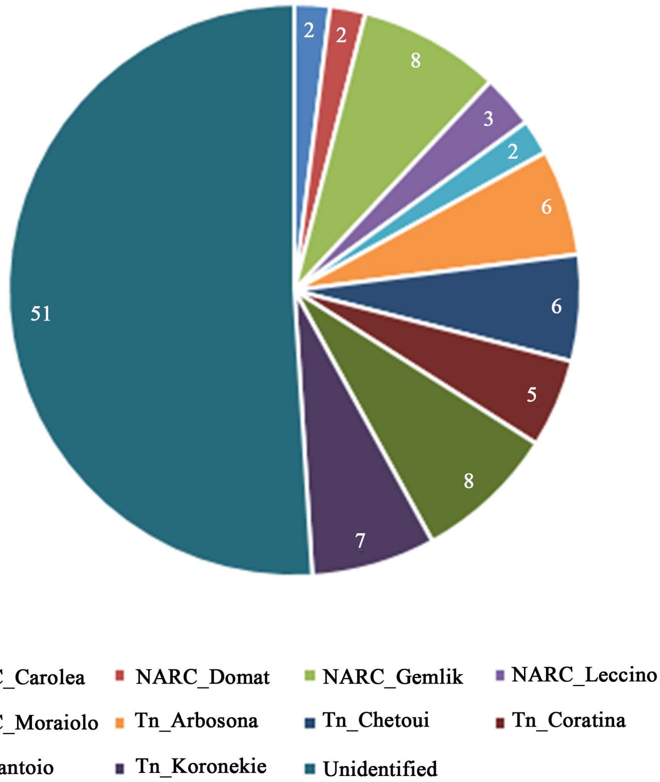

Figure 7. Graph showing the number of identified and unidentified olive plants on the basis of DNA sequence variations, multiple alignment and phylogenetic reconstruction. Frantoio and Gemlik revealed maximum matches of 8 each.

3.4. Identification of Olive Plants Using Multiple Alignments and Phylogenetics

Forty nine unknown plants were identified when a circular UPGMA (Unweighted Pair Group Mean Average) tree was reconstructed with MEGA6 (**Figure 3**). The remaining 51 unknown plants either clustered together or arranged separately but not with any of the known variety. They remained unidentified. The identified plants are written against their respective known variety in the **Table 3**.

Frantoio variety sampled from Ternab was clustered with 8 unknown plants. Gemlik sampled from NARC also clustered with 8 other unknown plants. It means, those plants that are clustered with Frantoio are all Frantoio. This is based on the similarity of the marker region and thus they clustered with their respective varieties. Five plants found to be Coratina, 6 were clustered with Arbosana, 6 with Chetoui. 2 were Carolea, 2 Domat, 2 clustered with Moraiolo and 3 found to be Leccino. A total of 49 unknown could be identified while the remaining 51 remained unidentified (**Figure 7**). They might also be identified by taking more reference controls.In order to find the closeness and differentiation at cultivar level, pairwise alignments were generated using BioEdit software to calculate the percent identity. In this connection, three plant cultivars represented in circular UPGMA tree (Tn_Arbosana, Tn_Frantoio and Tn_Koroneiki) were tested. The similarity is 99.26% - 99.81% between Tn_Arbosana and samples. It is 99.44% - 99.81% in Tn_Frantoio and its clustered plants. Similarly, Koroneiki and its samples are 98.17% - 99.16% identical as given in **Table 4**. It means these are closely related and represent one cultivar.

But surprisingly, the identity was less than 98% and even reduced to 96% between the different known cultivars. Tn_Arbosana and Tn_Frantoiohas 98% identity. Tn_Arbosana and Tn_Koroneikihas 97% and Tn_Frantoio and Tn_Koroneiki has 96% identity. Hence we can infer that 98% identity shows a different cultivar and above it is the same cultivar or plant.

4. Discussion

Varietal identification of olive plants is very important for further propagation and marketing of olive oil. The

Table 3. Identified olive varieties and number of plants from Tarnab olive orchard.

Sr. No.	Known cultivars	Identified unknown plants	Sr. No.	Known cultivars	Identified unknown plants
1	NARC_Carolea	Tn_Oc1_84 (R7-P5)[a] Tn_Oc1_93 (R8-P2)	2	NARC_Domat	Tn_Oc1_17 (R2-P12) Tn_Oc1_41 (R4-P8)
2	NARC_Leccino	Tn_Oc1_65 (R6-P2) Tn_Oc1_68 (R6-P5) Tn_Oc1_81 (R7-P2)	7	NARC_Moraiolo	Tn_Oc1_10 (R1-P10) Tn_Oc1_60 (R5-P13)
3	NARC_Gemlik	Tn_Oc1_28 (R3-P11) Tn_Oc1_30 (R3-P13) Tn_Oc1_40 (R4-P7) Tn_Oc1_42 (R4-P9) Tn_Oc1_43 (R4-P10) Tn_Oc1_47 (R4-P15) Tn_Oc1_58 (R5-P11) Tn_Oc1_87 (R7-P9)	8	Tn_Frantoio	Tn_Oc1_3 (R1-P3) Tn_Oc1_6 (R1-P6) Tn_Oc1_21 (R2-P16) Tn_Oc1_63 (R5-P16) Tn_Oc1_69 (R6-P6) Tn_Oc1_80 (R7-P1) Tn_Oc1_82 (R7-P3) Tn_Oc1_85 (R7-P6)
4	Tn_Arbosana	Tn_Oc1_8 (R1-P8) Tn_Oc1_16 (R2-P11) Tn_Oc1_31 (R3-P14) Tn_Oc1_49 (R5-P2) Tn_Oc1_62 (R5-P15) Tn_Oc1_70 (R6-P7)	9	Tn_Chetoui	Tn_Oc1_9 (R1-P9) Tn_Oc1_14 (R2-P6) Tn_Oc1_15 (R2-P9) Tn_Oc1_53 (R5-P6) Tn_Oc1_59 (R5-P12) Tn_Oc1_75 (R6-P12)
5	Tn_Koroneiki	Tn_Oc1_26 (R3-P9) Tn_Oc1_33 (R3-P16) Tn_Oc1_39 (R4-P6) Tn_Oc1_51 (R5-P4) Tn_Oc1_89 (R7-P13) Tn_Oc1_94 (R8-P4) Tn_Oc1_98 (R8-P8)	10	Tn_Coratina	Tn_Oc1_7 (R1-P7) Tn_Oc1_11 (R1-P11) Tn_Oc1_18 (R2-P13) Tn_Oc1_28 (R3-P11) Tn_Oc1_78 (R6-P15)

[a]Tn_Oc_1 (R1-P1) stands for Tarnab Orchard 1, Sample 1 in the Row 1 and Plant number 1.

Table 4. Sequence identity percentage calculated through pairwise alignment of the samples in three clusters of Tn_Arbosana, Tn_Frantoio and Tn_Koroneiki calculated in BioEdit software.

Tn_Arbosana		Tn_Frantoio		Tn_Koroneiki	
Unknown Plant	**Sequence Identity %**	**Unknown Plant**	**Sequence Identity %**	**Unknown Plant**	**Sequence Identity %**
Tn_16	99.63	Tn_3	99.81	Tn_26	99.16
Tn_49	99.81	Tn_63	99.44	Tn_51	99.08
Tn_70	99.26	Tn_80	99.44	Tn_33	98.79
Tn-31	99.81	Tn_82	99.63	Tn_39	98.17

majority of the cultivated olive plants present in Pakistan were brought from foreign countries, mostly Afghanistan and their variety name is not known and this is serious problem that the farmers are facing for years. They can differentiate these plants only by their morphology. They have no idea about the exact variety name or cultivar. As the morphological as well as biochemical parameters have limitations of being not reliable and very time consuming [27]. Thus it urged to develop a rapid, reliable and cost effective protocol for the accurate identification through DNA marker, an alternative. Molecular markers can detect DNA polymorphism to discriminate different cultivars in a very effective way [28].

The chloroplast genome of olive is the best platform for resolving the mixed and unknown plants of olive exactly into their varieties [29]. CpDNA is mostly conserved but has polymorphic regions enough to be used for this purpose. In this regard, the recent sequencing of the entire chloroplast genome of Frantoio cultivar is a big landmark. Marrioti and colleagues revealed 40 polymorphic regions in the CpDNA. Recent sequencing of plastid genome of the olive flaunts high resolution Cp markers for olive DNA fingerprinting [21]. Using this information, we designed a combination of chloroplast markers to amplify genes recruited in photosynthesis, ribosomal and NADH energy metabolism. Concatenated sequence of more than 100 unknown plants and 10 reference plants samples were analyzed using various bioinformatics and phylogenetic tools.

Scanning of entire chloroplast genome revealed 3 polymorphic regions. Multiple alignments of Frantoio and 5 NARC cultivars exhibited cultivar specific SNPs and deletions insertion that paved the way to extend this work to identify plants from 100 samples with more reference controls sampled from Ternab. Besnard and colleagues designed three markers in this region for identification of species or plants [30]. The plastid DNA regions screened by them showed a higher level of polymorphisms within the genus Olea than the rps16 and trnL-trnF sequences used in previous study [31]. The trnS-trnG intergenic spacer was the most variable region and was highly recommended for phylogenetic reconstructions of Oleaceae.

In this study, the marker region sequences of 100 unknown olive plants were analyzed. In order to investigate the evolutionary relationship, a phylogenetic tree was constructed taking 10 known reference plants. The tree clearly separated the samples into 10 clusters. These clusters include Arbosana, Carolea, Chetoui, Coratina, Domat, Frantoio, Gemlik, Koroneiki, Leccino and Moraiolo. This relationship shows that these plants have sequences similar to the known plants and might be the same variety. Multiple alignments were generated for the entire samples. The alignments revealed conservations groups in these plants on the basis of sequence similarities. This dataset was fragmented into smaller groups. Three clusters including Arbosana, Frantoio and Koroneiki were put under phylogenetic reconstruction again. There was a clear separation of these clusters along with unknown plants. This clustering was further validated using a neighbor net network in Splits Tree4 package. In order to find variety specific SNPs, a multiple alignment for these three clusters was generated. There was an obvious differentiation into three groups. "A" was specific to Koroneiki, "C" seemed to be preferable SNP for Frantoio. This is supported by pairwise alignments generated to calculate the percent identity between the samples of three clusters in circular phylogenetic reconstruction. The similarity is 99.26% - 99.81% between Tn_Arbosana and samples. It is 99.44% - 99.81% in Tn_Frantoio and its clustered plants. Similarly, Koroneiki and its samples are 98.17% - 99.16% identical. It means these are closely related and represent one cultivar. But surprisingly, the identity was less than 98% and even reduced to 96% between the different known cultivars. Hence we can infer that 98% identity shows a different cultivar and above it is the same cultivar or plant.

Taken together the data from all the approaches allow us to demonstrate that out of 100 plants 49 could be identified separated into 10 varieties. It is very important to mention that 51 plant samples could not be identified. They were not clustered into any of the known sequence clade. This means that there exist other varieties in these orchards for which we do not have any reference genome sequence. There are two solutions to this problem. First there is need to sequence more known varieties growing in Pakistan or to acquire the DNA of these varieties from other olive growing countries to be used as reference known genome. Secondly we need to sequence another nearby marker to expand gene region. Both the sequences will be joined. This is referred as concatenation of the sequences. It has more resolving power than a single sequence. Hence both sequences will be concatenated for alignment and phylogenetic reconstruction. This will generate more sequence diversity to get plants identified. An alternative strategy is to use nuclear markers (Cos markers) for which already many olive varieties have been sequenced. The implication of the above study is to identify all the fruit bearing unknown olive plants. The advent of high throughput genotyping through base calling SNPs has revolutionized the DNA fingerprinting. It is now possible to sequence the entire genome of the organisms and this technology is becoming cheaper ever passing day. This can be very practical for plants especially olive to sequence the entire plastome of all the samples.

5. Conclusion

In nutshell, our data reveal that the chloroplast genome of olive has polymorphic sites having variety specific SNPs and indels and they have resolving power to discriminate the olive plants at variety level. The Cp5 primer used successfully identified 49 varieties out of 100 unknown olive plants through mutations detection by alignment of the marker region sequences followed by the phylogenetic reconstruction with different bioinformatics software. This strategy can be further extended to characterize the olive tree germplasm reliably and efficiently with low costs which is distributed throughout the country in search of the better varieties. After the better varieties have been identified, this will enhance the olive oil and fruit production in Pakistan by the on-farm preservation and provision of the authentic germplasm to olive growers for the establishment to new olive orchards.

Acknowledgements

We are grateful to Dr. Muhammad Munir (National Project Director Olive), Dr. Nasir Cheema, PSO, Olive project at PARC for providing useful information aboit olive orchards and help there in sampling. We are thankful to Mr. Said Ahmed (Director, Olive Project KPK), Mr. Riaz Alam (Ternab), Dr. Muhammad Tariq (Director BARI), Mr. Muhammad Khalil (RA, BARI), Mr. Mukhtar (PSO, HRI, NARC), for providing the olive plants research material. Sincere appreciation and gratitude to Dr. Amir Ali Abbasi (NCB, QAU) for help in bioinformatics analysis. Italian funded project "Promotion of Olive Cultivation for Economic Development and Poverty" Pakistan Agricultural Research Council (PARC), Islamabad Pakistan is acknowledged for providing funding to Dr. MR Khan.

References

[1] Angiolillo, A., Mencuccini, M. and Baldoni, L. (1999) Olive Genetic Diversity Assessed Using Amplified Fragment Length Polymorphisms. *Theoretical and Applied Genetics*, **98**, 411-421. http://dx.doi.org/10.1007/s001220051087

[2] Doveri, S. and Baldoni, L. (2007) Olive. *Fruits and Nuts*, Springer, 253-264. http://dx.doi.org/10.1007/978-3-540-34533-6_13

[3] Golding, B., Reale, S., Doveri, S., Díaz, A., Angiolillo, A., Lucentini, L., Pilla, F., Martín, A., Donini, P. and Lee, D. (2006) SNP-Based Markers for Discriminating Olive (*Olea europaea* L.) Cultivars. *Genome*, **49**, 1193-1205.

[4] Green, P. (2002) A Revision of *Olea* L.(Oleaceae). *Kew Bulletin*, 91-140. http://dx.doi.org/10.2307/4110824

[5] Carriero, F., Fontanazza, G., Cellini, F. and Giorio, G. (2002) Identification of Simple Sequence Repeats (SSRs) in Olive (*Olea europaea* L.). *Theoretical and Applied Genetics*, **104**, 301-307. http://dx.doi.org/10.1007/s001220100691

[6] Lavee, S. (1985) *Olea europaea*. *Handbook of Flowering*, Vol. 6, CRC Press, Boca Raton, 423-434.

[7] Rugini, E. and Lavee, S. (1992) Olive. *Biotechnology of Perennial Fruit Crops*, CAB International, Wallingford, 371-382.

[8] Colmogro, S., Collins, G. and Sedgley, M. (2010) Processing Technology of the Table Olive. *Horticultural Reviews*, **25**, 235.

[9] Connels, J.H. (2005) History and Scope of the Olive Industry. *Olive Production Manual*, Vol. 3353, UCANR Publications, Oakland, 1-10.

[10] Sanz-Cortés, F., Martinez-Calvo, J., Badenes, M., Bleiholder, H., Hack, H., Llacer, G. and Meier, U. (2002) Phenological Growth Stages of Olive Trees (*Olea europaea*). *Annals of Applied Biology*, **140**, 151-157. http://dx.doi.org/10.1111/j.1744-7348.2002.tb00167.x

[11] Besnard, G., Breton, C., Baradat, P., Khadari, B. and Bervillé, A. (2001) Cultivar Identification in Olive Based on RAPD Markers. *Journal of the American Society for Horticultural Science*, **126**, 668-675.

[12] Lanza, B., Marsilio, V. and Martinelli, N. (1996) Olive Pollen Ultrastructure: Characterization of Exine Pattern through Image Analysis-Scanning Electron Microscopy (IA-SEM). *Scientia Horticulturae*, **65**, 283-294. http://dx.doi.org/10.1016/0304-4238(95)00868-3

[13] Barranco, D., Cimato, A., Fiorino, P., Rallo, L., Touzani, A., Castaneda, C., Serafin, F. and Truijillo, I. (2000) World Catalogue of Olive Varieties. International Olive Oil Council, Madrid.

[14] Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985) Individual-Specific "Fingerprints" of Human DNA. *Nature*, **316**, 76-79. http://dx.doi.org/10.1038/316076a0

[15] Giannoulia, K., Gazis, F., Nikoloudakis, N., Milioni, D. and Haralampidis, K. (2002) Breeding, Molecular Markers and Molecular Biology of the Olive Tree. *European Journal of Lipid Science and Technology*, **104**, 574-586. http://dx.doi.org/10.1002/1438-9312(200210)104:9/10<574::AID-EJLT574>3.0.CO;2-1

[16] Belaj, A., Caballero, J.M., Barranco, D., Rallo. L. and Trujillo, I. (2003) Genetic Characterization and Identification of New Accessions from Syria in an Olive Germplasm Bank by Means of RAPD Markers. *Euphytica*, **134**, 261-268. http://dx.doi.org/10.1023/B:EUPH.0000004954.93250.f5

[17] Besnard, G., Khadari, B., Villemur, P. and Bervillé, A. (2000) Cytoplasmic Male Sterility in the Olive (*Olea europaea* L.). *Theoretical and Applied Genetics*, **100**, 1018-1024. http://dx.doi.org/10.1007/s001220051383

[18] Besnard, G., Khadari, B., Baradat, P. and Bervillé, A. (2002) Combination of Chloroplast and Mitochondrial DNA Polymorphisms to Study Cytoplasm Genetic Differentiation in the Olive Complex (*Olea europaea* L.). *Theoretical and Applied Genetics*, **105**, 139-144. http://dx.doi.org/10.1007/s00122-002-0868-6

[19] Dumolin-Lapegue, S., Pemonge, M. and Petit, R. (1997) An Enlarged Set of Consensus Primers for the Study of Organelle DNA in Plants. *Molecular Ecology*, **6**, 393-397. http://dx.doi.org/10.1046/j.1365-294X.1997.00193.x

[20] Besnard, G., Casas, D., Rubio, R. and Vargas, P. (2003) A Set of Primers for Length and Nucleotide-Substitution Polymorphism in Chloroplastic DNA of *Olea europaea* L. (Oleaceae). *Molecular Ecology Notes*, **3**, 651-653. http://dx.doi.org/10.1046/j.1471-8286.2003.00547.x

[21] Mariotti, R., Cultrera, N.G., Díez, C.M., Baldoni, L. and Rubini, A. (2010) Identification of New Polymorphic Regions and Differentiation of Cultivated Olives (*Olea europaea* L.) through Plastome Sequence Comparison. *BMC Plant Biology*, **10**, 211. http://dx.doi.org/10.1186/1471-2229-10-211

[22] Olson, S.A. (1994) MacVector: An Integrated Sequence Analysis Program for the Macintosh. *Computer Analysis of Sequence Data*, Springer, 195-201. http://dx.doi.org/10.1385/0-89603-276-0:195

[23] Doyle, J.J. (1987) A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue. *Phytochemical Bulletin*, **19**, 11-15.

[24] Hall, T.A. (1999) BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95-98.

[25] Tamura, K, Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, **30**, 2725-2729. http://dx.doi.org/10.1093/molbev/mst197

[26] Huson, D.H. and Bryant, D. (2006) Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology Evolution*, **23**, 254-267. http://dx.doi.org/10.1093/molbev/msj030

[27] Tanksley, S.D. and Orton, T.J. (1983) Isozymes in Plant Breeding and Genetics. Elsevier.

[28] Busconi, M., Foroni, C., Corradi, M., Bongiorni, C., Cattapan, F. and Fogher, C. (2003) DNA Extraction from Olive Oil and Its Use in the Identification of the Production Cultivar. *Food Chemistry*, **83**, 127-134. http://dx.doi.org/10.1016/S0308-8146(03)00218-8

[29] Shaw, J., Lickey, E.B., Schilling, E.E. and Small, R.L. (2007) Comparison of Whole Chloroplast Genome Sequences to Choose Noncoding Regions for Phylogenetic Studies in Angiosperms: The Tortoise and the Hare III. *American Journal of Botany*, **94**, 275-288. http://dx.doi.org/10.3732/ajb.94.3.275

[30] Besnard, G., de Casas, R.R., Christin, P.A. and Vargas, P. (2009) Phylogenetics of Olea (Oleaceae) Based on Plastid and Nuclear Ribosomal DNA Sequences: Tertiary Climatic Shifts and Lineage Differentiation Times. *Annals of Botany*, 105. http://dx.doi.org/10.1093/aob/mcp105

[31] Wallander, E. and Albert, V.A. (2000) Phylogeny and Classification of Oleaceae Based on rps16 and trnL-F Sequence Data. *American Journal of Botany*, **87**, 1827-1841. http://dx.doi.org/10.2307/2656836

Abbreviations

Cp, Chloroplast
CTAB, Cetyl Trimethyl Ammonium Bromide
MEGA, Molecular Evolution Genetics Analysis
NARC, National Agricultural Research Centre
NCBI, National Center for Biotechnology Information
RAPD, Random Amplified Polymorphic DNA
SNP, Single Nucleotide Polymorphism
SSR, Simple Sequence Repeat
UPGMA, Unweighted Pair Grouped Method with Arithmetic Mean

Supplementary

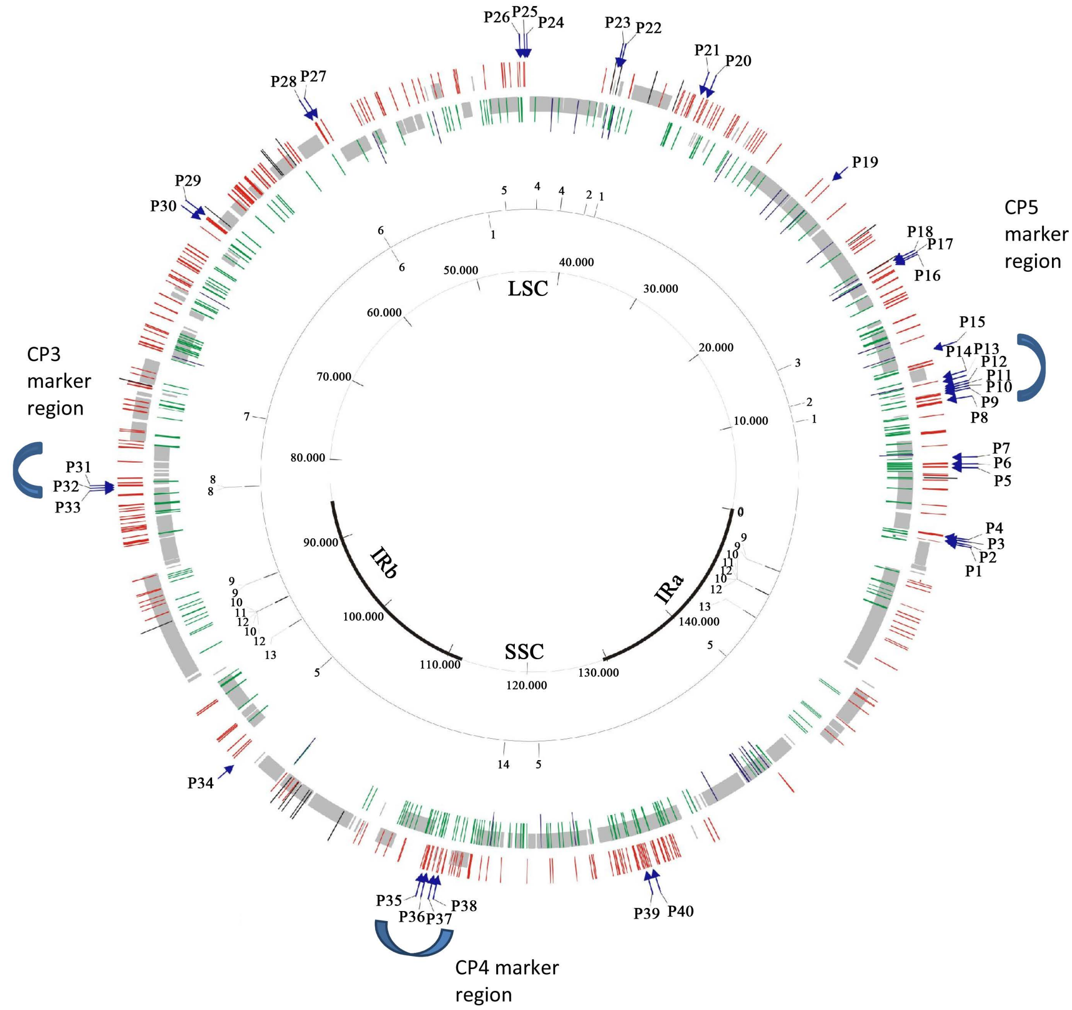

Figure S1. Polymorphic sites of olive chloroplast genome cv. Frantoio. (Adopted from Mariotti *et al.* (2010)). The three marker regions namely CP3, CP4 and CP5 are shown. The different colours indicate the four mono-nucleotide microsatellites (poly-T and poly-G are reported in the external circle, poly-A and poly-C in the internal circle), bar lengths correspond to the number of repetitions. Arrows indicate polymorphisms (base mutations, microsatellites and indels). The circle reports the interspersed repeats to the same number corresponds the same repetition. External or internal number position corresponds to the sense or anti-sense sequence direction.

Table S1. Scanning of plastid genome of Frantoio and selection of three polymorphic region for which 3 pairs of primers if were designed.

S. No	Marker Name	Sequence (5' to 3')	Product Size (bp)	Position in CP Genome (bp)
1	CP3-F	CCTTCTCGGAAAAGTATTTTCACA	740	83112-83852
	CP3-R	CATCCTTTGCATTGGAAGAATAGA		
2	CP4-F	GCTGAATAGACAGATTCATTGAAA	1334	101265-102599
	CP4-R	CCAGACTCTCTTCACTAAGTGTTA		
3	**CP5–F**	**CTGACAATTCATTTCTATTTCTAGA**	**720**	**8986-9705**
	CP5-R	**CATTATTTATCTATAATTCGTTGGA**		

Amplified length is also given, bold encloses he CP5 primer that was selected for further analysis.

```
Tn_1   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_2   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_3   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_4   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_5   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_6   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_7   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_8   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_9   AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_10  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_11  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_12  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_13  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_14  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_15  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_16  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_17  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_18  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_19  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACCCTGAAAC TCTTCGTTTA
Tn_20  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAA-TCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_21  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_22  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC CCTTCGTTTA
Tn_23  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_24  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_25  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_26  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_27  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_28  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_29  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_30  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_31  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_32  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_33  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACCCTGAAAC CCTTCGTTTA
Tn_34  AAGATCTATT CTCTTTTTTT TTTT-AAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTCCGTTTA
Tn_35  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAAGGCTT ACTCTGAAAC TCTTCGTTTA
Tn_36  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAAGGCTT ACCCTGAAAC CCTTCGTTTA
Tn_37  AAGATCTATT CTCTTTTTTT TTTT-AAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_38  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG GGTAAGGCTT ACTCTGAAAC TCTTCGTTTA
Tn_39  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACCCTGAAAC TCTTCGTTTA
Tn_40  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_41  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_42  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG GGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_43  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_44  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_45  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_46  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_47  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAAGGCTT ACTCTGAAAC TCTTCGTTTA
Tn_48  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAAGGCTT ACTCTGAAAC TCTTCGTTTA
Tn_49  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_50  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_51  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_52  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_53  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_54  AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
```

(a)

Tn_55	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_56	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_57	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_58	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAAGGCTT ACTCTGAAAC TCTTCGTTTA
Tn_59	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_60	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_61	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_62	AAGA-CTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_63	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_64	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_65	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_66	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_67	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_68	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_69	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_70	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_71	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_72	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_73	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_74	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_75	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_76	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_77	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_78	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_79	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_80	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_81	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_82	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_83	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_84	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_85	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_86	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATGG GGTAAGGCTT ACTCTGAAAC TCTTCGTTTA
Tn_87	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_88	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_89	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_90	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TGGGAGATTG GGTAATGCTT ACCCTGAAAC CCTCCGTTTA
Tn_91	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_92	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG GGTAAGGCTT ACTCTAAAAC TCTTCTTTTA
Tn_93	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_94	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC CCTCCGTTTA
Tn_95	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_96	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG GGTAAGGCTT ACTCTGAAAC TCTTCGTTTA
Tn_97	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_98	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_99	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_100	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Carolea	AGATCTATTT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAAATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Domat	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Gemik	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Leccino	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
NARC_Moraiolo	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_Arbosana	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_Chetoui	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_Coratina	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_Koroneiki	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA
Tn_Frantoio	AAGATCTATT CTCTTTTTTT TTTTCAAAAA AAAAATCATC TTGGAGATTG TGTAATGCTT ACTCTGAAAC TCTTCGTTTA

(b)

```
Tn_1    TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_2    TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_3    TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_4    TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_5    TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGACCCAGG ACGAAACCCG GGACGTGAAG
Tn_6    TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_7    TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_8    TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_9    TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_10   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_11   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_12   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_13   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_14   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCG GGACGTGAAG
Tn_15   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_16   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_17   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_18   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_19   TCCCGTAGTG ATTTTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGACCCGGG ACGTAATCCG GGACGTGAAG
Tn_20   TACCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_21   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_22   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCG GGACGTGAAG
Tn_23   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCT GGACGTGAAG
Tn_24   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_25   TACCGAAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA AGGATCCAGG ACGTAATCCG GGACGTGAAA
Tn_26   TCCCGAAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGACCCAGG ACGAAACCCG GGACGTGAAG
Tn_27   TCCCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCT GGACGTGAAG
Tn_28   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_29   TACCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_30   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_31   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCT GGACGTGAAG
Tn_32   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_33   TCCCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGACCCAGG ACGAAACCCG GGACGTGAAG
Tn_34   TCCCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCG GGACGTGAAG
Tn_35   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_36   AACCGAAGGG AAATTTTTTG TTTCCCCCTC CACCTTTGGA TCCCAATCTA ATGACCCAGG ACGTAACCCG GGACGGGAAG
Tn_37   TCCCGAAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_38   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCG GGACGTGAAA
Tn_39   TCCCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAACCCT GGACGTGAAG
Tn_40   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_41   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_42   TACCGAAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_43   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_44   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_45   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_46   TCCCGTAGTG ATATTTTTTG TTTCCCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAACCCG GGACGTGAAG
Tn_47   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_48   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA AGGACCCAGG ACGTAACCCG GGACGTGAAA
Tn_49   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_50   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_51   TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCAATCTA ATGATCCAGG ACGAAACCCG GGACGTGAAG
Tn_52   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_53   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_54   TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
```

(c)

Tn_55	TACCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_56	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAACCCT GGACGTGAAG
Tn_57	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_58	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_59	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_60	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_61	TACCGAAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_62	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_63	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_64	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_65	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCG GGACGTGAAG
Tn_66	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_67	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_68	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_69	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_70	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_71	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_72	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_73	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_74	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_75	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_76	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_77	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_78	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_79	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_80	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_81	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_82	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_83	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_84	TACCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_85	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_86	TACCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_87	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_88	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCG GGACGTGAAG
Tn_89	TCCCGTAGTG ATATTTTTTG TTTCCCCCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCG GGACGTGAAG
Tn_90	ACCCGAAGGG ATATTTTTTG TTTCCCCCTT CATCTTTGGA TTCCTATCTA AGGACCCAGG ACGAAACCCG GGACGTGAAA
Tn_91	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_92	TCCCGAAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAACCCG GGACGTGAAA
Tn_93	TACCGAAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
Tn_94	TCCCGAAGTG ATATTTTTTG TTTCCCCCTT CATCTTTGGA TTCCTATCTA AGGACCCAGG ACGAAACCCG GGACGGGAAG
Tn_95	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_96	TCCCGAAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA AGGACCCAGG ACGTAACCCG GGACGTGAAA
Tn_97	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_98	TCCCGTAGGG ATATTTTTTG TTTCCCCCTT CATCTTTGGA TTCCTATCTA AGGACCCAGG ACGAAACCCG GGACGTGAAG
Tn_99	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGAAATCCG GGACGTGAAG
Tn_100	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
NARC_Carolea	TACCGAAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
NARC_Domat	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
NARC_Gemlik	TACCGTAGGG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAA
NARC_Leccino	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
NARC_Moraiolo	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_Arbosana	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_Chetoui	TCCCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_Coratina	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG
Tn_Koroneiki	TCCCGAAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGACCCAGG ACGAAACCCG GGACGTGAAG
Tn_Frantoio	TACCGTAGTG ATATTTTTTG TTTCTCTCTT CATCTTTGGA TTCCTATCTA ATGATCCAGG ACGTAATCCT GGACGTGAAG

(d)

Tn_1	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_2	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_3	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_4	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_5	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT CCCACACGTT TAACTAAGAT
Tn_6	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_7	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_8	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT CCCACACGTT TAACTAAAAT
Tn_9	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_10	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_11	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_12	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_13	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_14	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_15	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_16	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_17	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_18	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_19	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACCCGTT TAACTAAGAT
Tn_20	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_21	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_22	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_23	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_24	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_25	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_26	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT CCCACACGTT TAACTAAAAT
Tn_27	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_28	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_29	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_30	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_31	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_32	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_33	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT CCCACACGTT TAACTAAAAT
Tn_34	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_35	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_36	AAAAAAACCC AAAGGGTTTT CCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTTT CCCCCCCGTT TAACAAAAAT
Tn_37	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_38	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_39	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACCCGTT TAACTAAAAT
Tn_40	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_41	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_42	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_43	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_44	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_45	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_46	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_47	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_48	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACCCGTT TAACTAAAAT
Tn_49	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_50	AATAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_51	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT CCCACCCGTT TAACTAAAAT
Tn_52	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT
Tn_53	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAAAT
Tn_54	AAAAAAATCC AAAGGGTTTT TCCTTGGTTA ATTTTCAAAT TTTCTTAGGA TTTTATCTAT TCCACACGTT TAACTAAGAT

(e)

Tn_55	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_56	TTCAAAAATT TGAAAAAAAA AAAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_57	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_58	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_59	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_60	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
Tn_61	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_62	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_63	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_64	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_65	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
Tn_66	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_67	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_68	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_69	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_70	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_71	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_72	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_73	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_74	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_75	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_76	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_77	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_78	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_79	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_80	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_81	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_82	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_83	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_84	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_85	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_86	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGACCCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_87	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_88	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCC-TCA ACGGACCCGG AAAGAGAGG- ATTCGAACCC TCGGTACGAA
Tn_89	TTCAAAAATT TGAAAAAAAA AAAAATAAAC CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_90	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCACCA CCGGACCCGG AAAGAGAGGG ATTCCAACCC TCGGTACAAA
Tn_91	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_92	TTCAAAAATT TGAAAAAAAA ATAAAAAAAT CAAGCCATCA CCGGACCCGG AAAAAGAGGG ATTCAACCCC CCGGTCCAAA
Tn_93	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTCCGAA
Tn_94	TTCAAAAATT TGAAAAATAA ATAAATAAAC CAAGCCCTCA CCGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_95	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_96	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAAAAAGGG ATTCAAACCC TCGGTACAAA
Tn_97	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_98	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCC-TCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_99	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_100	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
NARC_Carolea	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAAAAAGGG ATTCAAACCC TCGGTACGAA
NARC_Domat	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
NARC_Gemlik	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
NARC_Leccino	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
NARC_Moraiolo	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_Arbsana	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_Chetoui	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_Coratina	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_Koroneiki	TTCAAAAATT TGAAAAAAAA AAAAAAAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_Frantoio	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA

(f)

Tn_1	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_2	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_3	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_4	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_5	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_6	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_7	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_8	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_9	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_10	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_11	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_12	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_13	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_14	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_15	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_16	TTCAAAAATT TGAAAAAAAA AAAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_17	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
Tn_18	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_19	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_20	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_21	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_22	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_23	TTCAAAAATT TGAAAAAAAA AAAAAAAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_24	TTCAAAAATT TGAAAAAAAA AAAAAAAAAT CAAGCCCTCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_25	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_26	TTCAAAAATT TGAAAAAAAA AAAAAAAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_27	TTCAAAAATT TGAAAAAAAA AAAAAAAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_28	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_29	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_30	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_31	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_32	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_33	TTCAAAAATT TGAAAAAAAA AAAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_34	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_35	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_36	TTCAAAAATT TGAAAAAAAA AAAAAAAAAA AAAGCCCCCC CCGGCCCCGG AAGAAGAGGG AGTCGACCCC CCCCCCCAAA
Tn_37	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_38	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_39	TTCAAAAATT TGAAAAAAAA AAAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_40	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_41	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
Tn_42	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_43	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_44	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_45	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_46	TTCAAAAATT TGAAAAAAAA AAAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_47	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_48	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_49	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_50	TTCAAAAATT TGAAAAAAAA AAAAAAAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_51	TTCAAAAATT TGAAAAAAAA AAAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_52	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_53	TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_54	TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA

(g)

```
Tn_55           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_56           TTCAAAAATT TGAAAAAAAA AAAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_57           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_58           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_59           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_60           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
Tn_61           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_62           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_63           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_64           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_65           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
Tn_66           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_67           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_68           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_69           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_70           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_71           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_72           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_73           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_74           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_75           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_76           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_77           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_78           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_79           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_80           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_81           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_82           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_83           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_84           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_85           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_86           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGACCCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_87           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_88           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCC-TCA ACGGACCCGG AAAGAGAGG- ATTCGAACCC TCGGTACGAA
Tn_89           TTCAAAAATT TGAAAAAAAA AAAAATAAAC CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_90           TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCACCA CCGGACCCGG AAAGAGAGGG ATTCCAACCC TCGGTACAAA
Tn_91           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_92           TTCAAAAATT TGAAAAAAAA ATAAAAAAAT CAAGCCATCA CCGGACCCGG AAAAAGAGGG ATTCAACCCC CCGGTCCAAA
Tn_93           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTCCGAA
Tn_94           TTCAAAAATT TGAAAAATAA ATAAATAAAC CAAGCCCTCA CCGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_95           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_96           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAAAAAGGG ATTCAAACCC TCGGTACAAA
Tn_97           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_98           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCC-TCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_99           TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_100          TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGACCCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
NARC_Carolea    TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAAAAAGGG ATTCAAACCC TCGGTACGAA
NARC_Domat      TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACGAA
NARC_Gemlik     TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
NARC_Leccino    TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
NARC_Moraiolo   TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCAAACCC TCGGTACAAA
Tn_Arbsana      TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_Chetoui      TTCAAAAATT TGAAAAAAAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_Coratina     TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_Koroneiki    TTCAAAAATT TGAAAAAAAA AAAAAAAAAT CAAGCCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
Tn_Frantoio     TTCAAAAATT TGAAAAATAA ATAAATAAAT CAAGTCATCA ACGGAACCGG AAAGAGAGGG ATTCGAACCC TCGGTACGAA
```

(h)

Tn_1	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_2	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_3	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_4	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_5	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_6	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_7	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_8	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_9	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_10	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAAA -TTACTACCT
Tn_11	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_12	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGA-A ATTACTACCT
Tn_13	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_14	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_15	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_16	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_17	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_18	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_19	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_20	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACAACCT
Tn_21	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_22	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_23	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_24	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_25	TAACTCGTAC AACGGATTAG CAATCCAACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_26	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_27	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_28	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAAA ATTACTACCT
Tn_29	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_30	TAACTCGTAC AACGGATTAG CAATCCAACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAAA ATTACTACCT
Tn_31	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_32	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_33	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCCCCCAATT GAAAAAGAGA ATTACTACCT
Tn_34	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_35	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_36	AAACACCTCA AACGAAGAAG AAACCCCCCG CCTTTGCCCC CCCGCCCCCC CCCCCCACCT AAAAAAGAAA AAGACCCCCT
Tn_37	AAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_38	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAAA ATTACTACCT
Tn_39	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCCCCCAATT GAAAAAGAGA ATTACTACCT
Tn_40	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_41	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_42	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_43	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_44	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_45	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_46	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_47	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAAA ATTACTACCT
Tn_48	TAACTCGTAC AACGGATTAG CAATCCAACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_49	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_50	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_51	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_52	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_53	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_54	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_55	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAAA ATTACTACCT

(i)

Tn_56	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_57	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_58	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_59	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_60	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_61	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_62	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_63	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_64	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_65	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_66	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_67	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_68	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_69	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_70	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_71	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_72	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_73	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_74	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_75	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_76	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_77	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_78	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_79	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_80	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_81	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_82	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_83	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_84	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_85	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_86	TAACTCGTAC AACGGATTAG CAATCCAACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_87	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_88	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACAACCT
Tn_89	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_90	TAACTCGTAC AACGGATTAC CAACCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_91	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_92	AAACTCGTAC AACGGATTAG CAATCCAACC CTTTATCCCA CTCACCCATC TCTCCCAATT GAAAAAAAAA ATTACTCCCT
Tn_93	TAACTCGTAC AACGGATTAG CAATCCAACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_94	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_95	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_96	AAACTCGTAC AACGGATTAG CAACCCAACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAAAAA ATTACTACCT
Tn_97	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_98	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_99	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_100	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Carolea	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Domat	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Gemlik	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Leccino	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
NARC_Moraiolo	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_Arbosana	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_Chetoui	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_Coratina	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_Koroneiki	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGCCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT
Tn_Frantoio	TAACTCGTAC AACGGATTAG CAATCCGACG CTTTAGTCCA CTCAGCCATC TCTCCCAATT GAAAAAGAGA ATTACTACCT

(j)

Tn_1	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_2	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_3	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_4	TACATATAAT GTAAGGAGCC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_5	TACATATAAT GTAAGGAGCC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_6	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_7	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_8	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_9	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_10	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_11	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_12	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_13	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_14	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_15	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_16	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_17	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_18	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_19	TACATATAAT GTAAGGGGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT CCTTTTAGAT TAGATAATTA
Tn_20	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_21	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_22	TACATATAAT GTAAGGAGCC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_23	TACATATAAT GTAAGGAGTC TTTCTTTC-C TATTCTATAG AGAT-TACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_24	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_25	TACATATAAG GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAAAT TAAATAATTA
Tn_26	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_27	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_28	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_29	TACATATAAT GTAAGGAGTC TTTCTTTCCC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_30	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_31	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_32	TACATATAAT GTAAGGAGCC TTTCTTTCTC TATTCTATAG AGAT-TACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_33	TACATATAAT GTAAGGAGCC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_34	TACATATAAT GTAAGGAGCC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_35	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAAAT TAGATAATTA
Tn_36	CCCTTAAAAA AAAGGGAGGC CCTCTTTCCT TCCCCTCTCG AGAGAGAAAA ACAGAAAATT TCTTTCATGT AGAAATATAT
Tn_37	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_38	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_39	TACATATAAT GTAAGGAGCC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_40	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_41	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_42	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_43	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAAAT TAGATAATTA
Tn_44	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_45	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_46	TACATATAAT GTAAGGAGCC TTTCTTTC-C TATTCTATAG AGAT-TACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_47	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAAAT TAGATAATTA
Tn_48	TACATATAAG GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGAT-TACAA ATCAGGAATT TCTTTTAAAT TAGA-AATTA
Tn_49	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_50	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_51	TACATATAAT GTAAGGAGCC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_52	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_53	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_54	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA
Tn_55	TACATATAAT GTAAGGAGTC TTTCTTTCTC TATTCTATAG AGATATACAA ATCAGGAATT TCTTTTAGAT TAGATAATTA

(k)

Tn_56	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_57	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_58	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAAAT	TAGATAATTA
Tn_59	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_60	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_61	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAAAT	TAGATAATTA
Tn_62	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_63	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_64	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_65	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_66	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_67	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_68	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_69	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_70	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_71	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_72	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_73	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_74	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_75	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	T-TTCTATAG	AGAT-TACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_76	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_77	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_78	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_79	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_80	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_81	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_82	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_83	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_84	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_85	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_86	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_87	TACATATAAG	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_88	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_89	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGAT-TACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_90	TACATATAAG	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGA-ATACAA	ATCAGGAATT	TCTTTTAAAT	AAGATAATTA
Tn_91	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_92	TCCATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAA	AGATATCCAA	ATCAGGAATT	TCTTTTAAAT	TAAATAATTA
Tn_93	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_94	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_95	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_96	TACATATAAG	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAAAT	TAAATAATTA
Tn_97	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_98	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_99	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_100	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGAT-TACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
NARC_Carolea	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
NARC_Domat	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
NARC_Gemlik	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGAT-TACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
NARC_Leccino	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
NARC_Moraiolo	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_Arbosana	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_Chetoui	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_Coratina	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_Koroneiki	TACATATAAT	GTAAGGAGCC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA
Tn_Frantoio	TACATATAAT	GTAAGGAGTC	TTTCTTTCTC	TATTCTATAG	AGATATACAA	ATCAGGAATT	TCTTTTAGAT	TAGATAATTA

(l)

Tn_1	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAAACTT CTTTGGGTTG
Tn_2	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAAACTT CTTTGGGTTG
Tn_3	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_4	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_5	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGAT-T CTTTG-----
Tn_6	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGCTTCTTTG------
Tn_7	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGCTTCTTTG------
Tn_8	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AGGAGGCTTCTTT-------
Tn_9	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAGG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_10	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT TTTTGTGTTG
Tn_11	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_12	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_13	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_14	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_15	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGATAT CTTT------
Tn_16	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_17	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTT-
Tn_18	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT TTTT------
Tn_19	GATAAAGGAG GGGCTCGAAC GAGCCTTTAA ATAAATAAAG AAAAAAAAAA AGGAGGCCTT TTT-------
Tn_20	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAGG AAAAAAAAAA AGGAAGACTT CTT-------
Tn_21	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTCTTTG------
Tn_22	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_23	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_24	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_25	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTTG
Tn_26	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA A-GAAGA-AT CTTTG-----
Tn_27	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACAT CTTTGTGTG-
Tn_28	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAA-CTT TTTTGGGTTG
Tn_29	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_30	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGTGTTG
Tn_31	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_32	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT TTTT------
Tn_33	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_34	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_35	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAAAAT TTTTGGGTGG
Tn_36	AATAGAAGAG GCCCCCCCCC CCCCCACCAA ATAATTAAAA AAAAAAAAAA AAAAAAATCT TTTTTTTTT-
Tn_37	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAAT-T CTTTGGGTGG
Tn_38	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT TTTTGGGTTG
Tn_39	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACAT CTTTGTGTG-
Tn_40	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTTG
Tn_41	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTTG
Tn_42	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGTGTTG
Tn_43	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT TTTTGTGTTG
Tn_44	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_45	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT TTTT------
Tn_46	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_47	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT TTTTGGGTTG
Tn_48	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAA-TAAAG AAAAAAAAAA AAAAAAACTT TTTTGGGTA-
Tn_49	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG GAAAAAAAAA AAGAAGACAT CTTT------
Tn_50	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_51	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAGG AAAAAAAAAA AAAAAAACAT CTTTGTGTG-
Tn_52	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_53	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_54	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_55	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAAACAT CTTTGGGTGG

(m)

Tn_56	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_57	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGATAT CTTT------
Tn_58	GATAAAGGAA GGGCTCAAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTGG
Tn_59	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_60	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT TTTTGTGTTG
Tn_61	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAAC-- TCTTGGGGTA
Tn_62	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGATTC TTTG------
Tn_63	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_64	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_65	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACAT CTTTGTGTTG
Tn_66	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_67	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_68	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAAT-T CTTTGTGTGG
Tn_69	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_70	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AGGAAGACTT CTT-------
Tn_71	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AGGAAGACTT CTT-------
Tn_72	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG GAAAAAAAAA AAGAAGACAT CTTT------
Tn_73	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_74	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_75	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_76	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_77	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGA-AT CTTTG-----
Tn_78	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAGG AAAAAAAAAA AGGA-GACAT CTTT------
Tn_79	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_80	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_81	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTTG
Tn_82	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_83	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_84	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAAACAT CTTTGTGTTG
Tn_85	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_86	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACAT CTTTGTGTTG
Tn_87	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTTG
Tn_88	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAGACAT CTTTGTG---
Tn_89	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAGG AAAAAAAAAA AGGAAGACTT TTT-------
Tn_90	GATAAGGGAG GGGCTCCAAC GAGCCTATAA ATAAATAAAA AAAAAAAAAA ATAAAAATCC TT--------
Tn_91	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_92	AATAAAGGAA GGGCTCAAAC GACCCTATAA ATAAATAAAA AAAAAAAAAA AAAAAAACTT TTTTGGGTT-
Tn_93	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACAT CTTTGGGTG-
Tn_94	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAGATAT CTTTGTG---
Tn_95	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAATTAAGG AAAAAAAAAA AGGA-GACTT CTTT------
Tn_96	AATAAAGGAA GGGCTCAAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTG-
Tn_97	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AGGAAGACTT CTT-------
Tn_98	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACAT CTTTGTGTG-
Tn_99	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_100	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
NARC_Carolea	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAAACTT CTTG------
NARC_Domat	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGTGTT-
NARC_Gemlik	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAA-TAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTA-
NARC_Leccino	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACAT CTTTGTGTT-
NARC_Moraiolo	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAAAAAACTT CTTTGGGTT-
Tn_Arbosana	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACAT CTTT------
Tn_Chetoui	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AGGAAGACTT CTT-------
Tn_Coratina	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------
Tn_Koroneiki	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAGG AAAAAAAAAA AGGAAGACTT CTT-------
Tn_Frantoio	GATAAAGGAA GGGCTCGAAC GAGCCTATAA ATAAATAAAG AAAAAAAAAA AAGAAGACTT CTTT------

(n)

Figure S2. Alignment of 10 reference and 100 unknown plants marker region sequences.