*Pak. J. Bot.*, 51(1): 157-164, 2019. DOI: 10.30848/PJB2019-1(32)

# PHYLOGENETIC ANALYSIS OF *ROSA DAMASCENA* L. FROM TAIF USING DNA BARCODING APPROACH

**SHAWKAT M. AHMED**

*Biology Department, Faculty of Education, Ain Shams University, Cairo, Egypt*
*Biology Department, Faculty of Science, Ta'if University, Ta'if, Saudi Arabia*
*Corresponding author's email: shamahmoh@gmail.com*

**Abstract**

Little is known about the genome structure and the genetic origin of *Rosa damascene* L., widely distributed in Taif, Saudi Arabia for beauty and perfume production. For that, four specific barcodes; ITS2, mat*K*, rbcL and t-rnH were employed to genetically identify three *Rosa* species; *R. damascena*, *R. hybrida* and *R. damascena*Trigintipetala cultivated in Taif leading to the discrimination among them from one hand, and reviewing the previous suggestions of *R. damascena* origin from the other hand. rbcL, t-rnH and mat*K*revealed more effectiveness in sequence quality and species discrimination capability than ITS2 across the genome of *Rosa* depending on estimations of transition/transversion bias (R), Tajima relative evolutionary rate tests and the phylogenetic relationships using the Maximum Likelihood Tree method. *R. damascena* lineage exhibited accelerated evolutionary rates. The phylogenetic analyses proved the heterogeneous origin of *Rosa* cultivars from Taif doubting the validity of the previous suggestions about the genetic origin of *R. damascena*.

**Key words**: *Rosa damascena*, ITS2, mat*K*, rbcL, t-rnH, barcoding genes.

## Introduction

Roses nearly have 120 species and 20000 commercial cultivars worldwide that are popular for their beauty, medicinal purposes, decoration and perfume production (Rajapakse *et al.,* 2001). Rose resources were categorized into native species, exotic species, exotic varieties and locally developed varieties, of these classified groups only 10-20 species have developed the modern cultivars of rose (Riaz *et al.,* 2012). Despite their economic importance and for several reasons such as polyploidy, classical hybridization strategies, primitive selection, genetic mutations and insufficient funding for rose development projects, we have little information about the genome structure and subsequently the genetic origin of the three *Rosa* species; *R. damascena*, *R. hybrida* and *R. damascena* Trigintipetala cultivated in Taif. It was suggested that the summer Damask rose resulted from the recombination of *Rosa phoenicia* and *Rosa gallica* and *Rosa moschata* and *R. gallica* for the autumn Damask (Hurst, 1941), whereas Iwata *et al.,* (2000) suggested that *R. moschata, R. fedschenkoan* and *R. gallica* were the parentsforthe old Damask roses. Based on the above, different molecular strategies for studying the genetic structure and variability in roses in different regions have been employed such as isozymes, RAPD, SSR, AFLP, ISSR, SCAR and the specific DNA barcoding genes (Baydar *et al.,* 2004; Kiani *et al.,* 2008; Jabbarzadeh *et al.,* 2010; Schori & Showalter, 2011; Riaz *et al.,* 2012; Ahmed *et al.,* 2017).

The barcoding loci; ITS2, mat*K*, rbcL and t-rnH have been employed in the taxonomic and floristic studies as specific genes for plants. Their efficiency varied in quality and recovery and the discrimination ability among plant species (Tripathi *et al.,* 2013). Thus, many studies referred to the urgent need to use these loci among and within plant species to get accurate data for their genomic structure and subsequently good identification leading to valid taxonomic relationships. Taifcity inSaudi Arabiais famous for the cultivation of some *Rosa* species especially *Rosa damascena* that is denominated "Ward Taifi", *R. hybrida* and *R. damascene* Trigintipetala for beauty and production of perfumes. There is scant available knowledge about the sequence data of gene regions of *Rosa* species cultivated in Taif, Therefore, the effectiveness of four barcodes; ITS2, mat*K*, rbcL and t-rnH, were investigated for improved identification of Taif*Rosa* species and to distinguish them from other *Rosa* species retrieved from the GenBank database confirming or denying the mentioned suggestions of *R. damascena* origin.

## Materials and Methods

**Plant materials:** The three *Rosa* species, *Rosa damascena* "Ward Taifi", *R. hybrida* and *R. damascene* Trigintipetala, utilized in the investigation were selected from Taif governorate, Saudi Arabia.

**The extraction and amplification of DNA:** The DNA of fresh leaves of Taif *Rosa* species was extracted using CTAB method (Doyle & Doyle, 1987). The purified DNA was used in PCR amplification of ITS2, rbcL, matK and t-rnH loci using universal primers as mentioned in Table 1. Each reaction included 1 µl of each forward and reverse primer, 1 µl genomic DNA, 12 µl Master mix and 10 µl of distilled water. After mixing the previous components, the reaction was set up in the PCR thermal-cycler. The first step was denaturation that was performed at 94°C for 5 min then followed by 35 cycles of another denaturation at 94°C for 1 min, annealing at 52°C for 30 sec and extension at 72°C for 2 min with a final extension step of 72°C for 8 min.

**Table 1. List of primers used for the investigated DNA barcoding loci.**

| DNA locus | Primer name | | Primer sequences (5'–3') | Ann. temp. | Reference |
|---|---|---|---|---|---|
| ITS2 | ITS-S2F | F | ATGCGATACTTGGTGTGAAT | 52°C | Chen *et al*., (2010) |
| | ITS4 | R | TCCTCCGCTTATTGATATGC | | White *et al*., (1990) |
| rbcL | rbcLa | F | ATGTCACCACAAACAGAGACTAAAGC | 52°C | Levin *et al*., (2003) |
| | rbcLa | R | GTAAAATCAAGTCCACCRCG | | Kress & Erickson (2007) |
| mat*K* | matK-KIM1 | F | ACCCAGTCCATCTGGAAATCTTGGTTC | 52°C | Fazekas*et al*., (2012) |
| | matK-KIM3 | R | CGTACAGTACTTTTGTGTTTACGAG | | |
| T-rnH | psbAF | F | CGCGCATGGTGGATTCACAATCC | 52°C | Sang *et al*., (1997) |
| | t-rnH2 | R | GTTATGCATGAACGTAATGCTC | | Tate & Simpson (2003) |

**Table 2. Statistics derived from the amplification, sequencing and alignment processes for all loci.**

| Locus | ITS2 | rbcL | mat*K* | T-rnH |
|---|---|---|---|---|
| Number of Taif*Rosa* species examined | 3 | 3 | 3 | 3 |
| % PCR success | 100 | 100 | 100 | 100 |
| % Sequencing success | 100 | 100 | 66.6 | 100 |
| % Variable sites after alignment among *Rosa* species from Taif | 15 | 39 | 40 | 25 |
| Sequence length of *R. damascena* | 241 | 558 | 825 | 360 |
| Sequence length of *R. hybrida* | 347 | 549 | - | 544 |
| Sequence length of *R. d.*Trigintipetala | 330 | 531 | 813 | 380 |
| GC ratio in *R. damascena* | 61 | 44 | 34 | 24 |
| GC ratio in *R. hybrida* | 59 | 44 | - | 32 |
| GC ratio in *R. d.*Trigintipetala | 59 | 44 | 38 | 26 |
| Number of the retrieved *Rosa* species from the GenBank with lower E value | 34 | 24 | 38 | 14 |
| Mean query cover among *R. damascena* and the retrieved species | 93 | 97 | 99 | 90 |
| % Mean identity among *R. damascena* and the retrieved species | 95 | 99 | 99.5 | 91 |

**The sequencing of PCR products:** The PCR products of Taif*Rosa* species for the four barcoding loci were sent for purification and sequencing process at Macrogen Inc., South Korea.

**The sequences alignment and phylogenetic analysis:** The sequences of ITS2, rbcL, matK and t-rnH of *R. damascena* were subjected to BLAST on http://blast.ncbi.nlm.nih.gov/Blast.cgi to confirm them from the other related *Rosa* species existed on database of the GenBank. Numbers of the retrieved *Rosa* species according to each barcode locus were shown in Table 2. MUSCLE algorithm was utilized for sequence alignments that were performed using the MEGA6 program (Edgar, 2004; Tamura *et al*., 2013). Sequence alignments were adjusted manually and gaps were eliminated in the phylogenetic analysis. The parameters of equality of evolutionary rate between sequences of *R. damascena* and *R.hybrida,* with sequence of *R. damascena* Trigintipetala used as an outgroup were estimated by Tajima's relative rate test (Tajima, 1993). Nucleotide substitution rates and Transition/Transversion bias (R) were calculated by Maximum Likelihood method using MEGA6 (Tamura *et al.,* 2013). Means and levels of significance among values were determined by t-test using SPSS software (SPSS, 2011). The Maximum likelihood bootstrap (MLB) analysis were also performed for building the phylogenic trees. A total of 1000 bootstrap replicates were performed. Bootstrap support from maximum likelihood analysis ≥ 50% was detected on the branches of phylogenic trees and could be considered as good support.

## Results and Discussion

DNA barcoding loci are modern important tools for developing the genetic relationships among plant species leading to rapid accurate approaches for crops improvement especially those possessing high economic value like roses. For that, ITS2, mat*K*, rbcL and t-rnH loci were investigated to identify and distinguish between *R. damascena* and two *Rosa* species from Taif, and its putative ancestors retrieved from the GenBank database.

Although the extraction of *Rosa* DNA was difficult, due to the large amounts of secondary products, the quantity and quality of DNA were sufficient for PCR amplification and sequencing. Statistics that derived from the amplification, sequencing and alignment processes of the four loci were shown in Table 2. The efficiency of both amplification (Fig. 1) and sequencing were 100% for the four candidate loci except sequencing of *R. hybrida* did not succeed for locus mat*K* (66.6%). The sequences of *R. damascena* for ITS2, rbcL, mat*K* and t-rnH were submitted to BLAST at the GenBank database for sequence symmetry and assertion of the amplified loci. Numbers of the retrieved *Rosa* species were 34, 24, 38 and 14 for ITS2, rbcL, mat*K* and t-rnH respectively, that showed 95, 99, 99.5 and 91% similarity to sequences of *R. damascena* with lower E value (≤1e).

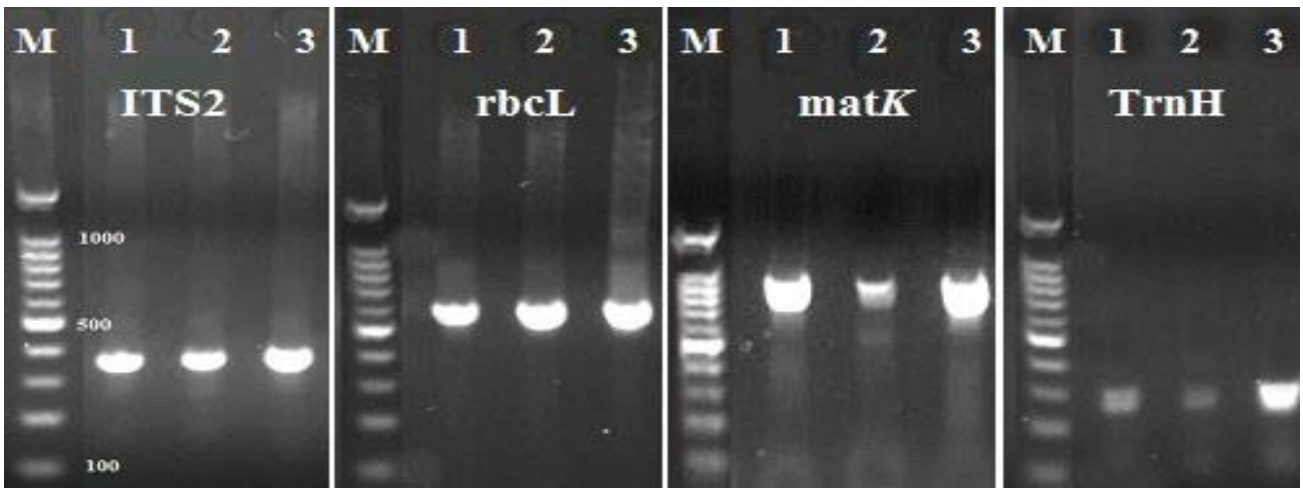


Fig. 1. The agarose gels showing PCR amplifications of ITS2, rbcL, mat*K*and t-rnHloci of Taif*Rosa* species; 1- *R. damascena*, 2- *R. hybrida*, 3- *R. damascena*Trigintipetala. M- Marker.

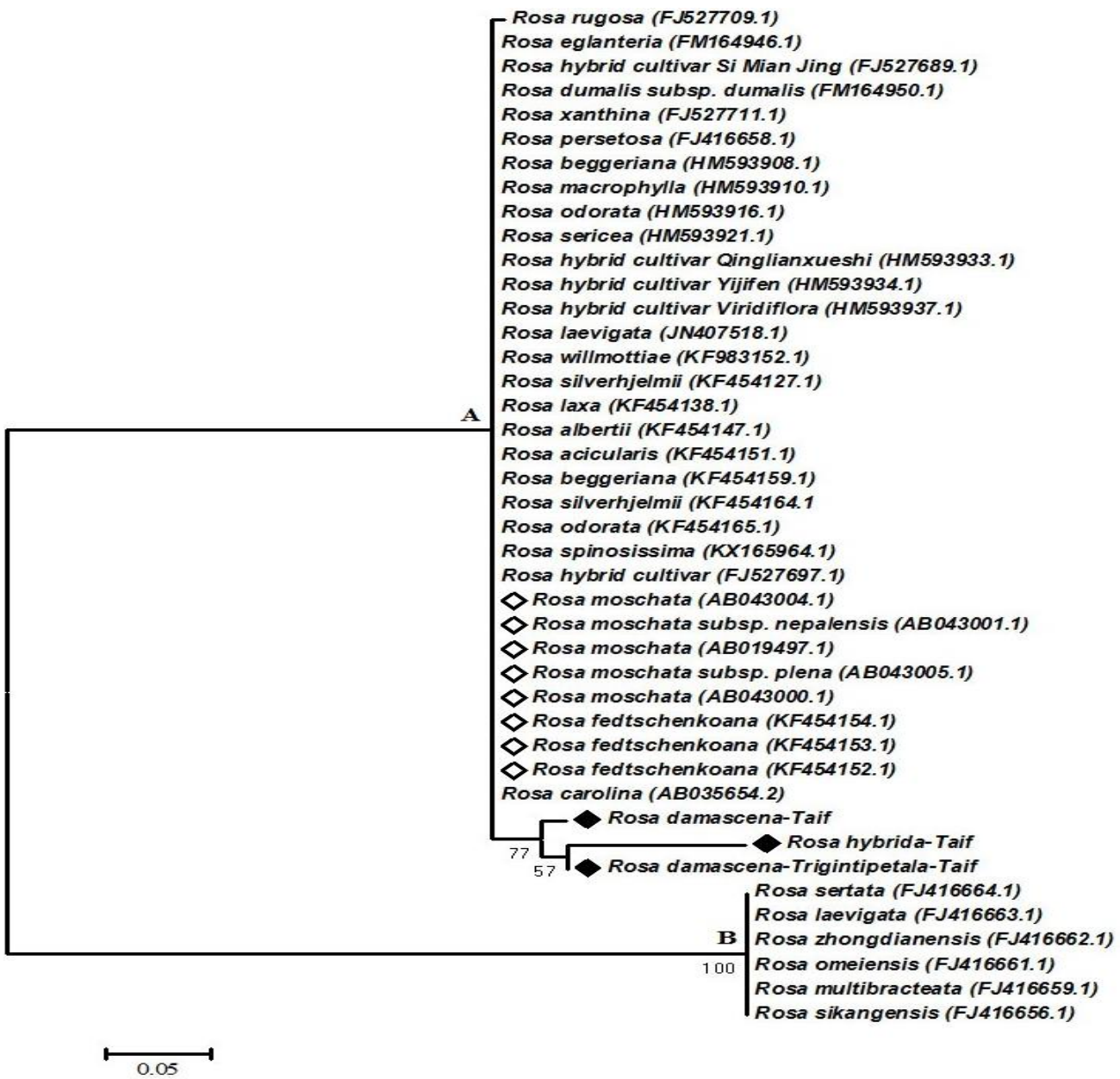


Fig. 2. Phylogram based on ITS2 locus for Taif roses and the retrieved species using Maximum Likelihood Tree method.

**Table 3. Mean nucleotide substitution rates and transition/transversion bias (R) in the four loci for *Rosa* species calculated by Maximum Likelihood method.**

| Species | Locus | Transition | | | Transversion | | | | | | (R) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | A→G | G→A | T→C | C→T | A→T | T→A | A→C | C→A | T→G | G→T | C→G | G→C | |
| Overall | ITS2 | 11.7 | 6.5 | 10.8 | 5.7 | 5.7 | 5.8 | 10.9 | 5.8 | 10.3 | 5.7 | 10.3 | 10.9 | 0.45 |
| | rbcL | 15.9 | 18.9 | 13.2 | 18.9 | 4.5 | 4.5 | 3.4 | 4.5 | 3.8 | 4.9 | 3.8 | 3.4 | 1.96 |
| | T-rnH | 6.5 | 16.0 | 7.3 | 29.0 | 7.9 | 7.6 | 2.0 | 7.6 | 3.1 | 7.9 | 3.1 | 2.0 | 0.99 |
| | mat*K* | 9.4 | 16.7 | 10.2 | 21.1 | 7.5 | 6.6 | 3.6 | 6.6 | 3.7 | 7.5 | 3.7 | 3.6 | 1.17 |
| | Mean | 10.9* | 14.5* | 10.4** | 18.7* | 6.4** | 6.1** | 5.0 | 6.1** | 5.2 | 6.5** | 5.2 | 5.0 | 1.14* |
| Taif cultivars | ITS2 | 3.4 | 2.1 | 5.0 | 3.0 | 8.7 | 7.6 | 14.3 | 7.6 | 12.7 | 8.7 | 12.7 | 14.3 | 0.14 |
| | rbcL | 16.3 | 19.6 | 13.7 | 18.1 | 4.6 | 4.5 | 3.5 | 4.5 | 3.7 | 4.6 | 3.7 | 3.5 | 2.04 |
| | T-rnH | 7.8 | 22.1 | 9.3 | 21.9 | 6.8 | 7.3 | 2.9 | 7.3 | 2.6 | 6.8 | 2.6 | 2.9 | 1.07 |
| | mat*K* | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Mean | 6.6 | 10.9 | 4.4 | 10.0 | 6.7* | 6.5* | 6.9 | 6.5* | 6.3 | 6.7* | 6.3 | 6.9 | 1.1 |
| *R. damascena*a and its supposed ancestors | ITS2 | 7.8 | 4.6 | 6.7 | 4.2 | 7.8 | 6.7 | 12.4 | 6.7 | 11.4 | 7.8 | 11.4 | 12.4 | 0.29 |
| | rbcL | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | T-rnH | 6.6 | 16.6 | 7.6 | 31.0 | 7.7 | 6.8 | 1.9 | 6.8 | 2.7 | 7.7 | 2.7 | 16.6 | 1.09 |
| | mat*K* | 9.0 | 16.4 | 10.1 | 7.0 | 7.5 | 7.0 | 3.8 | 7.0 | 3.9 | 7.5 | 3.9 | 3.8 | 1.08 |
| | Mean | 7.8** | 12.5 | 8.1* | 14.1 | 7.7** | 6.8** | 6.0 | 6.8** | 6.0 | 7.7** | 6.0 | 10.9 | 0.82 |

Significance:* ≤ 0.05, ** ≤ 0.01

After alignment process, the sequence length range (from 544 to 360 bp) of t-rnH was higher than that of ITS2 (from 347 to 241 bp) and rbcL (from 558 to 531 bp), whereas, the percentages of variable sites were higher in mat*K* (40%) and rbcL (39%) than those of ITS2 and t-rnH among *Rosa* species from Taif. The GC ratio scored in locus ITS2 was the highest and followed by those of rbcL, mat*K*and t-rnH loci as shown in Table 2. Knowledge of the rates of transitions to transversions for species sequences after alignment is important to explain the substitution mutations in plant genomes for studying the molecular evolution of these sequences of species under consideration (Kimura, 1981; Gojobori *et al*., 1982). Table 3 revealed that transitions occurred in general more than transversions. This is because the process of transition may be more easy than that of transversion supporting the point of view that nucleotide substitution bias prefer transitions more than transversions and some nucleotides mutate greater than others (Zhao & Boerwinkle, 2002; Zhang & Gerstein, 2003). It was also noticed that transition/transversion bias (R) of rbcL, t-rnH and mat*K* was higher than that of ITS2. Luo *et al.,* (2016) mentioned that the occurrence of transition–transversion bias differs based on genome region. Moreover, the Tajima relative evolutionary rate tests (Tajima, 1993) were estimated to investigate whether *R. damascena* and its relative species evolve at the same rate for the four specific loci (Table 4). The test showed that the evolutionary rates were not constant in *R. damascena* and its relative species. In spite of using different species as outgruops, *R. damascena* lineage exhibited accelerated evolutionary rates in all loci. Except in case of *R. hybrida* as an outgroup, *R. damascena* and *R. damascena* Trigintipetala showed a similar evolutionary rates because the *P*-values were higher than 0.05 in ITS2, rbcL and t-rnH loci. The null hypothesis of equal evolution rates between *R. damascena* and its supposed ancestors; *R. moschata*, *R. gallica* and *R. fedschenkoana*, was rejected because the *P*-values were lower than 0.05 in ITS2, rbcL and mat*K*r evealing the accelerated evolutionary rate of *R. damascena* and subsequently reflecting the variance among them. The present results revealed that rbcL, t-rnH and mat*K*were more effective loci in sequence quality and also in species discrimination capability than ITS2 (Kress & Erickson, 2007; Piredda *et al*., 2011; De Vere *et al*., 2012) across the genome of *Rosa*.

For further clarification of the phylogenetic relationships of *Rosa* species under study, sequence data of the four loci was used to reconstruct the phylogenetic trees using the Maximum Likelihood Tree method (Figs 2, 3, 4 and 5). These loci have been used in many species for the development of the phylogenetic relationships due to their great discriminatory power (Dong *et al*., 2014; Shinwari *et al*., 2014; Al-Hemaid *et al*., 2015; Bolson *et al*., 2015; Tang *et al*., 2016; Khan *et al*., 2016; Zahra *et al*., 2016; Shinwari *et al*., 2018). Figure 2 demonstrated the phylogeny tree resulting from ITS2 data that split the *Rosa* species into two groups (A and B). Two subgroups were formed within group A showed a low level of diversity within the group. The

first subgroup had the putative ancestors; *R. moschata* and *R. fedtschenkoana,* of *R. damascena* within thirty three retrieved species from the GenBank, whereas, the three *Rosa* species from Taif were placed in the second subgroup showing low interspecific divergence among them. The phylogenetic tree derived from rbcL data categorized the three roses from Taif into single species, *R. hybrid* formed a separate operational taxonomic units (OTU) in cluster D and *R. damascena* and *R. damascena* Trigintipetala into two separate clades within cluster C (Fig. 3). The tree showed high divergence among the three species. Sequences of ancestors of *R. damascena* for rbcL gene were not represented in the GenBank database. t-rnH and mat*K* trees showed complete divergence between *R. damascena* and its putative ancestors; *R. moschata* and *R. gallica*, into two distinct groups (Figs. 4 and 5). On the contraryt-rnH tree which grouped *R. damascena* and *R. damascena* Trigintipetala together in one group (F), the mat*K* tree separated the two species into two different subgroups within group H. It could be concluded that, incontrastwith ITS2 tree, the phylogenetic trees depending upon rbcL, t-rnH and mat*K* sequencing data revealed clear discrimination among the three *Rosa* species cultivated in Taif, on one hand, and between *R. damascena*and its supposed ancestors on the other hand. These results confirmed those of Ahmed *et al.,* (2017) depending on SSR and ISSR data. However, they were not in accordance with those of Iwata *et al*., (2000) who suggested that *R. moschata*, *R. gallica*and *R. fedschenkoana* were the parental species of Damask varieties depending on ITS and *psb*A-*t-rnH* sequences. This means that their suggestion may be uncertain and needs further studies or that *R. damascena*cultivated in Taif was originated from a hybridization process from other parents. The phylogenetic studies of the present *Rosa* cultivars are difficult and complex due to their heterogeneous origin resulting from free pollination in nature or the traditional strategies which permit the hybridization among cultivars or hybrids without declaring their parental form (Lidia & Irina, 2009). This confirmed the point of view that it is preferable to sequence more specific loci to obtain accurate results instead of one gene which is less reproducible to reconstruct the tree of evolution (Sandersson&Driskell, 2003). The obvious divergence of *R. damascena*from other species of *Rosa* may return to the geographic isolation, the traditional selection for the commercial traits and the adaptation to the environmental conditions in Taif through some mutational events.

**Table 4. Tajima relative rate tests of the four loci for *Rosa* species.**

| Locus | Outgroup | Testing group | | RI | RD | RA | RB | χ2 | *P* value |
|---|---|---|---|---|---|---|---|---|---|
| | | (A) | (B) | | | | | | |
| ITS2 | *R. d.*Trigintipetala | *R. damascena* | *R. hybrida* | 201 | 3 | 7 | 19 | 5.54 | <0.05 |
| | *R. hybrida* | *R. damascena* | *R. d.*Trigintipetala | 201 | 3 | 7 | 7 | 0.00 | >0.05 |
| | *R. fedtschenkoana* | *R. damascena* | *R. moschata* | 209 | 0 | 8 | 0 | 8.00 | <0.05 |
| | *R. moschata* | *R. damascena* | *R. fedtschenkoana* | 209 | 0 | 8 | 1 | 5.44 | <0.05 |
| rbcL | *R. d.*Trigintipetala | *R. damascena* | *R. hybrida* | 249 | 0 | 1 | 200 | 197 | <0.05 |
| | *R. hybrida* | *R. damascena* | *R. d.*Trigintipetala | 249 | 0 | 1 | 2 | 0.33 | >0.05 |
| | *R.gallica* | *R. damascena* | *R. moschata* | 154 | 2 | 115 | 0 | 115 | <0.05 |
| | *R. moschata* | *R. damascena* | *R. gallica* | 154 | 2 | 115 | 1 | 112 | <0.05 |
| T-rnH | *R. d.*Trigintipetala | *R. damascena* | *R. hybrida* | 138 | 4 | 4 | 95 | 83.7 | <0.05 |
| | *R. hybrida* | *R. damascena* | *R. d.*Trigintipetala | 138 | 4 | 4 | 2 | 0.67 | >0.05 |
| mat*K* | *R. gallica* | *R. damascena* | *R. moschata* | 417 | 0 | 276 | 0 | 276 | <0.05 |
| | *R. moschata* | *R. damascena* | *R. gallica* | 417 | 0 | 276 | 1 | 273 | <0.05 |

The Tajima relative rate test was used to examine the equality of evolutionary rate between *R.damascena* and other relative species with different outgorups
RI is the identical sites in all three sequences
RD is the divergent sites in all three sequences
RA is the number of unique differences in the sequence A
RB is the number of unique differences in the sequence B
χ2 test statistic more than 3.841 ($p$<0.05) indicates accelerated evolution
*P* value greater than 0.05 is often used to accept the null hypothesis of equal rates between lineages

## Conclusion

This study is a pioneering attempt at producing DNA sequences from ITS2, mat*K*, rbcL and t-rnH barcoding loci in three *Rosa* species cultivated in Taif. mat*K*, rbcL and t-rnH were more applicable barcodes than ITS2 for the authentication and diversification of *Rosa* species. Like animal genomes, the plant genomes showed higher transition rates than the transversion rates. Mutation accumulation, little gene flow and geographic isolation encouraged *R. damascena*to gain accelerated evolutionary rates and subsequently considerable divergence from other related species to genus *Rosa*.

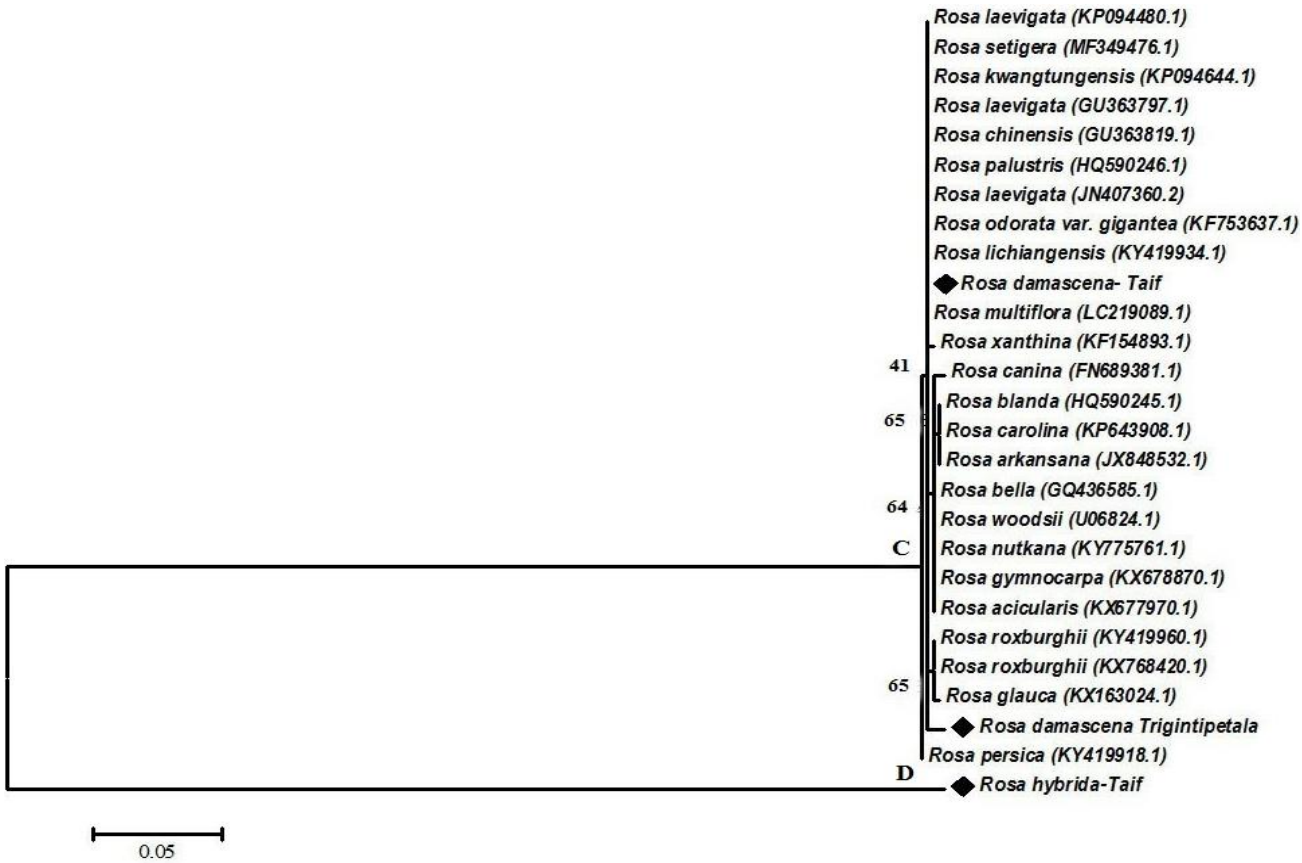


Fig. 3. Phylogram based on rbcL locus for Taif roses and the retrieved species using Maximum Likelihood Tree method.

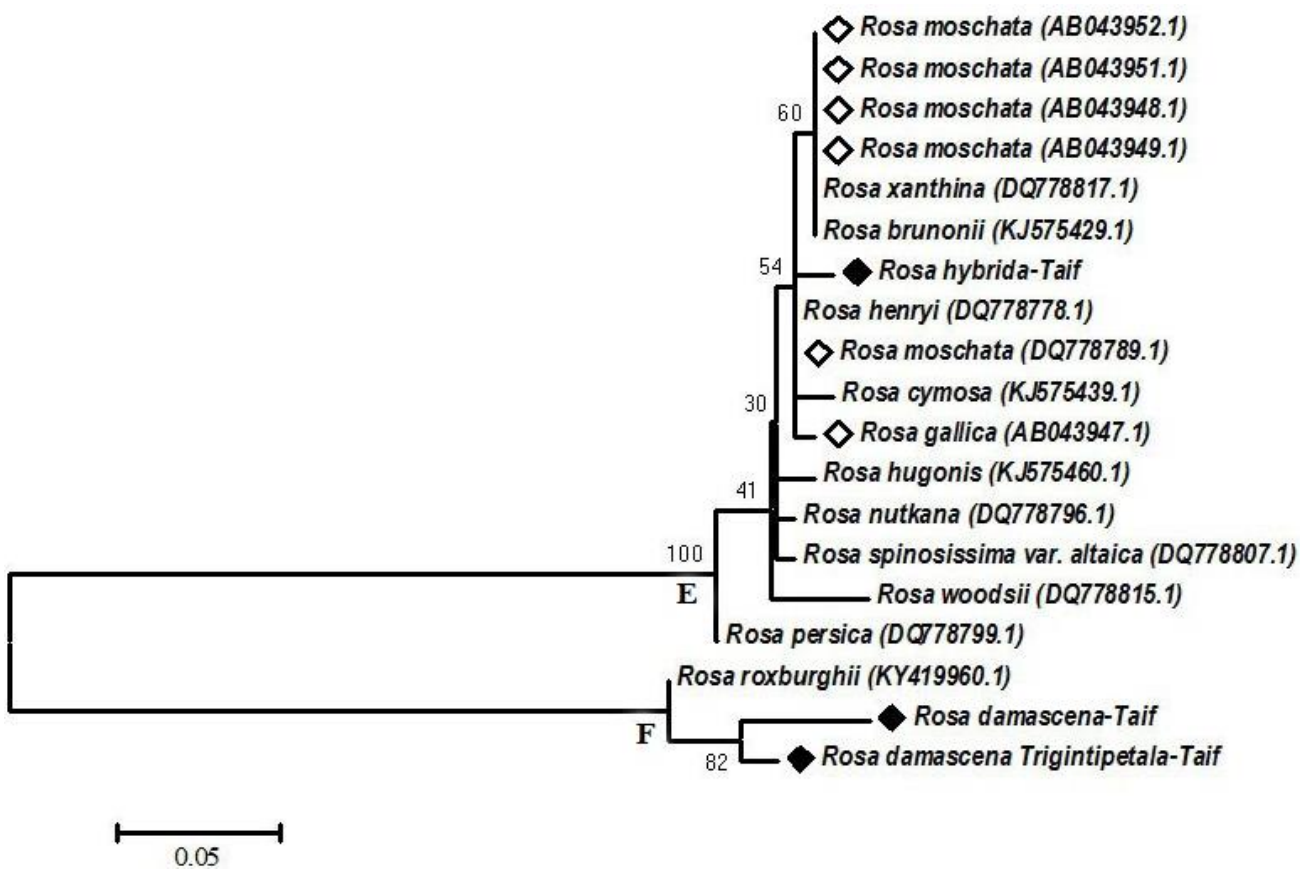


Fig. 4. Phylogram based on t-rnH locus for Taif roses and the retrieved species using Maximum Likelihood Tree method.

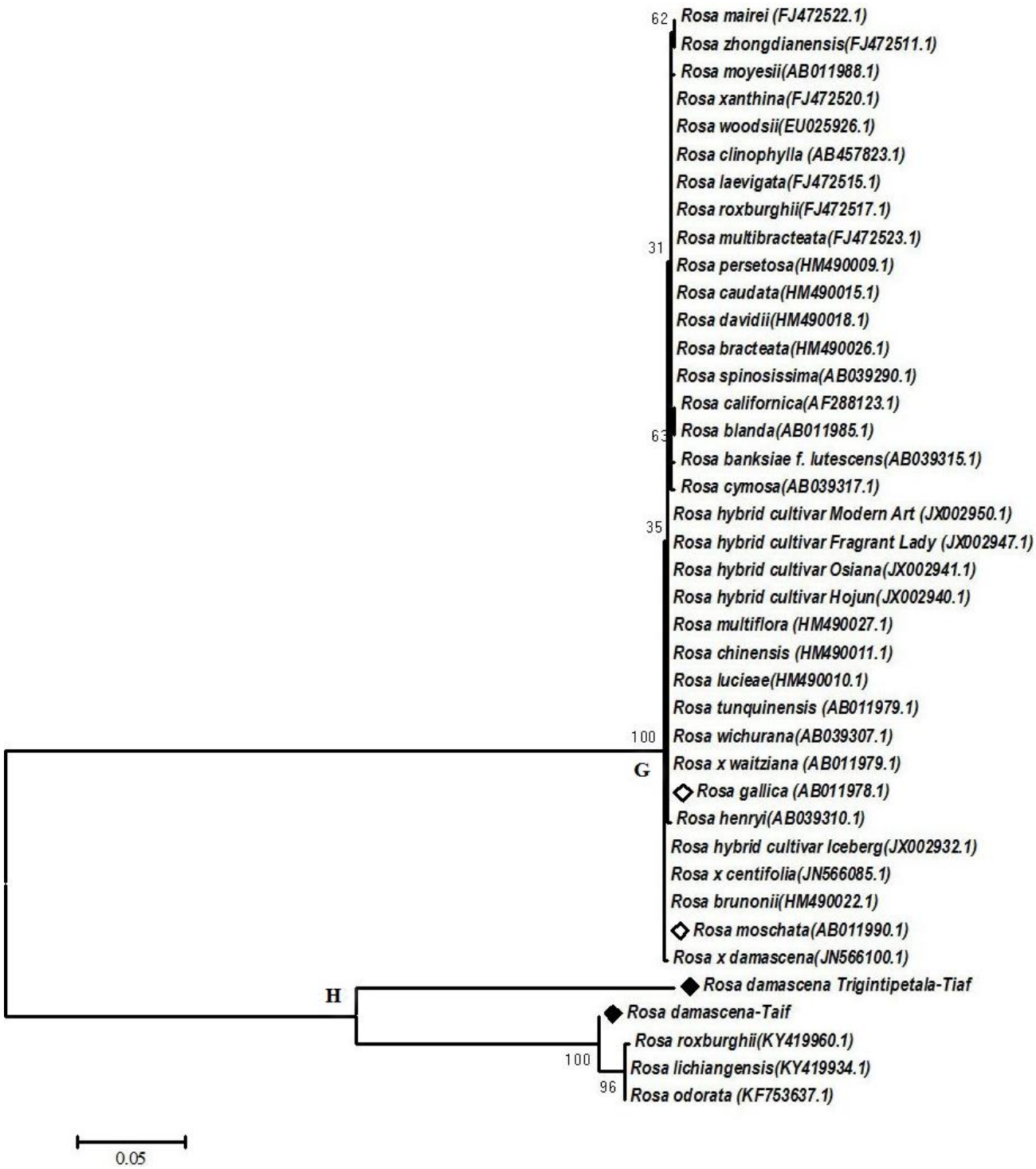


Fig. 5. Phylogram based on mat*K* locus for Taif roses and the retrieved species using Maximum Likelihood Tree method.

**Acknowledgement**

The author is grateful to the chair of research and development studies for Taif rose, Taif University, Kingdom of Saudi Arabia, for supporting this work.

**References**

Ahmed, S.M., H.Y. Darwish and K.H. Alamer. 2017. Microsatellite, inter simple sequence repeat and biochemical analyses of *Rosa* genotypes from Saudi Arabia. *Afr. J. Biotech*., 16(12): 552-557.

Al-Hemaid, F.M.A., M.A. Ali, J. Lee, S. Kim and M. Oliur Rahman. 2015. Molecular evolutionary relationships of *Euphorbia scordifolia*jacq. within the genus inferred from analysis of internal transcribed spacer sequences. *Bangl. J. Plant Taxon*., 22(2): 111-118.

Baydar, N.G., H. Baydar and T. Debener. 2004. Analysis of genetic relationships among *Rosa damascena* plants grown in Turkey by using AFLP and microsatellite markers. *J. Biotech*., 11: 1263-267.

Bolson, M., E.C. Smidt, M.L. Brotto and V. Silva-Pereira. 2015. ITS and t-rnH-psbA as efficient DNA barcodes to identify threatened commercial woody angiosperms from southern Brazilian atlantic rainforests. *PLoS One*, 10(12): e0143049.

Chen, S., H. Yao, J. Han, C. Liu, J. Song and L. Shi. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One*, 5(1): e8613.

De Vere, N., T.C. Rich, C.R. Ford, S.A. Trinder and C. Long. 2012. DNA barcoding the native flowering plants and conifers of Wales. *PLoS One,* 7: e37945.

Dong, W., T. Cheng, C. Li, C. Xu, P. Long, C. Chen and S. Zhou. 2014. Discriminating plants using the DNA barcode rbclb: an appraisal based on a large data set. *Mol. Ecol. Resour*., 14: 336-343.

Doyle, J.J. and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull*., 19: 11-15.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*., 32(5): 1792-1997.

Fazekas, A.J., M.L. Kuzmina, S.G. Newmaster and P.M. Hollingsworth. 2012. DNA barcoding methods for land plants. *Methods Mol Biol.*, 858: 223-252.

Gojobori, T., W.H. Li and D. Graur. 1982. Patterns of nucleotide substitution in pseudogenes and functional genes. *J. Mol. Evol.,* 18: 360-369.

Hurst, C.C. 1941. Notes on the origin and evolution of our garden roses. *J. Roy. Hort. Soc.,* 66: 73-289.

Iwata, H., T. Kato and S. Ohno. 2000. Triparental origin of Damask roses. *Gene*, 259: 53-59.

Jabbarzadeh, Z., M. Khosh-khui, H. Salehi and A. Saberivand. 2010. Inter simple sequence repeat (ISSR) markers as reproducible and specific tools for genetic diversity analysis of rose species. *Afr. J. Biotech*., 9(37): 6091-6095.

Khan, M.Q., A.T. Khalil and Z.K. Shinwari. 2016. Searching for DNA Barcodes in Plants. American-*Eurasian J. Agri. & Environ. Sci.,* 15(4): 504-513.

Kiani, M, Z. Zamani, A. Khalighi, R. Fatahi and D.H. Byrne. 2008. Wide genetic diversity of *Rosa damascena*Mill. germplasm in Iran as revealed by RAPD analysis. *Sci. Hort*., 115: 386-392.

Kimura, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci.USA*, 78: 454-458.

Kress, W.J. and D.L. Erickson. 2007. A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding t-rnH-psbA spacer region. *PLoS One,* 2: e508.

Levin, R.A., W.L. Wagner, P.C. Hoch, M. Nepokroeff, J.C. Pires and E.A. Zimmer. 2003. Family-level relationships of Onagraceae based on chloroplast rbcL and ndhF data. *Amer. J. Bot.*, 90(1): 107-15.

Lidia, A. and S. Irina. 2009. Theoretical considerations upon the origin and nomenclature of the present rose cultivars. *J. Plant Develop.,* 16: 103-108.

Luo, G., X. Li, Z. Han, Z. Zhang, Q. Yang, H. Guo and J. Fang. 2016. Transition and transversion mutations are biased towards GC in transposons of *Chilosuppressalis* (Lepidoptera: Pyralidae). *Genes*, 7: 72.

Piredda. R.S.M., M. Attimonelli, R. Bellarosa and B. Schirone. 2011. Prospects of barcoding the Italian wild dendroflora: oaks reveal severe limitations to tracking species identity. *Mol. Ecol. Resour*., 11: 72-83.

Rajapakse S., D.H. Byrne, L. Zhang, N. Anderson, K. Arumuganathan and R.E. Ballard. 2001. Two genetic linkage maps of tetraploid roses. *Theor. Appl. Genet*., 103: 575-583.

Riaz S., B. Sadia, F.S. Awan, I.A. Khan, H.A. Sadaqat and I.A. Khan. 2012. Development of a species-specific sequence-characterized amplified region marker for roses. *Genet. Mol. Res*., 11(1): 440-447.

Sanderson, M.J. and A.C. Driskell. 2003. The challenge of constructing large phylogenetic trees. *Tr. Plant Sci.,* 8: 374-379.

Sang, T., D. Crawford and T. Stuessy. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *Amer. J. Bot.,* 84(9): 1120-1136.

Schori, M. and A.M. Showalter. 2011. DNA barcoding as a means for identifying medicinal plants of Pakistan. *Pak. J. Bot*., 43: 1-4.

Shinwari Z.K., S.A. Jan, A.T. Khalil, A. Khan, M. Ali, M. Qaiser and N.B. Zahra. 2018. Identification and phylogenetic analysis of selected medicinal plant species from Pakistan: DNA barcoding approach. *Pak. J. Bot.*, 50(2): 553-560.

SPSS. 2011. "*SPSS Statistics for Windows*". Version 20.0. Chicago, IL, USA, SPSS Inc.

Tajima, F. 1993. Simple methods for testing molecular clock hypothesis. *Genetics*, 135: 599-607.

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*., 30: 2725-2729.

Tang, Y., Y. Wu, R. Huang, N. Chao, Y. Liu, P. Xu, K. Li, D. Cai and Y. Luo. 2016. Molecular identification of *Uncaria* (Gouteng) through DNA barcoding. *Chin. Med.,* 11: 3.

Tate, J.A. and B.B. Simpson. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *System Bot.,* 28(4): 723-37.

Tripathi, A.M., A. Tyagi, A. Kumar, A. Singh and S. Singh. 2013. The Internal transcribed spacer (ITS) region and t-rnHh-psba are suitable candidate loci for DNA Barcoding of tropical tree species of India. *PLoS One,* 8(2): 1-11.

Wen, X.P. and X.X. Pang. 2004. Characterization of genetic relationships of Rosa roxburghiiTratt and its relatives using morphological traits, RAPD and AFLP markers. *Sci. Hortic.*, 79: 189-196.

White, T.J., T. Bruns, S. Lee and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: Guide Methods Appl.*, 18: 315-322.

Zabta K.S., K. Jamil and N. B. Zahra. 2014. Molecular systematics of selected genera of family Fabaceae.*Pak. J. Bot*., 46(2): 591-598.

Zahra, N.B., Z.K. Shinwari and M. Qaiser. 2016. DNA Barcoding: A tool for standardization of herbal medicinal products (HMPs) of Lamiaceae from Pakistan. *Pak. J. Bot*., 48(5): 2167-2174.

Zhang, Z. and M. Gerstein. 2003. Patterns of nucleotide substitution, insertion and deletion in the human genome inferred from pseudogenes. *Nucleic Acids Res.,* 31: 5338-5348.

Zhao, Z. and E. Boerwinkle. 2002. Neighboring-nucleotide effects on single nucleotide polymorphisms: A study of 2.6 million polymorphisms across the human genome. *Genome Res.,* 12: 1679-1686.

(Received for publication 15 November 2018)