Chapter

# Genetic Resources of The Universal Flavor, Vanilla

*Minoo Divakaran and N.T. Fathima Rafieah*

## Abstract

Commercially cultivated vanilla (*V. planifolia*) is native to Mexico and its cultivation and breeding programmes face major bottlenecks. This study reports presence of important agronomic characters in two important and endangered species of Vanilla, *V. aphylla* and *V. pilifera*, indigenous to India. *V. aphylla* was tolerant to Fusarium wilt and had longer flower life than the cultivated vanilla. *V. pilifera* flowers were fragrant, showed signs of insect pollination and had large fruit size. The species were amenable to interspecific hybridization and successful reciprocal crosses were done. Sequence similarity studies indicated the clustering of leafy and leafless species separately.

**Keywords:** interspecific hybridization, *V. apylla*, *V. pilifera*, sequence similarity

## 1. Introduction

The genus Vanilla includes about 110 species and the species have been treated in various monographic works [1, 2] including the life history of *V. planifolia* [3]. *Vanilla planifolia* (Salisb.) Ames (syn. *V. fragrans* Andrews.), is a tropical climbing orchid known for yielding the delicate popular flavor, vanilla [4] and is the second most expensive spice traded in the world market [5] (Spices Board 2000). The major vanilla producing countries are Madagascar, Comoro, Indonesia, Mexico and the Reunion, of which, Madagascar holds the prominent position.

Vanilla was introduced to Europe from Mexico, in about 1500 and its reputation of being an aphrodisiac followed it to countries where it was introduced. The importance of vanilla since early times in Mexico, is evident by the mention of offering vanilla as a medicinal beverage as part of a tribute during reign of Itzcóatl (Aztec Emperor) in 1427 and citing vanilla as a remedy for fatigue in Badianus manuscript in 1552 [6]. *Vanilla planifolia*, which yields the vanilla of commerce, is native to Mexico and parts of Central America and the history of origin of cultivated vanilla suggests that the entire stock outside Mexico may be from a single genetic source. For the last 400 years, humans have been playing important role in the dispersal and spread of vanilla in the New World.

## 2. Species of Vanilla

Studies of divergence among species of agronomic importance have been receiving greater attention. Genomics-based tools are efficient to characterize and

identify genetic diversity in Vanilla and act as a significant tool for genomics-assisted plant breeding [7]. RAPD polymorphism was used to estimate the level of genetic diversity and interrelationships among few related species *Vanilla planifolia*, including both leafy and leafless types such as *V. tahitensis*, *V. andamanica*, *V. pilifera* and *V. aphylla*. Studies revealed that there is very limited variation within collections of *V. planifolia*, indicative of its narrow genetic base [8]. The British introduced *Vanilla planifolia* into India about 200 years ago whereas five other species are native viz., *V. pilifera* Holt., *V. andamanica* Rolfe., *V. aphylla* Blume., *V. walkeriae* Wight. and *V. wightiana* Lindl.

*V. pilifera* originally described from Malaya, recorded in Thailand is found in the Mikir hills of Northeast India. *V. aphylla*, an endangered species, previously known from Thailand is found in South India [9]. *V. andamanica* is endemic to Andaman group of Islands and is believed to be same as *V. albida* [10]. *V. tahitensis*, which is commonly exploited throughout the tropics, is indigenous to the Tahiti Islands. The presence and absence of leaves, and floral characters (colors of flower, lip, hairs on lip and ovary-pedicel etc.,), morphologically distinguish these species (**Table 1**).

A preliminary analysis of the various characters of Vanilla species including the above species, showed presence and absence of leaves formed an important part in the classification of the genus which in general had the basic chromosome number x = 16. Most of the Indian species were leafless, except *V. pilifera* which was intermediate in character, i.e., leafless in early stages and long narrow leaves at maturity and the chromosome number in *V. aphylla* is 2n = 64, whereas the cultivated vanilla and *V. tahitensis* had a somatic chromosome number of 2n = 32 [10, 11]. Differences in floral characters existed in flower color and lip characters (**Table 2**). In *V. pilifera* vines, leaves developed as the vine grew with flowers that were narrower (2.8 x 0.8 cm) with distinct pure white ovary-pedicel (**Figures 1** and **2**), pale green tepals, purplish violet and longer (6 mm approx.) hairs on white lip. *V. aphylla* is leafless (with scales-1.8 cm) and yellowish-cream flowers (petal size 3 x 1.2 cm approx.) having tuft of hairs that are cream near tip, deep reddish inside (2–3 mm) and light green ovary-pedicel (**Figures 3** and **4**).

| Species | Leaf type | | Internode | Median ridge |
|---|---|---|---|---|
| | Shape | Size | | |
| *V. aphylla* | Scale leaves to leafless | 2.1 cm in fresh shoots | 6.5 cm | Absent in fresh shoots |
| *V. pilifera* | Narrow (intermediate to *V. planifolia* and *V. aphylla*) | L (8.5–16.5 cm) B (1.6–3.1 cm) | 7 cm | Present all along the stem |

**Table 1.**
*Vegetative characters of* Vanilla aphylla *and* V. pilifera.

| Species | Petal color | Ovary-pedicel | Tuft of hair on the lip | Nature | Fruit size after pollination – 4 weeks |
|---|---|---|---|---|---|
| *V. aphylla* | Yellowish cream, L-3 cm, B-1.2 cm | Light green | Cream near tip, reddish brown inside (2–3 mm) | Longer life | 14 cm (L) 3 cm (B) |
| *V. pilifera* | Pale green, narrower, L-2.8 cm, B-0.8 cm | White | Violet and longer (6 mm) | More brittle | 11.5 cm (L) 3.3 cm (B) |

*L, Length; B, Breadth.*

**Table 2.**
*Variations in floral characters.*

**Figure 1.**
*Members of* V. aphylla *inflorescence arranged sequentially (Inset: Close-up of an opened flower).*

**Figure 2.**
*Members of V. pilifera inflorescence arranged sequentially (Inset: Close-up of an opened flower).*

**Figure 3.**
*L.S. of flowers of* V. aphylla *(L) and* V. pilifera *(R).*

**Figure 4.**
*Comparison of dissected out flowers of* V. aphylla *(L) and* V. pilifera *(R).*

## 3. Biotechnological applications

Micropropagation and *in vitro* conservation techniques for the different species of Vanilla [12] and interspecific hybridization as a tool for gene flow of desirable characters from wild species into cultivated species, through pollen, have been reported [13]. Genetic interrelationships studies, using RAPD profiles [8], among different species revealed that the leafless forms of vanilla, *V. aphylla* and *V. pilifera* formed a separate sub-cluster. All the other leafy vanilla types formed a separate sub-cluster. *V. pilifera*, which showed an intermediate leaf character, showed only 50–56.1% similarity to *V. planifolia* but closely resembled *V. aphylla* (76.8%). Thus, the present study reveals the presence of important agronomic characters for introgression into cultivated vanilla and which can be utilized to overcome major bottlenecks in vanilla breeding. The presence of fragrance which attracts insects, coupled with signs of fruit set without hand pollination, holds *V. pilifera* as a potential candidate for breeding programmes, to overcome the problem of lack of natural seed set in vanilla. *V. aphylla* which was tolerant to *Fusarium oxysporum* [8] and its crossability to cultivated vanilla can be utilized as a bridging species and to help wipe out diseases arising out of monoculture. Interspecific hybridization has been reported and hence transfer of these desirable traits into cultivated vanilla,

*V. planifolia*, may not be hindered. The advent of biotechnological tools, offers techniques for transfer of these characters into *V. planifolia*, thus making the dream of transforming vanilla into a fragrant, natural seed setting, disease tolerant commercially important orchid can be turned into a reality.

The identification of a hydratase/lyase type enzyme as being a vanillin synthase offers new opportunities for the Vanilla pod-based industries. The accumulation of vanillin glucoside in the capsules of cultivated vines in response to environmental challenges may now be assessed at the molecular level. Likewise, the basis for development of genetic markers for the selection of vanilla orchid varieties with improved aromatic properties has now been laid down. Vanillin produced biologically is termed 'natural' vanillin and has a high economic value compared with chemically synthesized vanillin. Likewise, in the transition towards a bio-based economy, it is important to develop sustainable production systems to replace those currently based on fossil fuels. The demonstration that a single enzyme in the vanilla pod catalyzes the conversion of ferulic acid and ferulic acid glucoside into vanillin and vanillin glucoside provides several options for biotechnological applications [14].

## 4. Materials and methods

### 4.1 Genomic DNA isolation

Genomic DNA was isolated from approximately 100 mg fresh leaves by grinding in a pestle and mortar using liquid Nitrogen and following the procedure using DNeasy® Plant Mini Kit (Qiagen, USA). The ground sample powder (100 mg) was transferred to microfuge tubes. Followed by addition of 400 μl AP1 buffer and 4 μl RNase A and mixed by vortex. The tubes were incubated at 65°C for 10 min in a water bath with intermittent mixing 2–3 times by inverting the tubes. Added 130 μl buffer P3 to the tube, mixed and incubated for 5 min on ice. The lysate was centrifuged for 5 min at 14,000 rpm. The samples were then loaded onto the QIAshredder spin columns and centrifuged at 14,000 rmp for 2 min. The flow-through was transferred to a new tube without disturbing the pellet. Added 1.5 volume of buffer AW1 and mixed by pipetting. The contents were then loaded in 650 μl fractions onto the DNeasy mini spin column and centrifuged at 8000 rmp for 1 minute. The flow-through was discarded. The spin column was placed into a new 2 ml collection tube and added 500 μl buffer AW2, followed by centrifugation for 1 min at 8000 rpm. This last step with buffer AW2 step was repeated, with centrifugation at 14,000 rpm for 2 min. The spin columns were placed in fresh microfuge tubes and 100 μl AE buffer was added onto the membranes and incubated at room temperature for 5 min. The tubes were then centrifuged at 8000 rpm for 1 min. This step was repeated with another 100 μl of AE buffer. The eluted samples were stored at −20°C.

### 4.2 Measurement of purity and DNA concentration

Quality and quantity of genomic DNA was monitored by using UV/Vis. Spectrophotometry and quality was confirmed by using 0.8% Agarose Gel Electrophoresis. Each of the sample DNA was diluted to 5 ng/μL in double distilled water for use as a PCR template.

### 4.3 PCR amplification of DNA barcoding region and sequencing

PCR reactions were carried out using universal primers for the DNA barcode regions matK, nrDNA-ITS, rbcL and trnH-psbA. All the specific locus primers were

purchased with universal M13 primer sequence at their 5′ ends, thus enabling the direct sequencing of the PCR products using the universal M13 primers. The PCR amplification was performed in a 20 μl reaction mixture, consisting of 1X PCR buffer (2 mM Mgcl2), 200 μM each of dATP, dCTP, dGTP, dTTP; 0.5 μM of each forward and reverse primers and 1 U of Taq polymerase (TakaRa-Taq), and (5–20 ng) DNA template. DNA amplification was performed in a thermal cycler (Eppendoff, Germany). When the reaction has finished, the tubes were stored at 4°C. PCR products were separated by agarose gel electrophoresis (1.8%). The list of primers, their nucleotide sequences, annealing temperature and the specific PCR cycling conditions are shown in **Table 3**. A large volume PCR reaction (100 μl) per sample loci was done and PCR purification was done using Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel, Germany). The purified PCR products were sequenced using M13 universal primers (M13 forward and M13 reverse primers) on ABI 3730xl

| S. No | DNA region | Primers Name | Sequence (5′-3′) | Reference | PCR reaction conditions |
|---|---|---|---|---|---|
| 1 | *matK* | 390F | CGATCTATTCATTCAATATTTC | [15] | 95°C for 4 min.<br>95°C for 45 sec.<br>48°C for 30 sec<br>72°C for 50 sec.<br>72°C for 8 min. } 35X |
| | | 1326R | TCTAGCACACGAAAGTCGAAGT | | |
| 3 | ITS | ITS4 | TCCTCCGCTTATTGATATGC | [16] | 94°C for 4 min.<br>94°C for 40 sec.<br>55 for 40 sec.<br>72°C for 1 min.<br>72°C for 8 min. } 36x |
| | | ITS5 | GGAAGTAAAAGTCGTAACAAGG | [17] | |
| 4 | ITS | ITS-P5 | CCTTATCAYTTAGAGGAAGGAG | [18] | 94°C for 4 min.<br>94°C for 30 sec.<br>55°C for 40 sec.<br>72°C for 1 min.<br>72°C for 10 min. } 34x |
| | | ITS-u4 | RGTTTCTTTTCCTCCGCTTA | | |
| 5 | *rbcL* | rbcL_1F | ATGTCACCACAAACAGAAAC | [19] | 95°C for 4 min.<br>95°C for 45 sec.<br>55°C for 30 sec.<br>72°C for 50 sec.<br>72°C for 8 min. } 35x |
| | | rbcL_724R | TCGCATGTACCTGCAGTAGC | [20] | |
| 6 | *rbcL* | rbcL_1F | ATGTCACCACAAACAGAAAC | [19] | 95°C for 1 min.<br>95°C for 30 sec.<br>51°C for 30 Section<br>72°C for 1 min.<br>72°C for 5 min. } 35x |
| | | rbcLa_r | CTTCTGCTACAAATAAGAATCGATCTC | [21] | |
| 7 | *rbcL* | rbcL_1F | ATGTCACCACAAACAGAAAC | [19] | 95°C for 1 min.<br>95°C for 30 sec.<br>51°C for 30 sec.<br>72°C for 1 min.<br>72°C for 5 min. } 35x |
| | | rbcLa_SI_Rev | GTAAAATCAAGTCCACCRCG | [22] | |
| 8 | *rbcL* | rbcL_1F | ATGTCACCACAAACAGAAAC | [19] | 95°C for 1 min.<br>95°C for 30 sec.<br>51°C for 30 sec.<br>72°C for 1 min.<br>72°C for 5 min. } 35x |
| | | rbcLaj634R | GAAACGGTCTCTCCAACGCAT | [23] | |
| 9 | trnH-psbA | psbA3_f | GTTATGCATGAACGTAATGCTC | [24] | 92°C for 4 min.<br>94°C for 1 min.<br>52°C for 1 min.<br>64°C for 1 min.<br>64°C for 8 min. } 35x |
| | | trnHf_05 | CGCGCATGGTGGATTCACAATCC | [25] | |

**Table 3.**
*List of primers used for amplification of different loci and their PCR conditions.*

DNA sequencer at AgriGenome labs facility, Kochi, India. Each DNA barcode region was sequenced.

The different universal primers used in this study for the amplification are shown in **Table 4** along with the amplified product size.

The final edited sequences are provided in FASTA format below for each of the successful sequencing reactions:

>Vanilla S matk.

TCTCACATTTAAATTATGTGTCAGATCTACTAATACCCTATCCCATACATC
TGGAAATCTTAGTTCAAATTCTTCAATGCTGGGTCAAAGATGTTCTTTCTTTG
CATTTATTGCGATTGTTTTTTCACGAATATCATAATTTGAATAGTCTCGTTAC
TTCAAAGAAATCTATTTATGTCTTTTCAAAAATAAATAAAAGATTTTTTTTAT
TCCTACATAATTTTTATGTATATGAATCCGAATATCTATTCCTGTTTCTTCGT
AAACAGTCTTCTTATTTACGATCAACATCTTCTGGAGTGTTTCTTGAACAAA
CACATTTCTATGTAAAAATAGAACATATTCATCTTATAGTAGTAGTGTGTTG
TAATTCTTTCAAAAGGGACCTATGGTTTCTCGAAGATCCTTTCATGCATTAT
GTTCGATATCAAGGAAAAGCTATTCTGGGTTCAAAAGGAACTCTTATTCTGG
TGAATAAATGGAAATATTATCTTATTAATTTTTGGCAATCTTATTTTCACTTT
TGGTCTCAACCAGATAGGATCTATAGAAAGCAATTCTCCGACTATTCCTTTT
CTTTCCTGGGGTATTTTTCAAGTGTATTAAAAAATACTTTGGTAGTCAGAAA
TCAAATGCTAGAGAATTGCTTTCTCATAAATACTCCGACTCAGAAATTAGAT
ACCATAGCCCCGGTTATTTCTCTTATTGGATCCTTGTCGAAGGCAAAATTTT
GTACGTTAATGGGTCATCCCATTAGTAAACCGATCTGGACCGATTTATCGGA
TTCTGAGATTATTGATCGATTTTGTCGAATATGTAGAAATCTTTGTCGTTATC
ACAGTGGATCCTCAAAAAAACAGGTTT.

>VG matk.

TTCTCACATTTAAATTATGTGTCAGATCTACTAATACCCTATCCCATACAT
CTGGAAATCTTAGTTCAAATTCTTCAATGCTGGGTCAAAGATGTTCTTTCTTT
GCATTTATTGCGATTGTTTTTTCACGAATATCAGAATTTGAATAGTCTCGTTA
CTTCAAAGAAATCTATTTATGTCTTTTCAAAAAAAAATAAAAGATTCTTTTTA
TTCCTACATAATTTTTATGTATATGAATTCGAATATCTATTCATGTTTCTTCG
TAAACAGTCTTCTTATTTACGATCAACATCTTCTGGAGTGTTTCTTGAACAAA
CACATTTTTATGGAAAAATAGAACATATTCATCTTATAGTAGTAGTGTGTTT
TAATTCTTTAAAAAGCGACCTATGGTTTCTCGAAGATCCTTTCATGCATTAT
GTTCGATATCAAGGAAAAGCTATTCTGGGTTCAAAAGGAACTCTTATTCTGT
TGAATAAATGGAAATATTATATTATTTATTTTTTGCAATCTTATTTTCACTTT
TGGTCTCAACCAGATAGGATCTATAGAAAGCAATTCTCTGACTATTCCTTTT
CTTTCCTGGGGTATTTTTCAAGTGTATTAAAAAATACTTTGGTAGTCAGAAA
TCAAATGCTAGGGAATTGCTTTCTCATAAATATTCCGATTCAGAAATTAGAT
ACCACAGCCCCGGTGATTTCTCTTATTGGATCCTTGTCGAAGGCAAAATTTT
GTACGTTAATGGGTCATCCCATTAGTAAACCGATCTGGACTGATTTATCGGA
TTCTGAGATTATTGATCGATTTTGTCGAATATGTAGAAATCTTTGTCGTTATC
ACAGTGGA.

>VP matk.

TTCTCACATTTAAATTATGTGTCAGATCTACTAATACCCTATCCCATACAT
CTGGAAATCTTAGTTCAAATTCTTCAATGCTGGGTCAAAGATGTTCTTTCTTT
GCATTTATTGCGATTGTTTTTTCACGAATATCAGAATTTGAATAGTCTCGTTA
CTTCAAAGAAATCTATTTATGTCTTTTCAAAAAAAAATAAAAGATTCTTTTTA
TTCCTACATAATTTTTATGTATATGAATTCGAATATCTATTCATGTTTCTTCG
TAAACAGTCTTCTTATTTACGATCAACATCTTCTGGAGTGTTTCTTGAACAAA
CACATTTTTATGGAAAAATAGAACATATTCATCTTATAGTAGTAGTGTGTTT
TAATTCTTTAAAAAGCGACCTATGGTTTCTCGAAGATCCTTTCATGCATTAT
GTTCGATATCAAGGAAAAGCTATTCTGGGTTCAAAAGGAACTCTTATTCTGT
TGAATAAATGGAAATATTATATTATTTATTTTTTGCAATCTTATTTTCACTTT

| S. No | Plant Name | Matk | ITS | | rbcL Primers | | | | trnH-psbA |
|---|---|---|---|---|---|---|---|---|---|
| | | (product size) | (product size) | (product size) | (product size) | (product size) | (product size) | (product size) | (product size) |
| 1 | *V1* | 390F/1326R (1000 bp) | ITS4/ITS5 (800 bp) | — | rbcL_1F/rbcL_724R (NA*) | rbcLa_f/rbcLa_r (700 bp) | rbcLa_f/rbcLa_SI_Rev (600 bp) | rbcLa_f/rbcLaj634R (650 bp) | psbA3_f/trnHf_05 (750 bp) |
| 2 | *VG* | 390F/1326R (1000 bp) | ITS4/ITS5 (800 bp) | ITS-P5/ITS-u4 (800 bp) | rbcL_1F/rbcL_724R (NA*) | rbcLa_f/rbcLa_r (NA*) | rbcLa_f/rbcLa_SI_Rev (600 bp) | rbcLa_f/rbcLaj634R (650 bp) | psbA3_f/trnHf_05 (800 bp) |
| 3 | *VP* | 390F/1326R (1000 bp) | ITS4/ITS5 (800 bp) | ITS-P5/ITS-u4 (800 bp) | rbcL_1F/rbcL_724R (NA*) | rbcLa_f/rbcLa_r (NA*) | rbcLa_f/rbcLa_SI_Rev (600 bp) | rbcLa_f/rbcLaj634R (650 bp) | psbA3_f/trnHf_05 (800 bp) |

**NA, No amplification.*
*The Bold text indicates successful sequencing was done for these samples.*
*The ITS region was very problematic while sequencing and only V1 was completed, while sequencing is pending for VG and VP samples.*
*Loci rbcL and trnH-psbA has been successfully amplified but has not yet been sent for sequencing.*

**Table 4.**
*List of primers pairs used for amplification of different barcode loci and its estimated product sizes in agarose gel, for the* Vanilla *species under study.*

TGGTCTCAACCAGATAGGATCTATAGAAAGCAATTCTCTGACTATTCCTTTT
CTTTCCTGGGGTATTTTTCAAGTGTATTAAAAAATACTTTGGTAGTCAGAAA
TCAAATGCTAGGGAATTGCTTTCTCATAAATATTCCGATTCAGAAATTAGAT
ACCACAGCCCCGGTGATTTCTCTTATTGGATCCTTGTCGAAGGCAAAATTTT
GTACGTTAATGGGTCATCCCATTAGTAAACCGATCTGGACTGATTTATCGGA
TTCTGAGATTATTGATCGATTTTGTCGAATATGTAGAAATCTTTGTCGTTATC
ACAGTGGA.

>VS ITS.

AGTGGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAG
GATCATTGACGAGAGCTATGACTGATCGAGTGATCTGTGCAACCTGTGGGG
GTGCGACGGCTGTTTGATGTCGCATTCTTCCATCGCAGAGCTCCTGCTTCCA
GGGGGAGCTCGATGCTGTGGGGGGATAAACAACAGCCTATGGGCGTGGTCA
TACGCCAAGGGAGAGCAAATGTTAAGCCGCCAACGGGTGTGTTGTGCGTCG
CCAGGCCCAGTGGGGTATGGCAAACGAACACTGAACGACTCTCGACAACGG
ATATCTTGGCTCTCGCATCGATGAAGAACGCAGCGAAATGCGATACGTGTTG
TGAATTGTAGAATCCCGTGAACCATCCATTTTTTGAACGCAAGTTGCGCCCG
AGGATGCAAGCCGAGGGCACTCCTGCATGGGTGTAATGCGTTCTGTCGCTC
CTCGCGCAGGCATGGAATCGTTGGTTTAGATCAGCGGCCCCTCGCCAGGAT
GCGATCGATGGCACCCTGTGCTACGGCATGGCGTGTTCAAGCGTTGGGCGA
TGGTCGGCTGTAGACACGGCAAGAGGTGGATGCCACCGAGTGTTGTGGTGT
TGGCCAGTAGGAACCGATGTTGCAGTGCGACAAGGTGATGCCCCTTGCAAA
TCCAACTCCATGCTCCATGGTGTGGAATCGTGACCCCATGTTAGGTGAGGCT
ACCCGCTGAGTTTAAGCATATCAATAAGCGGA.

## 5. Result and discussion

### 5.1 Presence of important agronomic characters

Among the different species of vanilla studied *Vanilla aphylla* Blume and *V. pilifera* Holtt., flowered synchronously (**Figure 5**). *V. aphylla* occurs naturally in South India (**Figure 6**) and *V. pilifera* (**Figure 7**) in Assam, Northeast India. Flowers of both the species opened sequentially and lasted for one day in *V. pilifera*, whereas it lasted for 2 days in *V. aphylla*. In the former, signs of fruit set were observed even without pollination (**Figures 1** and **8**) whereas *V. apylla* flowers did not set fruit (**Figure 3**), ruling out the possibility of natural fruit set in this species, which is thus similar to *V. planifolia* (**Table 5**).

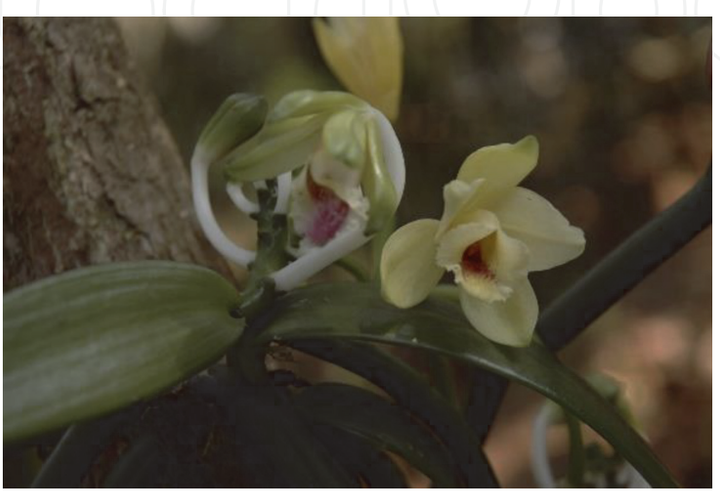

**Figure 5.**
V. pilifera *(T) inflorescence in comparison with that of* V. aphylla *(B). Arrow Indicates signs of natural fruit set without pollination in* V. pilifera.

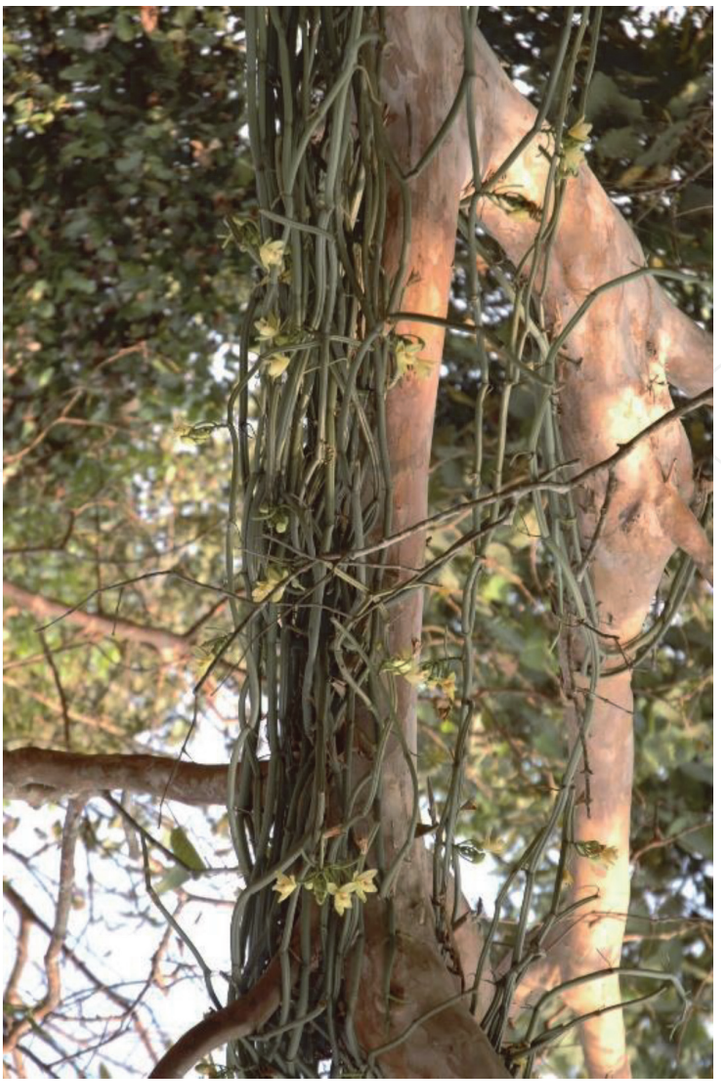

**Figure 6.**
V. pilifera *flower with a leaf (which develops at maturity).*

**Figure 7.**
*Vine of* V. aphylla *in bloom.*

**Figure 8.**
*Cross section of the ovary pedicel of* V. aphylla *without pollination, and after 24 hrs.*

| Species | Disease resistance | Fragrance | Natural seed set | Crossability | Fruit size | Flower life (hours) |
|---|---|---|---|---|---|---|
| *V. aphylla* | Tolerant to *Fusarium oxysporum* | — | Not seen | Crossable to *V. pilifera* and *V. planifolia* | | > 36 |
| *V. pilifera* | — | Highly fragrant | Symptoms seen | Crossable to *V. aphylla* | Larger | ~24 |
| *V. planifolia* | Susceptible | | Not seen | Crossable to *V. aphylla* | | < 24 |

**Table 5.**
*Important agronomic characters.*

Cross sections of the ovary pedicel were observed after closing of the flowers. Persistent perianth is characteristic to the genus and also indicative of effective pollination. In flowers where pollination is not effected, the perianth is shed after the flower closes. Perianth in *V. pilifera* were found to persist even without pollination and the cross sections indicated initiation of seed set (**Figures 9** and **10**), whereas *V. apyhlla* did not show any indications (**Figures 11** and **12**). Since rostellum is present in both the species, natural pollination without an aid is ruled out. It can be suspected, that the fragrance of the *V. pilifera* flowers attracts insects (which were found to frequent the flowers often) to visit them and bring about effective pollination.

Pollinations both self and interspecific hybridizations between the two species were done and fruits set was observed (**Figure 13**).

## 5.2 Sequence analysis

General observations from the experiment

1. The matK sequences of VG and VP are identical.

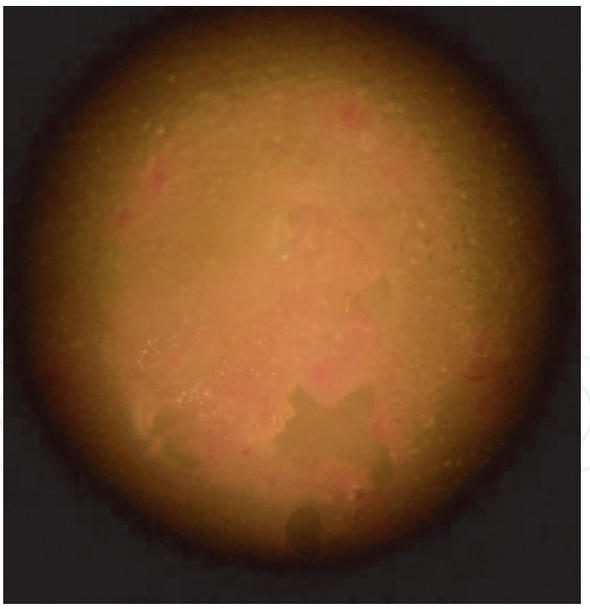

**Figure 9.**
*Cross section of the ovary pedicel of* V. aphylla *without pollination, and after 24 hrs.*

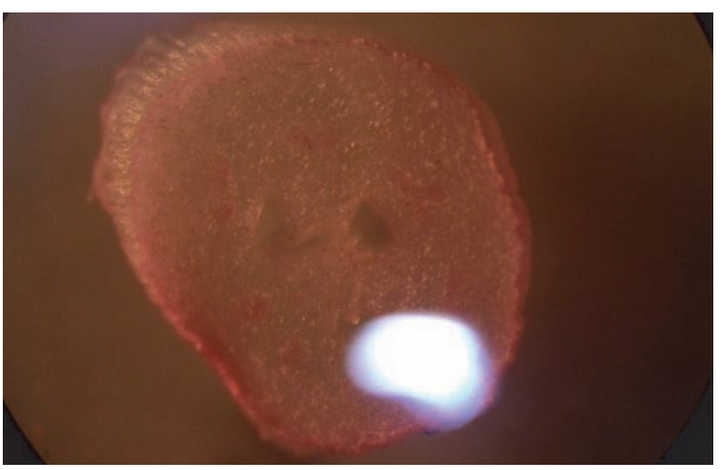

**Figure 10.**
*C.S. of ovary-pedicel of* V. pilifera *without pollination, showing indications of seed set.*

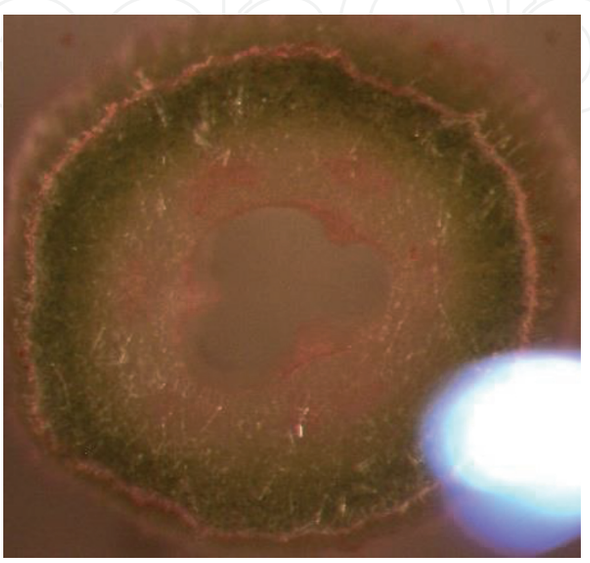

**Figure 11.**
*C.S. of ovary-pedicel of* V. pilifera *without pollination, showing indications of seed set.*

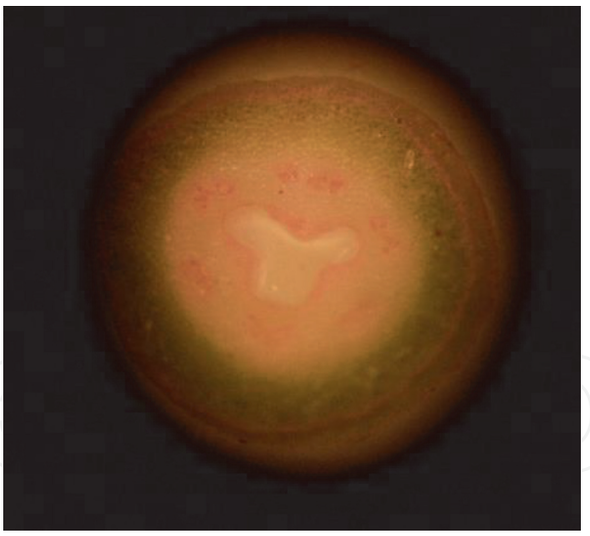

**Figure 12.**
*Flowers of* V. pilifera *in comparison with* V. aphylla.

**Figure 13.**
V. aphylla *inflorescence with fruit set after interspecific hybridization.*

2. The matK sequence of V1 was different from VG/VP at 21 nucleotide positions as shown in **Table 6** below.

| **Nucleotide position** | **1** | **1** | **2** | **2** | **2** | **3** | **3** | **3** | **3** | **3** | **4** | **4** | **4** | **5** | **5** | **6** | **6** | **6** | **6** | **6** | **7** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **3** | **9** | **0** | **3** | **5** | **2** | **2** | **6** | **7** | **8** | **7** | **9** | **9** | **0** | **6** | **4** | **6** | **6** | **8** | **9** | **7** |
| | **7** | **0** | **3** | **8** | **2** | **3** | **8** | **7** | **7** | **3** | **1** | **2** | **8** | **5** | **3** | **0** | **1** | **7** | **5** | **5** | **3** |
| VS matk | T | T | T | C | C | C | T | G | C | G | G | C | A | G | C | A | C | C | T | T | C |
| VG matk | G | A | C | T | A | T | G | T | A | C | T | A | T | T | T | G | T | T | C | G | T |
| VP matk | G | A | C | T | A | T | G | T | A | C | T | A | T | T | T | G | T | T | C | G | T |

**Table 6.**
*Matk sequence analysis.*

3. Blast search of the matk sequences of VG/VP in the NCBI blast search gave 100% match with *Vanilla planifolia* (Accession No. KJ566306.1), as in the NCBI search results shown below.

| Description | Max score | Total score | Query cover | E value | Ident | Accession |
|---|---|---|---|---|---|---|
| *Vanilla planifolia* chloroplast, complete genome | 1563 | 1563 | 100% | 0.0 | 100.00% | KJ566306.1 |
| Vanilla planifolia chloroplast, complete genome | 1546 | 1546 | 100% | 0,0 | 99.65% | MF197310.1 |
| Vanilla planifolia tRNA-Lys (trnK) gene, partial sequence; and maturase K (matK) gene, compl | 1524 | 1524 | 100% | 0.0 | 99.17% | JN181462.1 |
| Vanilla planifolia maturase K (matK) chloroplast pseudogene, partial sequence | 1507 | 1507 | 100% | 0.0 | 98.82% | AF263687.1 |
| Vanilla planifolia chloroplast matK pseudogene | 1507 | 1507 | 100% | 0.0 | 98.82% | AJ310079.1 |
| Vanilla planifolia plastid partial matK gene for maturase K, specimen voucher Chase O-199 K | 1423 | 1423 | 95% | 0.0 | 98.40% | AJ581443.1 |
| Vanilla somae voucher KFBG290 maturase K (matK) gene, partial cds: plastid | 1419 | 1419 | 100% | 0.0 | 96.93% | KY966974.1 |
| Vanilla aphylla chloroplast DNA, complete genome | 1419 | 1419 | 99% | 0.0 | 97.03% | LC085348.1 |
| Vanilla pilifera voucher V5 maturase K (matK) gene, partial cds: chloroplast | 1354 | 1354 | 99% | 0,0 | 95.64% | FJ816099.1 |
| Vanilla siamensis voucher V2 maturase K (matK) gene, partial cds: chloroplast | 1315 | 1315 | 97% | 0,0 | 95.23% | FJ816097.1 |
| Vanilla planifolia voucher SBB-0324 maturase K (matK) gene, partial cds: chloroplast | 1314 | 1314 | 84% | 0.0 | 100.00% | JN004635.1 |
| Vanilla planifolia isolate AD7LN25 maturase K (matK) qene. Partial cds: chloroplast | 1284 | 1284 | 89% | 0.0 | 97.11% | MF349972.1 |

4. Blast search of the matk sequences of VS in the NCBI blast search gave maximum identity with *Vanilla somae* (Accession No.KY966974.1). See the NCBI search results below.

| Description | Max score | Total score | Query cover | E value | Ident | Accession |
|---|---|---|---|---|---|---|
| Vanilla somae voucher KFBG290 maturase K (matK) gene, partial cds: plastid | 1570 | 1570 | 99% | 0.0 | 99.54% | KY966974.1 |
| Vanilla aphylla chloroplast DNA, complete genome | 1570 | 1570 | 99% | 0.0 | 99.54% | LC085348.1 |
| Vanilla pilifera voucher V5 maturase K (matK) gene, partial cds: chloroplast | 1489 | 1489 | 100% | 0.0 | 97.81% | FJ816099.1 |

| Description | Max score | Total score | Query cover | E value | Ident | Accession |
|---|---|---|---|---|---|---|
| *Vanilla pompona* chloroplast. Complete genome | 1474 | 1474 | 100% | 0,0 | 97.45% | MF197310.1 |
| Vanilla planifolia chloroplast. Complete genome | 1469 | 1469 | 100% | 0.0 | 97.34% | KJ566306.1 |
| Vanilla planifolia tRNA-Lys (trnK) gene, partial sequence: and maturase K (matK) gene, complel | 1452 | 1452 | 100% | 0.0 | 96.99% | JN181462.1 |
| Vanilla siamensis voucher V2 maturase K (matK) gene, partial cds: chloroplast | 1447 | 1447 | 98% | 0.0 | 97.33% | FJ816097.1 |
| Vanilla planifolia maturase K (matK) chloroplast pseudogene, partial sequence | 1435 | 1435 | 100% | 0.0 | 96.64% | AF263687.1 |
| Vanilla planifolia chloroplast matK pseudogene | 1435 | 1435 | 100% | 0.0 | 96.64% | AJ310079.1 |
| Vanilla planifolia isolate AD7LN25 maturase K (matK) gene, partial cds: chloroplast | 1408 | 1408 | 90% | 0.0 | 99.23% | MF349972.1 |
| Vanilla roscheri voucher NMK:838.10216 maturase K (matK) gene, partial cds: chloroplast | 1395 | 1395 | 93% | 0.0 | 97.89% | KU748308.1 |
| Vanilla aphylla voucher V1 maturase K (matK) gene, partial cds: chloroplast | 1391 | 1391 | 96% | 0.0 | 96.79% | FJ816096.1 |

5. Blast search of the ITS sequences of VS in the NCBI blast search gave maximum identity with Vanilla shenzhenica (Accession No. JF796930.1). See the NCBI search results below.

| Description | Max score | Total score | Query cover | E value | Ident | Accession |
|---|---|---|---|---|---|---|
| Vanilla somae voucher KFBG290 maturase K (matK) gene, partial cds: plastid | 1570 | 1570 | 99% | 0.0 | 99.54% | KY966974.1 |
| Vanilla aphylla chloroplast DNA, complete genome | 1570 | 1570 | 99% | 0.0 | 99.54% | LC085348.1 |
| Vanilla pilifera voucher V5 maturase K (matK) gene, partial cds: chloroplast | 1489 | 1489 | 100% | 0.0 | 97.81% | FJ816099.1 |
| Vanilla pompona chloroplast. Complete genome | 1474 | 1474 | 100% | 0,0 | 97.45% | MF197310.1 |
| Vanilla planifolia chloroplast. Complete genome | 1469 | 1469 | 100% | 0.0 | 97.34% | KJ566306.1 |
| Vanilla planifolia tRNA-Lys (trnK) gene, partial sequence: and maturase K (matK) gene, complel | 1452 | 1452 | 100% | 0.0 | 96.99% | JN181462.1 |
| Vanilla siamensis voucher V2 maturase K (matK) gene, partial cds: chloroplast | 1447 | 1447 | 98% | 0.0 | 97.33% | FJ816097.1 |
| Vanilla planifolia maturase K (matK) chloroplast pseudogene, partial sequence | 1435 | 1435 | 100% | 0.0 | 96.64% | AF263687.1 |
| Vanilla planifolia chloroplast matK pseudogene | 1435 | 1435 | 100% | 0.0 | 96.64% | AJ310079.1 |

| Description | Max score | Total score | Query cover | E value | Ident | Accession |
|---|---|---|---|---|---|---|
| Vanilla planifolia isolate AD7LN25 maturase K (matK) gene, partial cds: chloroplast | 1408 | 1408 | 90% | 0.0 | 99.23% | MF349972.1 |
| Vanilla roscheri voucher NMK:838.10216 maturase K (matK) gene, partial cds: chloroplast | 1395 | 1395 | 93% | 0.0 | 97.89% | KU748308.1 |
| Vanilla aphylla voucher V1 maturase K (matK) gene, partial cds: chloroplast | 1391 | 1391 | 96% | 0.0 | 96.79% | FJ816096.1 |

6. The ITS sequences matching with different Vanilla sp. were downloaded and subjected to analysis using MEGA7.0 software (**Table** 7).

| | *VS ITS* | *JF796930.1_Vanilla_shenzhenica* | *KY966687.1_Vanilla_somae* | *AF151006.1_Vanilla_aphylla* | *JF825978.1_Vanilla_siamensis* | *FJ425830.1_Vanilla_imperialis* | *FJ425840.1_Vanilla_roscheri* | *FJ425835.1_Vanilla_barbellata* | *FJ425834.1_Vanilla_africana* | *EU498163.1_Vanilla_bahiana* | *GQ867241.1_Vanilla_planifolia* | *GQ867237.1_Vanilla_pompona* | *AF391785.1_Vanilla_hirsuta* | *EU498165.1_Vanilla_edwallii* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *VS ITS* | 0 | | | | | | | | | | | | | |
| *JF796930.1_Vanilla_shenzhenica* | 20 | 0 | | | | | | | | | | | | |
| *KY966687.1_Vanilla_somae* | 20 | 0 | 0 | | | | | | | | | | | |
| *AF151006.1_Vanilla_aphylla* | 27 | 19 | 19 | 0 | | | | | | | | | | |
| *JF825978.1_Vanilla_siamensis* | 42 | 39 | 39 | 41 | 0 | | | | | | | | | |
| *FJ425830.1_Vanilla_imperialis* | 40 | 33 | 33 | 41 | 47 | 0 | | | | | | | | |
| *FJ425840.1_Vanilla_roscheri* | 41 | 34 | 34 | 45 | 53 | 40 | 0 | | | | | | | |
| *FJ425835.1_Vanilla_barbellata* | 55 | 48 | 48 | 55 | 66 | 57 | 58 | 0 | | | | | | |
| *FJ425834.1_Vanilla_africana* | 63 | 51 | 51 | 56 | 66 | 56 | 60 | 76 | 0 | | | | | |
| *EU498163.1_Vanilla_bahiana* | 96 | 89 | 89 | 96 | 101 | 97 | 99 | 107 | 100 | 0 | | | | |
| *GQ867241.1_Vanilla_planifolia* | 107 | 100 | 100 | 107 | 119 | 106 | 113 | 110 | 110 | 45 | 0 | | | |
| *GQ867237.1_Vanilla_pompona* | 99 | 92 | 92 | 99 | 104 | 98 | 103 | 107 | 104 | 20 | 45 | 0 | | |
| *AF391785.1_Vanilla_hirsuta* | 101 | 92 | 92 | 102 | 114 | 101 | 102 | 108 | 100 | 39 | 31 | 43 | 0 | |
| *EU498165.1_Vanilla_edwallii* | 132 | 133 | 133 | 136 | 143 | 136 | 143 | 141 | 146 | 139 | 150 | 136 | 146 | 0 |

**Table 7.**
*Estimates of evolutionary divergence between sequences.*

The number of base differences per sequence from between sequences are shown. The analysis involved 14 nucleotide sequences. All ambiguous positions were removed for each sequence pair. There was a total of 657 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (**Figure 14**).

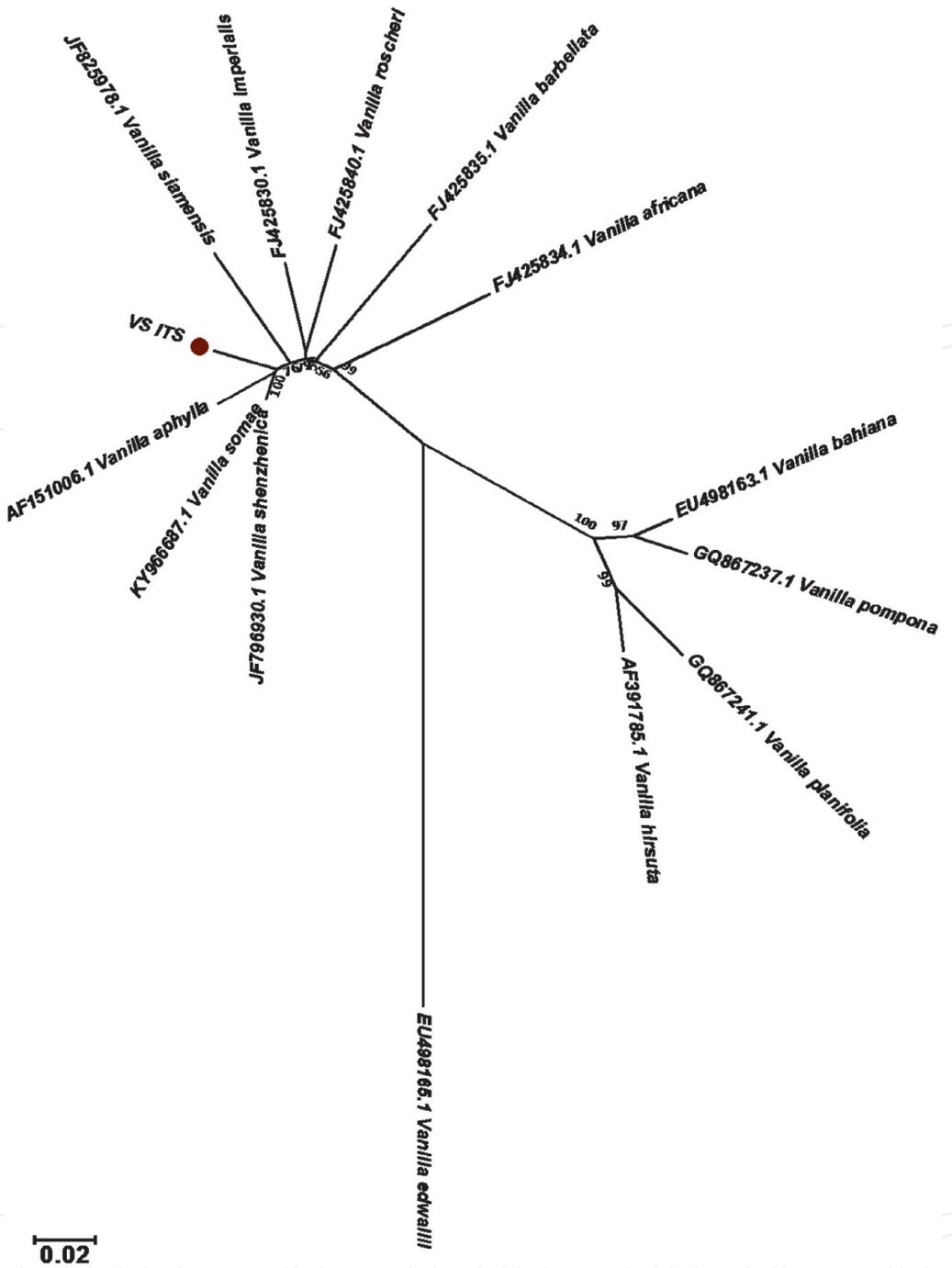


**Figure 14.**
*Phylogenic analysis of the ITS sequences inferred using the neighbor-joining method, computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 14 nucleotide sequences. All ambiguous positions were removed for each sequence pair. There was a total of 657 positions in the final dataset. The analyses were conducted in MEGA7.*

## 6. Conclusions

The analysis involved 14 nucleotide sequences. All ambiguous positions were removed for each sequence pair. There was a total of 657 positions in the final dataset. The analyses were conducted in MEGA7. The phylogeny analysis also revealed the separate clustering offer leafy and leafless species. *Vanilla siamensis*, a leafy species, indicating signs of self-pollination in its wild, in Thailand, clustered with leafless *V. aphylla* species.

The studies further reveal the complexity of the biosynthesis of the natural vanillin synthesis. However, it is to be further analyzed whether leafy character is associated with enhanced photosynthetic products that indirectly affect the vanillin synthesis too. This reiterates the need for conservation of the genetic resources [12] of Vanilla across the continents, for implementing meaningful breeding programs, to enhance vanillin productivity in addition to disease resistance and reproductive behavior.

## Acknowledgements

The first author acknowledges the former Directors of Indian Institute of Spices Research, Kozhikode, Kerala from where she initiated these studies. The authors are thankful for the financial support rendered by University Grants Commission (MRP and CPE funds), DST-FIST funds, for taking up these research initiatives, in their laboratory and express gratitude to Scientists at Centre for Medicinal Plant Research, Kottakkal, Malappuram, for support in Sequencing work.

## Author details

Minoo Divakaran* and N.T. Fathima Rafieah
Department of Botany, Providence Women's College, Kozhikode, Kerala, India

*Address all correspondence to: minoodivakaran@gmail.com

## References

[1] Correll, D. S. (1953). Vanilla-its botany, history, cultivation and economic importance. Economic Botany, 7(4), 291–358. DOI:10.1007/bf02930810

[2] Bouriquet G (1954) Le vanillier et la vanille dans le monde. Encyclope'die Biologique XLVI. Lechevalier, Paris VI

[3] Swamy, B. G. L. (1947). On the life-history of Vanilla planifolia. Botanical Gazette, 108(3), 449–456. DOI:10.1086/335429

[4] Purseglove JW, Brown EG, Green CL, Robbins SRJ (eds.) 1981. Spices. Vol. 2. Longman Inc., New York, pp 644-735.

[5] Spices Board 2000. Vanilla Status paper. Spices Board, Cochin, India, pp. 33.

[6] Lubinsky P (2004) Vanilla diversity in Mexico. In: Bakto Fla-vors, Rutgers University (eds) Vanilla 2004 Europe, Second International Congress, Cannes, France, 30 Sept–01 Oct 2004.

[7] Hu, Y., Resende, M. F. R., Jr., Bombarely, A., Brym, M., Bassil, E., and Chambers, A. H. (2019). Genomics-based spanersity analysis of Vanilla species using a Vanilla planifolia draft genome and genotyping-by-sequencing. Scientific Reports, 9(1). DOI:10.1038/s41598-019-40144-1

[8] Minoo, D., Jayakumar, V. N., Veena, S. S., Vimala, J., Basha, A., Saji, K. V., Nirmal Babu, K., and Peter, K. V. (2007). Genetic variations and interrelationships in Vanilla planifolia and few related species as expressed by RAPD polymorphism. Genetic Resources and Crop Evolution, 55(3), 459–470. DOI: 10.1007/s10722-007-9252-3

[9] Satish K and Manilal K.S. 1993. Vanilla aphylla Blume. Amer. Orch. Soc. Bull. 62, 394–397.

[10] Seidenfaden G. 1978. Orchid genera in Thailand VI Neottioideae. Dansk. Bot. Ark. 32, 138–146.

[11] Franklin, W.M., 1963. Chromosome number and behavior in a Vanilla hybrid and several Vanilla species. Bull. Torrey Bot. Club 90, 416–417

[12] Minoo, D., Nirmal Babu, K., and Peter, K. V. (2006). Conservation of Vanilla species, in vitro. Scientia Horticulturae, 110(2), 175–180. DOI: 10.1016/j.scienta.2006.07.003

[13] Minoo, D., Nirmal Babu, K., Ravindran, P. N., and Peter, K. V. (2006). Interspecific hybridization in vanilla and molecular characterization of hybrids and selfed progenies using RAPD and AFLP markers. Scientia Horticulturae, 108(4), 414–422. DOI: 10.1016/j.scienta.2006.02.018

[14] Gallage, N. J., Hansen, E. H., Kannangara, R., Olsen, C. E., Motawia, M. S., Jørgensen, K., Holme, I., Hebelstrup, K., Grisoni, M., and Møller, B. L. (2014). Vanillin formation from ferulic acid in Vanilla planifolia is catalysed by a single enzyme. Nature Communications, 5(1). DOI:10.1038/ncomms5037

[15] Cuenoud, P., Savolainen, V., Chatrou, L. W., Powell, M., Grayer, R. J., and Chase, M. W. (2002). Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB, and matK DNA sequences. American Journal of Botany, 89(1), 132–144. DOI:10.3732/ajb.89.1.132

[16] White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). AFmplification and direct sequencing of fungal ribosomal rna genes for phylogenetics. In PCR Protocols (pp. 315–322). Elsevier. doi: 10.1016/b978-0-12-372180-8.50042-1

[17] Sang, T., Crawford, D. J., and Stuessy, T. F. (1995). Documentation of

reticulate evolution in peonies (Paeonia) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. Proceedings of the National Academy of Sciences, 92(15), 6813–6817. DOI:10.1073/pnas.92.15.6813

[18] Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., and Zhou, S. (2015). Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. Molecular Ecology Resources, 16(1), 138–149. DOI:10.1111/1755-0998.12438

[19] Olmstead, R. G., Michaels, H. J., Scott, K. M., and Palmer, J. D. (1992). Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of rbcL. Annals of the Missouri Botanical Garden, 79(2), 249. DOI:10.2307/2399768

[20] Fay, M. F., Swensen, S. M., and Chase, M. W. (1997). Taxonomic affinities of Medusagyne oppositifolia (Medusagynaceae). Kew Bulletin, 52(1), 111. DOI:10.2307/4117844

[21] Kress, W. J., and Erickson, D. L. (2007). A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS ONE, 2 (6), e508. DOI:10.1371/journal.pone.0000508

[22] Kress, W. J., Erickson, D. L., Jones, F. A., Swenson, N. G., Perez, R., Sanjur, O., and Bermingham, E. (2009a). Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy of Sciences, 106(44), 18621–18626. DOI:10.1073/pnas.0909820106

[23] Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., Percy, D. M., Hajibabaei, M., and Barrett, S. C. H. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS ONE, 3(7), e2802. DOI:10.1371/journal.pone.0002802

[24] Sang, T., Crawford, D. J., and Stuessy, T. F. (1997). Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). American Journal of Botany, 84(8), 1120–1136. DOI:10.2307/2446155

[25] Tate, J.A., Simpson, B.B. (2003). Paraphyly of Tarasa (Malvaceae) and diverse origins of the Polyploid species. Systematic Botany. 28. 723-737. DOI: 10.1043/02-64.1