VANDA TRICOLOR LINDL. CONSERVATION IN JAVA, INDONESIA: GENETIC AND GEOGRAPHIC STRUCTURE AND HISTORY

LAUREN M. GARDINER

Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 United Kingdom
l.gardiner@uea.ac.uk

Introduction

Vanda tricolor Lindl. is widespread in cultivation in South East Asia, being relatively easily cultivated in the garden, and is often seen growing floriferously on garden trees, fence posts and verandas. Whilst the species is widespread in cultivation in its native regions of Java and Bali, wild populations are small and highly fragmented. The author's recent field trip to Java and Bali found few wild examples of the species, and the few plants found were in the most inaccessible trees, due to massive over-collection, often with no other *V. tricolor* plants to be found in the area.

A study of *V. tricolor* specimens, collected in the wild by the Indonesian botanic gardens, using several population genetics methods elucidates whether the specimens exhibit any genetic structure and how that might relate to their geographical origins. The results show several interesting points including the putative geographical origin of the species and the dates of the divergence of the species from this origin and into other regions.

Investigation

Thirty four specimens of *V.tricolor* with original collection locality information were collected from all four of the Indonesian Botanic Gardens – Bogor, Cibodas and Purwodadi in Java and Eka Karya in Bali (see Table 1). Leaf specimens were collected for DNA extraction and where plants were in flower, flowers were preserved in spirit for morphological analysis. *Vanda luzonica* Loher ex Rolfe has previously been identified as a putative sister taxon to *V.tricolor* (Gardiner, unpublished data) and was used throughout the genetic study as an outgroup. The *V.luzonica* DNA sequence (referred to as 'VL' throughout) was ampli-

fied from DNA sample 19632 in the RBG Kew DNA bank, collected by the author from Motes' Orchids in Florida, USA, and originally from the Philippines.

The original collecting localities of the *V.tricolor* specimens divided into five clear, non-overlapping geographic regions so each specimen was assigned to a geographic region – West Java, Central Java (including D.I. Yogyakarta), East Java, Bali, Central Sulawesi, (Fig. 1). At the national park at Kaliurung on Mount Merapi in Java, there is an effort to cultivate and reintroduce *V. tricolor* to 'protected' areas of the park on a small scale when funds and time are available. The specimens originally collected from 'Central Java' and used in this study are a selection of those being reintroduced.

A portion of the single copy chloroplast DNA region trnL-F was amplified using primers c and f of Taberlet et al (1991), as in Table 2, and sequenced using just primer c. The region amplified includes the 3' end of the 5' trnL (UAA) exon, and most of the trnL intron, and is between 270-319 base pairs in length. During phylogenetic work on the genus
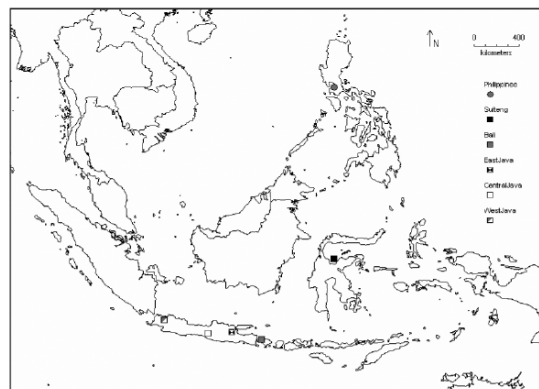

FIGURE 1. Geographic regions from which *V.tricolor* (and *V.luzonica* - outgroup) specimens were collected

TABLE 1. Specimens used in study

Botanic garden	Collection number	Sequence ref	Origin	floral material collected
Cibodas, Java	LG138	w1	West Java	o
Cibodas, Java	LG141	w3	West Java	o
Cibodas, Java	LG147	w2	West Java	x
Bogor, Java	LG129	e8	East Java	x
Bogor, Java	LG150	w4	West Java	o
Bogor, Java	LG153	w5	West Java	o
Bogor, Java	LG154	w6	West Java	o
Bogor, Java	LG157	c7	Central Java	x
Bogor, Java	LG159	c8	Central Java	x
Bogor, Java	LG163	c9	Central Java	x
Bogor, Java	LG164	w7	West Java	x
Bogor, Java	LG175	e9	East Java	x
Purwodadi, Java	LG177	e1	East Java	o
Purwodadi, Java	LG181	e2	East Java	o
Purwodadi, Java	LG182	e3	East Java	o
Purwodadi, Java	LG183	e4	East Java	o
Purwodadi, Java	LG184	e5	East Java	o
Purwodadi, Java	LG185	e6	East Java	o
Purwodadi, Java	LG186	e7	East Java	o
Purwodadi, Java	LG230	c1	Central Java	o
Purwodadi, Java	LG231	c2	Central Java	o
Purwodadi, Java	LG232	c3	Central Java	o
Purwodadi, Java	LG233	c4	Central Java	o
Purwodadi, Java	LG234	c5	Central Java	o
Purwodadi, Java	LG235	c6	Central Java	o
Eka Karya, Bali	LG242	b1	Bali	o
Eka Karya, Bali	LG243	b2	Bali	o
Eka Karya, Bali	LG244	b3	Bali	o
Eka Karya, Bali	LG245	b4	Bali	x
Eka Karya, Bali	LG246	b5	Bali	o
Eka Karya, Bali	LG247	b6	Bali	o
Eka Karya, Bali	LG249	b7	Bali	x
Eka Karya, Bali	LG251	s1	Sulawesi	x
Eka Karya, Bali	LG255	b8	Bali	x

Vanda by the author and by Kocyan (pers. comm. 2006), it had been found that this sequence contains a 'hypervariable' region, with sequences from different species proving to be highly variable and difficult to align, but was considered to be sufficiently variable for a single species population level study.

TABLE 2. Primers used for amplification and sequencing

Primer	Sequence	Author
c	CGAAATCGGTAGACGCTACG	Taberlet *et al.* 1991
f	ATTTGAACTGGTGACACGAG	Taberlet *et al.* 1991

Total cellular DNA was extracted from silica-gel dried leaf material (Chase and Hills, 1991), according to the method of Doyle and Doyle (1987), without an RNA treatment. The PCR programme consisted of 1 min at 94°C (initial denaturation), 35 cycles of 1 min at 94°C (denaturation), 1 min at 50°C (annealing) and 1 min 10 s at 72°C, and 5 min at 72°C final extension. Amplification was performed using 50 µl reactions, with 1 µl of each primer, 5 µl NH_4, 3 µl $MgCl_2$ (2.5 µl for trnL-F), 1 µl dNTPs, 0.2 µl *Taq* polymerase 1 µl DNA (approximately 20-70ng).

	indel 1	indel 2	indel 3	
TAATATAATA	----------	------------------------	-------------------------------------	GAGTAATATAATA
TAATATAATA	GTATAT	------------------------	-------------------------------------	GAGTAATATAATA
TAATATAATA	GTATAT	GAGTAATATAATA	-------------------------------	GAGTAATATAATA
TAATATAATA	GTATAT	GAGTAATATAATA	GAGTAATAATATAATA	GAGTAATATAATA

Can be coded as:

```
                123
TAATATAATA 000 GAGTAATATAATA
TAATATAATA 100 GAGTAATATAATA
TAATATAATA 110 GAGTAATATAATA
TAATATAATA 111 GAGTAATATAATA
```

Where: 0 = absence of sequence, 1 = presence of sequence

FIGURE 2. Example of coding indels

Amplified products were purified using QIAquick PCR purification kit (Qiagen). Cycle sequencing reactions were performed using Big Dye terminator chemistry (PE Biosystems), under the following conditions: an initial denaturation step at 96°C for 1 min followed by 25 cycles of 10 s at 96°C, 5 s at 50°C, and 4 min at 60°C using 10 µl reactions made with 2 µl Big Dye 3.1, 1 µl Big Dye buffer, 1 µl primer c, and 1 µl DNA (approximately 20-25 ng). Sequences were run on an ABI3700 capillary sequencer (PE Biosystems). Chromas™ and BioEdit™ software programs were used to edit and align sequences, with the initial alignment performed using the Clustal X algorithm in BioEdit™, followed by manual optimisation.

Phylogenetic analysis is not usually used at the population level for a number of reasons. Two reasons are that firstly the lower level of genetic variation within species often means that relationships cannot be resolved due to lack of information in the DNA sequences examined. Secondly, species populations are less likely to be geographically and genetically isolated than distinct species, with gene flow and therefore reticulated relationships possible and even likely. The bifurcating nature and the inherent assumption of non-reticulation and gene flow of phylogenetic analyses means that they are often highly inappropriate for population level analyses. Simple phylogenetic analyses can be carried out however to investigate whether there is strong phylogenetic signal in the data, which might imply reduced gene flow between individuals/populations, and to help guide the investigation.

Neighbour-joining (NJ) and maximum parsimony (MP) analyses were performed using PAUP* version 4.0b10 (Swofford, 2002), with and without rooting with the outgroup, *V.luzonica*. The MP analyses were heuristic searches with 100 replicates, with equal weighting on all characters, ACCTRAN optimization and TBR branch swapping. The initial MP analysis showed that the phylogenetic relationships between the specimens were completely unresolved, and the NJ analysis connected all specimens in a perfect ladder which implies that the relationships are not resolved (the NJ method will always connect taxa step-by-step in this way if there is no phylogenetic resolution, which is supported by the lack of resolution in the MP analysis).

There are several clearly identifiable insertion/deletion events ("indels") in the sequence alignment, these can carry information about phylogenetic relationships which can otherwise be obscured by the individual bases present or absent in those positions. In order to deal with these indels, in a manner which reduces the influence of the individual bases and emphasises the presence/absence of a section of sequence at a particular position, they can be "coded" using various methods. Here the indels were coded simply as presence/absence by the use of an extra character inserted at the end of the alignment for each indel, and the character "0" used to denote the
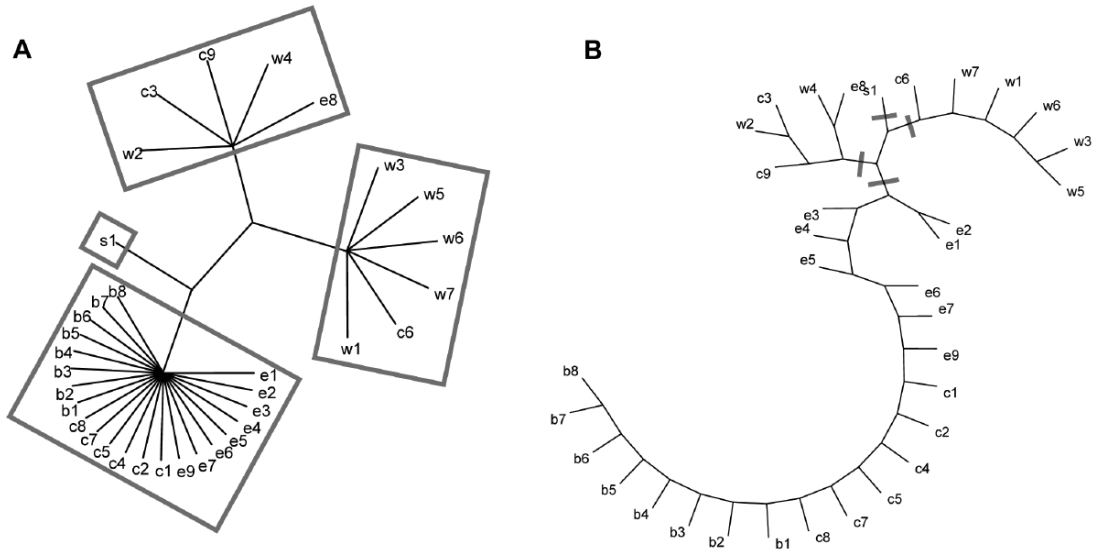
FIGURE 3. Unrooted consensus analyses of *Vanda tricolor* specimens, without outgroup, **A**: MP, **B**: NJ, showing four corresponding clades (denoted by gray boxes/lines)
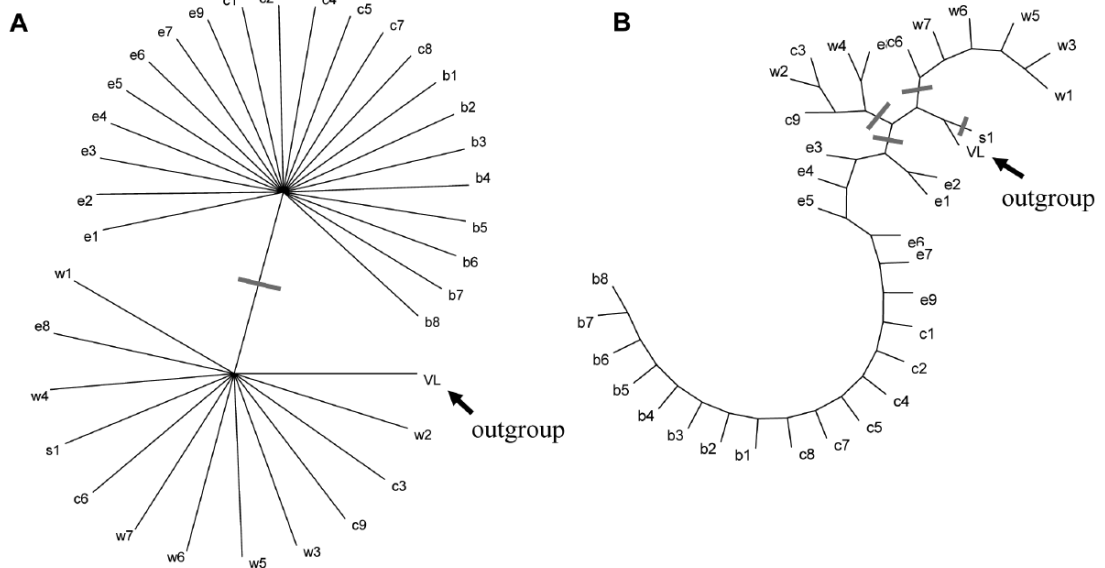

FIGURE 4. Unrooted consensus analyses of Vanda tricolor specimens, including outgroup (VL), A: MP showing only two clades, B: NJ showing four clades previously identified (denoted by gray lines)

absence of sequence at the indel region, and the character "1" to denote the presence of sequence at the indel region, as shown in (Fig. 2).

After coding the indels using this simple presence/absence 0/1 method, the analyses showed much clearer relationships between the specimens. Excluding the outgroup, MP analysis shows the specimens resolved into four clades (as in Fig. 3A) and the NJ analysis shows groupings which correspond to these four clades (Fig. 3B).
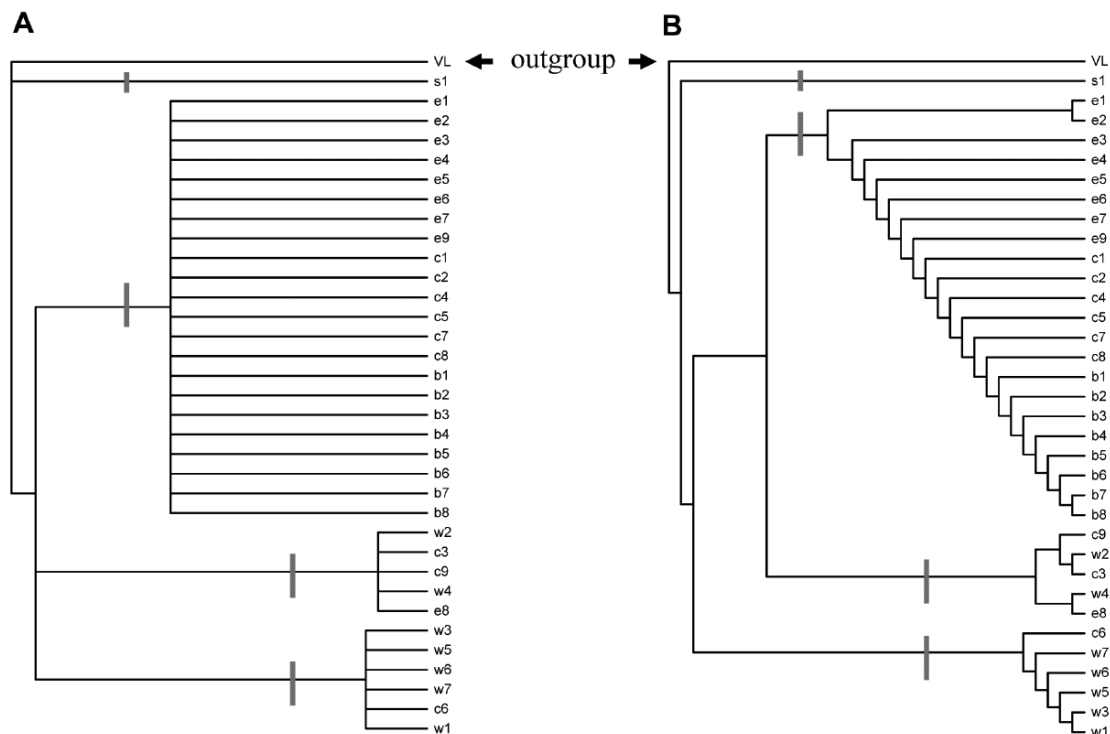
FIGURE 5. Rooted consensus analyses of *Vanda tricolor* specimens (rooted with outgroup, *V.luzonica*), A: MP, B: NJ, showing the four clades also identified in unrooted analyses (indicated by gray boxes/lines)

When the outgroup, *V.luzonica*, was included in the analyses but not used to root them, two clades congruent with the four clades previously seen were present in the MP consensus tree (figure 4A), but all four clades were present in the NJ consensus tree (Fig. 4B). The outgroup appears to be mostly closely related to the *V.tricolor* specimen from Sulawesi, s1, in figure 4B, and with the [w1, w3, w5, w6, w7, c6] clade. Rooting the analyses with the outgroup results in all four clades being resolved in both the MP and the NJ consensus trees, and the outgroup most closely related to specimen s1 (figures 5A and 5B).

The coded sequence alignment was analysed using two programs used for population genetic analyses – TCS (Clement et al 2000) and Network (Bandelt et al 1999). TCS estimates gene genealogies which may include multifurcations and/or reticulations, producing networks using a method of network estimation called Statistical Parsimony (Templeton et al 1992). Statistical parsimony calculates the probability for DNA pairwise differences until the probability exceeds 0.95 ("95% cutoff"), emphasising the shared characters among haplotypes by connecting the most similar haplotypes first, then those with one mutation, then two, and so forth. The number of mutational differences associated with this probability just before the 95% cutoff is the maximum number of mutational connections between pairs of sequences that can be justified by the criterion. TCS only processes sequences using the standard DNA code, so the 0/1 coded characters used previously were converted into ACGT characters prior to TCS analysis. Because TCS uses the differences between the sequences, not the actual bases in the sequence (ie. it does not infer information from *which particular* base is at a position, rather than whether the base is *different* to that in the other sequences at the same position), the 0 and 1 characters were substituted with A, C, G or T. Network uses a method of network estimation called Median Joining, combining minimum spanning trees into a single network using median vectors (nodes) which represent missing intermediate haplotypes, but does rely on the absence of recombination in the data.

Both methods of estimating the genealogy of the

TABLE 3. Specimens used in study

Haplotype	Sequence reference	Origin
1	c1	Central Java
1	c2	Central Java
1	c4	Central Java
1	c5	Central Java
1	c7	Central Java
1	c8	Central Java
1	e1	East Java
1	e2	East Java
1	e3	East Java
1	e4	East Java
1	e5	East Java
1	e6	East Java
1	e7	East Java
1	e9	East Java
1	b1	Bali
1	b2	Bali
1	b3	Bali
1	b4	Bali
1	b5	Bali
1	b6	Bali
1	b7	Bali
1	b8	Bali
2	w2	West Java
2	c3	Central Java
2	c9	Central Java
3	w3	West Java
3	w5	West Java
3	w6	West Java
3	w7	West Java
3	c6	Central Java
4	s1	Sulawesi
5	w4	West Java
6	e8	East Java
7	w1	West Java
8	VL (V.luzonica)	Philippines

specimens produced the same network, collapsing the data into 8 haplotypes (numbered 1-8), the eighth being the *V.luzonica* sequence (Fig. 5). The specimens fall into the haplotypes as listed in Table 3.

The haplotypes are not individually restricted to a single geographic region, but appear to be distributed in a non-random manner (Table 4), with the most haplotypic diversity in West Java (four haplotypes present – 2, 3, 5, 7), the least in Bali and Sulawesi (only one haplotype present in each, 1 and 4 respec-

tively), and haplotype 1 is the most widespread, found in Central and East Java and in Bali.

Both methods of analysis included the same loop which cannot conclusively be broken using the guidelines of Crandall and Templeton (1993) (Fig. 6). Singletons (ie. unique haplotypes or those present at low frequencies) are more likely to be at the tips of networks than internally, so the loop is most likely to be broken at point A or B than at the other two possible places. The unrooted phylogenetic analyses (MP and NJ) showed that haplotypes 8 and 4 (the *V.luzonica* outgroup sequence, VL, and *V.tricolor*, s1) are more closely related to haplotype 3 (w3, w5, w6, w7, c6) than to haplotype 1 (c1, c2, c4, c5, c7, c8, e1, e2, e3, e4, e5, e6, e7, e9, b1, b2, b3, b4, b5, b6, b7, b8), supporting network 5B as the more likely of the two.

The network was subdivided into clades as described in Templeton (1998) (Fig. 7) and Nested Clade Phylogenetic Analysis (NCPA) carried out using the program GeoDis (Posada et al 2000). The results of the contingency analysis of clade 1-3 (equivalent to clade 2-2, and included in clade 3-1 which includes the whole network) is significant at the 5% level, with all other clades being non-significant. When interpreted using Templeton's latest revised inference key (http://darwin.uvigo.es), the outcome of the nested contingency analysis of geographical associations infers that clade 1-3 has been affected by restricted gene flow with isolation by distance. No conclusions can be evaluated from the data for the other clades, probably as a result of low sample size, to which NCPA is sensitive.

The geographic pattern of the *V.tricolor* specimens subdivision and structure was tested by analysis of molecular variance (AMOVA) (Excoffier et al 1992), with ø-statistics calculated analogous to the F-statistics of population genetics. Many different structural hierarchies were assessed, grouping the haplotypes according to geography and phylogenetic clades, and the results showed that the most appropriate subdivision of the haplotypes accounting for 88.66% of the variation among the groups was that which divided the specimens into the four clades found in the phylogenetic and haplotypic analyses, and shown in the network grouped as in figure 8: [s1], [w1, w3, w5, w6, w7, c6], [w4, c3, c9, e8], and [c1, c2, c4, c5, c7, c8, e1, e2, e3, e4, e5, e6, e7, e9, b1, b2, b3, b4, b5, b6, b7, b8].
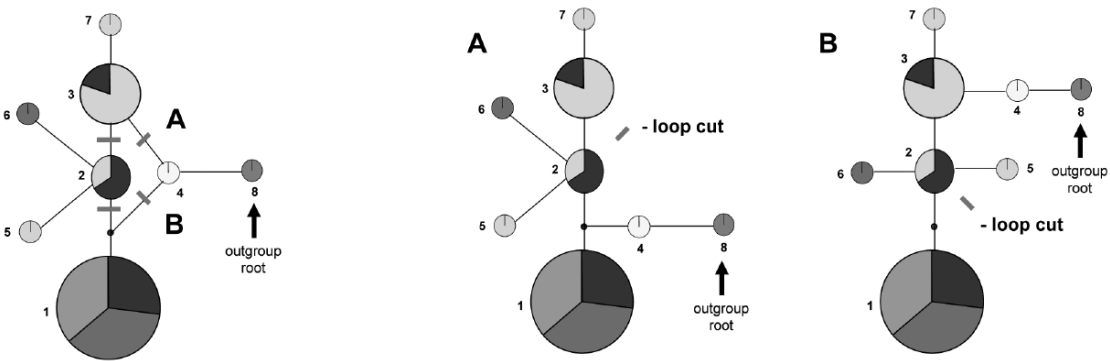
FIGURE 6. Statistical parsimony and Median-joining network, gray bars indicate the four places where the ambiguous loop could be cut, with alternative topologies - A: ambiguous loop cut at position A, B: ambiguous loop cut at position B.
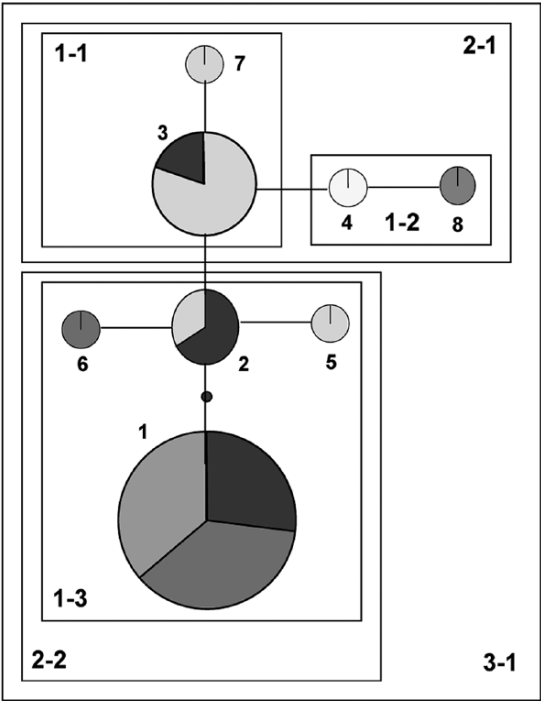

FIGURE 7. Network coded for Nested Clade Phylogenetic Analysis

Discussion

The results suggest that geographically the most recent common ancestor of *V.tricolor* and *V.luzonica* originated in or near to Sulawesi and the Philippines, and colonised West and Central Java first, before spreading into East Java and Bali. As the species spread across the region, mutational change resulted in new haplotypes evolving, so we could predict that haplotype 4 is the oldest *V.tricolor* haplotype, followed by haplotype 3, and then haplotype 2, with the terminal haplotypes 1, 5, 6, 7 being the most recently evolved. These divergences can be used to infer when the species spread from Sulawesi, into West and Central Java, and then into East Java and Bali, using previously published mutational rates of change in chloroplast DNA sequences and information about the sequences used. The method of Saillard et al (2000) was used to calculate these divergence times, and the aligned sequence length of 271 base pairs and the mutation rate of 4.87×10^{-10} substitutions per site per year (Richardson et al 2001) were used. The calculated date of the movement of *V.tricolor* from Sulawesi into West and Central Java is 1,894,269

TABLE 4. Specimens used in study

Locality	Number of haplotypes	Total number of individuals	Haplotype (number of individuals)
West Java	4	7	2 (1), 3 (4), 5 (1), 7 (1)
Central Java	3	9	1 (6), 2 (2), 3 (1)
East Java	2	9	1 (8), 6 (1)
Bali 1	8	1 (8)	
Sulawesi	1	1	4 (1)
Philippines (outgroup)	1	1	8 (1)
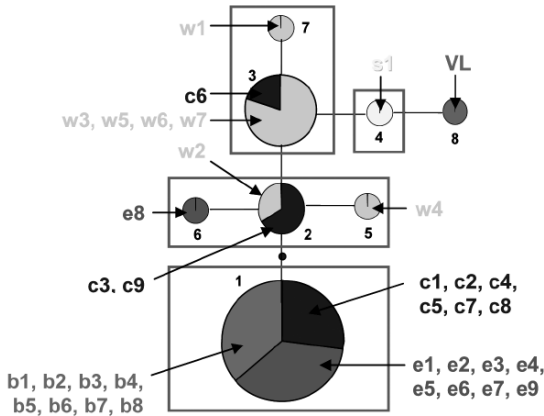
FIGURE 8. Haplotypic network with four clades that account for 88.66% of the variation among groups

FIGURE 9. *Vanda tricolor* var *suavis* growing in Purwodadi Botanic Garden, Java

years ago +/- 568,281 (between 1,231,275 – 2,462,550 years ago), and from West and Central Java into East Java and Bali is 1,470,844 years ago +/- 134,500 (between 1,336,345 – 1,605,344 years ago). These dates are at the end of and after the Pleistocene, when there was a shallow Java Sea in the region and there is evidence of land bridges which would have facilitated the migration of plants and animals from continental Asia.

The main restriction of this investigation has unfortunately been the relatively low sample size, totalling 34 *V.tricolor* specimens, due to the lack of specimens with collection localities recorded and the lack of populations in the wild available for sampling. As statistical and modelling-based analyses are sensitive to sample sizes, the results must be interpreted with caution. However, the results to suggest an interesting colonisation pattern which may warrant further investigation in this and other species, and shows that the remaining populations of V.tricolor in Java and Bali appear to be at least partially genetically isolated, with widespread haplotypes and others only found in one region.

Conservation efforts to reintroduce and/or increase populations of *V.tricolor* should take into account the diversity of haplotypes in Java and Bali in order to conserve genetic diversity and potential. This study shows that not all specimens of *V.tricolor* in Java and Bali are genetically identical and that different regions have different genetic varieties. The next stage in this investigation will examine the floral morphology of the specimens using the spirit collection material collected at the same time as the DNA material, and determine whether there is any relationship between morphological characters observed in the flowers and their genetic and/or geographic structure. There is taxonomic confusion about the status of species/varieties/subspecies such as varieties *suavis*, *planilaberis* and *purpurea*, which may be clarified when examined in relation to the genetics and geographic origins of the taxa.

The specimens from the reintroduction efforts on Mount Merapi, Java, are genetically diverse, with the three most common haplotypes found making up the sample, conserving more genetic potential than if the specimens all belonged to the same haplotype. It could be advised that efforts should be made to self-pollinate these plants in order to preserve the genetic potential, rather than to cross-pollinate them with each other whilst they are in cultivation and before they are reintroduced into the national park, and even to reintroduce plants into geographically separate sites in the park to reduce natural cross-pollination.

Harry Wiriadinata at Herbarium Bogoriense; Mrs Elizabeth Handini and Mr Pepen of Bogor Botanic Garden; Dr Djauhar Asikin, Miss Siti Nurfadilah, Mr Tarmudji and Mr Yusuf of Purwodadi Botanic Garden; Mr Milo Migliavacca and Mr Stephen Monkhouse of Milo Bali Adelaide Orchids; Mr Suriyadi and Mr Amin, Bedugal, Bali. Thanks for supervision and assistance in the UK must go to Dr Brent Emerson and Miss Lorenza Legarreta, University of East Anglia; Dr David Roberts, RBG Kew. Additional thanks must go to Mr Peter O'Byrne, Singapore, and Dr Jaap Jan Vermeulen, Leiden.

LITERATURE CITED

Bandelt H. J., P. Forster & A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Molec. Biol. Evol. 16: 37-48.

Chase, M.W. & H.G. Hills. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. Taxon 40: 215-215.

Clement, M., D. Posada & K.A Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9 (10): 1657-1660.

Crandall, K.A. & A.R. Templeton. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134: 959-969.

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

Posada, D., K.A. Crandall & A.R. Templeton. 2000. GeoDis: A program for the cladistic nested analysis of the geographic distribution of genetic haplotypes. Mol. Ecol. 9: 487-488.

Saillard, J., P. Forster, N. Lynnerup, H.J. Bandelt & S. Nørby. 2000. mtDNA variation among Greenland Eskimos: the edge of the Beringian Expansion. Am. J. Human Genetics 67: 718-726.

Swofford, D.L. 2002. PAUP*. Phylogenetic analysis using parsimony (* and other methods) Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Richardson, J.E., R.T. Pennington, T.D. Pennington & P.M. Hollingsworth. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. Science 293: 2242-2245.

Taberlet, P., G. Ludovic, G. Pautou & J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17: 1105-1109.

Templeton, A.R., K.A. Crandall & C.F. Sing. 1992. A cladistic analysis of phenotypic association with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132: 619-633.

Templeton, A.R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Mol. Ecol. 7:

Lauren Gardiner is micropropagation assistant in the Micropropagation Unit at RBG Kew and is currently writing up her PhD on the Phylogenetics and conservation of the genus Vanda. Educated at the Universities of Cambridge, Reading and East Anglia, she is particularly interested in taxonomy and systematics of the Orchidaceae and practical conservation methods.