


ORIGINAL ARTICLE

Open Access



Phylogeography of *Begonia luzhaiensis* suggests both natural and anthropogenic causes for the marked population genetic structure

Yu-Hsin Tseng^{1†}, Han-Yau Huang^{1,2†}, Wei-Bin Xu³, Hsun-An Yang¹, Ching-I Peng^{1^}, Yan Liu^{3*} and Kuo-Fang Chung^{1,2*} 

Abstract

Background: Sino-Vietnamese limestone karsts (SVLK) are a biodiversity hotspot rich in endemic plant species associated with caves and cave-like microhabitats. Based on phylogenetic studies of *Begonia* sect. *Coelocentrum*, a species-rich and characteristic SVLK clade, geographic isolation caused by extensive and continuous karstification was proposed as the major driving force triggering population diversification and geographic speciation. To test this proposition, population genetics and phylogeography of *Begonia luzhaiensis* were investigated using EST-SSR markers and the chloroplast *trnC-ycf6* intergenic spacer.

Results: *F* statistics, Bayesian clustering analysis, AMOVA, and PCoA of both data sets all indicated substantial population differentiation and significant isolation by distance. Nested clade phylogeographic analyses inferred that historical fragmentations have been prominent, congruent with Guangxi's geohistory of karstification as well as suggesting a mountain chain in northeastern Guangxi could have also acted as a major geographic barrier. A Bayesian skyline plot (BSP) indicated a slight decline in effective population size at 75,000 years ago (75 Kya), coinciding with the last glacial period during which the increased aridity in East Asia had retarded karstification, negatively affecting the populations of *B. luzhaiensis*. However, BSP detected a continuous and further population decline until the present time even though summer monsoons have resumed since the end of the last glacial maximum.

Conclusions: The microevolution patterns of *B. luzhaiensis* support that limited gene flow would have greatly enhanced the effects of random genetic drift and has been a major factor promoting diversification in *Begonia*, highly congruent with previous proposition. Based our study, we further propose that the arrival of Paleolithic *Homo sapiens* whose activities centered around limestone caves could have had further impacts on the populations of *B. luzhaiensis*, resulting in additional population decline. Further habitat destruction could have resulted from the transition from hunter gathering to food-producing societies ca. 20–10 Kya and the development of agriculture ca. 10 Kya in South China. Implications of the current study for SVLK plant conservation are also discussed.

*Correspondence: gxiably@163.com; bochung@gate.sinica.edu.tw

†Yu-Hsin Tseng and Han-Yau Huang contributed equally to this work

[^]Ching-I Peng—Deceased

¹ Research Museum and Herbarium (HAST), Biodiversity Research Center, Academia Sinica, Taipei, Taiwan

³ Guangxi Key Laboratory of Plant Conservation and Restoration Ecology in Karst Terrain, Guangxi Institute of Botany, Guangxi Zhuang Autonomous Region and Chinese Academy of Sciences, Guilin, Guangxi, China

Full list of author information is available at the end of the article

Keywords: *Begonia* sect. *Coelocentrum*, Chloroplast marker, EST-SSR, Flora of Guangxi, Glacial periods, Limestone karst conservation, Sino-Vietnamese limestone karsts (SVLK)

Background

Karstification is the geomorphological process whereby soluble and porous bedrock (i.e., limestone, dolomite, gypsum, etc.) is dissolved and infiltrated by water via chemical and mechanical forces through which the surface features and subterranean drainage network of karsts are developed (Day and Urich 2000; Ford and Williams 2007; Williams 2008). In Southeast Asia and South China, more than 800,000 km² of land are limestone karst, characterized by splendid towers [i.e., fengcong/cockpit karst (Waltham 2008)], peaks [i.e., fenglin/tower karst (Waltham 2008)], gorges, depressions, caves, underground tunnels, and subterranean rivers (Fig. 1a–d). Many of these scenic landscapes are not only UNESCO World Heritage Sites attracting tourists from around the world (Fig. 1a) (Day and Urich 2000; Williams 2008), but also major biodiversity hotspots rich in marvelous flora and fauna (Clements et al. 2006; Peng et al. 2014; Xu et al. 2017).

Stretching from South China to northern Vietnam, Sino-Vietnamese limestone karsts (SVLK) are particularly noted for the picturesque landscapes (Fig. 1a) and rich biodiversity with extremely high levels of endemism. In China, 61 of the ca. 250 Chinese endemic plant genera are found only in Guangxi (Qin and Liu 2010), a province known for its spectacular limestone karsts (Sweeting 1978) (Fig. 1a), with a majority of these endemic genera associated with limestone ecosystems (Hou et al. 2010). The limestone flora of Guangxi is especially noted for a suite of plant genera with exceedingly high levels of endemism and restricted distribution almost exclusively confined to the ‘twilight zone’ (Poulson and White 1969) of caves and cave-like microhabitats (Fig. 1c) (Xu et al. 2012; Chung et al. 2014; Monro et al. 2018). These narrowly distributed limestone plants—e.g., *Aspidistra* Ker Gawl., *Begonia* L., the fern genus *Polysticum* Roth, *Impatiens* L., and gesneriad genus *Primulina* Hance, etc.—also present intriguing study systems for investigating plant speciation (Chung et al. 2014). However, the major driving forces that created this plant diversity and the great richness of these narrow-endemic plants in caves of limestone karsts in South China remain poorly understood.

Begonia L. (Begoniaceae), comprising more than 1900 species (Moonlight et al. 2018), is the 5th or 6th largest flowering plant genus distributed in tropical and subtropical Asia, Africa and America. Of the 70 sections of *Begonia* (Moonlight et al. 2018), *Begonia* sect. *Coelocentrum* Irmsch. is a remarkable example of a limestone

plant group with high endemism and species diversity (Gu et al. 2007; Chung et al. 2014; Qin et al. 2017b). Plants of this section (i.e., sect. *Coelocentrum* s.str.), characterized by their parietal placentation, are found in caves and cave-like microhabitats of SVLK (Fig. 1c), with the highest diversity in Guangxi (Gu et al. 2007; Chung et al. 2014). In the past two decades, explorations and taxonomic studies in the region have greatly expanded the number of species in the section from 15 (Ku 1999) to 78 (Chung et al. 2014; Peng et al. 2014, 2015a, b; Li et al. 2016; Qin et al. 2017b; Chen et al. 2018). As a great portion of SVLK remains unexplored, it is fully expected that further exploration will undoubtedly uncover more novelties (Chen et al. 2018).

Based on phylogenetic analyses of Asian *Begonia*, Chung et al. (2014) showed that all sampled species of rhizomatous SVLK *Begonia* are grouped in a strongly supported and yet internally poorly-resolved clade (Clade SVLB) dominated by sect. *Coelocentrum*, indicating its single evolutionary origin, radiation, and adaptation to limestone substrate and suggesting a strong phylogenetic niche conservatism (Wiens 2004) of Clade SVLB. Divergent times estimated by a Yule model in BEAST (Drummond et al. 2012) indicated that the mean divergent crown age of Clade SVLB was estimated to be 8.06 Mya (million years ago), with HPD (highest posterior density) date ranging from 5.09 to 11.21 Mya (Chung et al. 2014), corresponding to the onset of the East Asian monsoon triggered by the coeval uplift of the Himalaya-Tibetan Plateau (An et al. 2001). Geological studies suggest that, since the Miocene, the increased summer precipitation brought by the East Asian monsoon has provided the excessive moisture essential for karstification (Liu 1997; Liu et al. 2001). This geomorphological process that has generated the splendid Sino-Vietnamese limestone karsts would have also created copious cave and cave-like microhabitats suitable for the spread of limestone karst plants such as *Begonia* (Chung et al. 2014). As karstification proceeds, caves develop, collapse, and eventually diminish (Palmer 1991; Ford and Williams 2007), resulting in geographic isolation of limestone karst plants because strong niche conservatism would have forced limestone plants to track their ancestral niche (Wiens 2004; Kozak et al. 2006; Chung et al. 2014). These cave and cave-like microhabitats thus have functioned (Chung et al. 2014) as “paleoreugia, now-fragmented relicts of a formerly widespread matrix community” (Nekola 1999). Moreover, because *Begonia* species are generally poor



Fig. 1 Guangxi's limestone karsts (**a–d**) and habitats of *Begonia luzhaiensis* (**e–h**). **a** The picturesque limestone fengcong karst (Waltham 2008) along the Li River in Guangxi, China is a habitat for many limestone plants, including *B. luzhaiensis*. **b** Aerial photography taken before descending into Guilin Airport, Guangxi, showing the fragmented and cockpit-like karst landscape. **c** The entrance of a limestone cave in Luzhai County, Guangxi where the LZ population (Fig. 2) was sampled. This cave is also the type locality of one rare gesneriad, *Chirita luzhaiensis* (\equiv *Primulina luzhaiensis*) (Huang et al. 2010). **d** Rice terraces in Tian'e County, Guangxi, showing the typical agricultural practices in the region and degraded limestone forest in a limestone karst. **e** *B. luzhaiensis* grows in the limestone cave (Fig. 1c) in Luzhai County (LZ in Fig. 2). **f** *B. luzhaiensis* grows on a limestone outcrop in Xincheng (XC in Fig. 2). **g** *B. luzhaiensis* grows in the entrance of a limestone cave in Rongshui County (RS in Fig. 2). **h** *B. luzhaiensis* grows on the base of a limestone tower along the Li River (YS in Fig. 2)

dispersers (Hughes and Hollingsworth 2008) of low vagility (Kozak et al. 2006; Fouquet et al. 2007) prone to diversify via geographic isolation (Hughes and Hollingsworth 2008), restricted gene flow among conspecific populations could effectively lead to population differentiation (Slatkin 1993). Consequently, geographic fragmentation resulting from continuous karstification and restricted gene flow jointly would promote population differentiation in *Begonia* as well as other limestone karst plants (Gao et al. 2015), resulting in the proliferation of new species manifesting as species radiation in SVLK (Chung et al. 2014). However, because the major driving force of this evolutionary process is geographic isolation, this diversification pattern is better characterized as ‘non-adaptive radiation’ (Gittenberger 1991; Kozak et al. 2006; Rundell and Price 2009), a pattern reported in organisms of low vagility and strong niche conservatism (Gittenberger 1991; Comes et al. 2008).

The microevolutionary scenario postulated by Chung et al. (2014) predicts a marked population genetic structure and phylogeographic pattern in SVLK plant species inhabiting caves and cave-like microhabitats. The scenario predicted by Chung et al. (2014) has been supported by analyses of the *Primulina eburnea* (Hance) Yin Z.Wang species complex that is distributed in limestone karsts of southern China (Gao et al. 2015). To further test these hypotheses, expressed sequence tag-simple sequence repeat (EST-SSR) markers (Tseng et al. 2017) and cpDNA sequences of *trnC-ycf6* intergenic spacer were collected for *Begonia luzhaiensis* T.C.Ku (Fig. 1e–h). *Begonia luzhaiensis* is the only species of sect. *Coelocentrum* s.str. with a relatively wide distribution (Fig. 2) suitable for phylogeographic and population genetic analyses; the rest species of Clade SVLB are known from a single or a few localities except for *B. cavaleriei* H.Lév. and *B. leprosa* Hance that were formerly classified under sect. *Diploclinium* and sect. *Leprosae*, respectively (Ku 1999; Gu et al. 2007). As limestone ecosystems are increasingly threatened worldwide, especially in Southeast Asia (Clements et al. 2006; Hughes 2017) and South China (Monro et al. 2018), the extent of this species’ genetic diversity inferred from the present study will also provide important insights into conservation and management strategies for the precious Sino-Vietnamese limestone karst flora.

Methods

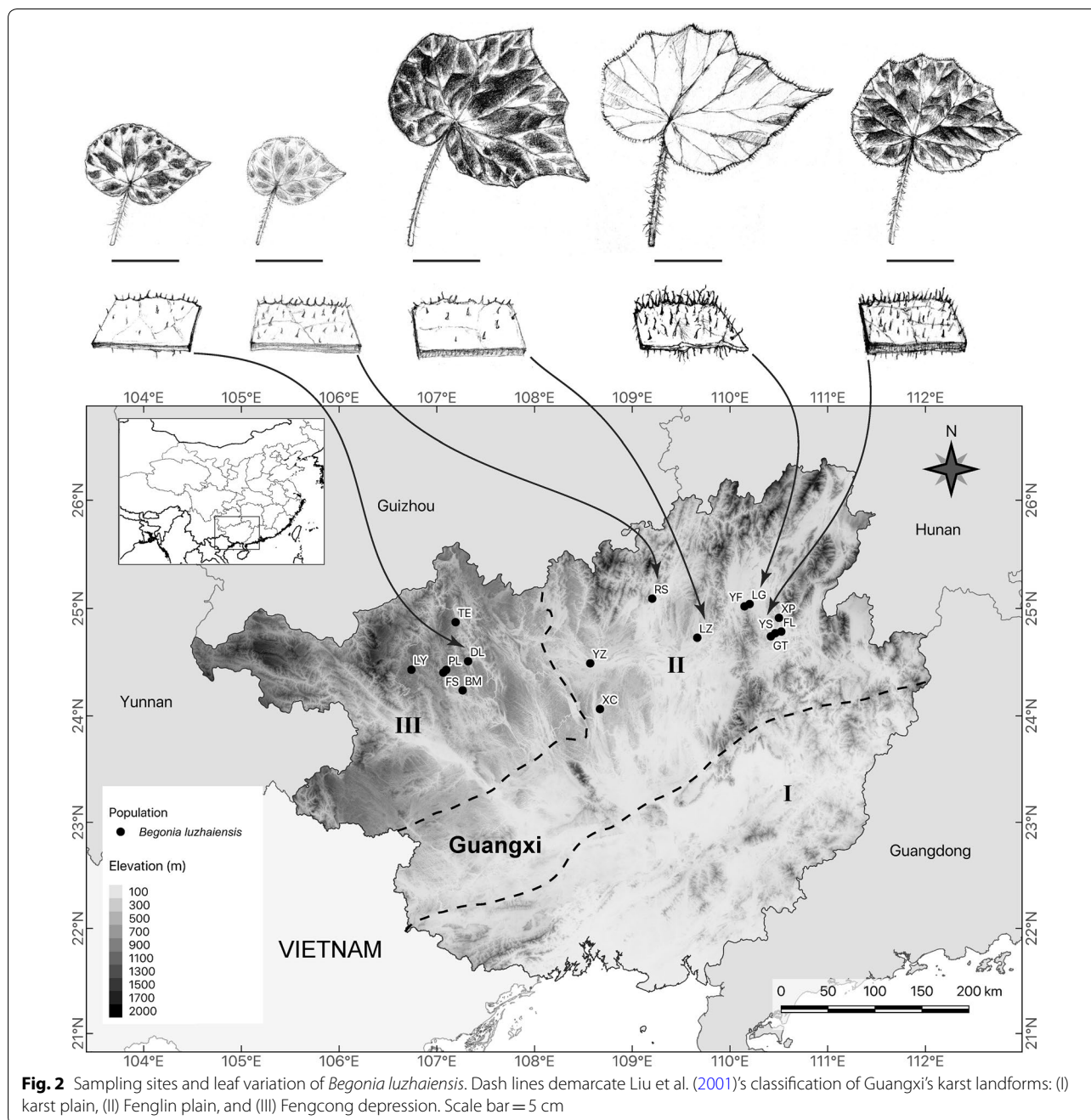
Sample collection

Begonia luzhaiensis is not included in the “Threatened Species List of China’s Higher Plants” (Qin et al. 2017a). For the purpose of scientific research, permits were not required by authorities of the collecting areas. *Begonia luzhaiensis* is endemic to Guangxi (Gu et al. 2007).

Consulting Liu et al. (2001)’s classification of the regional tropical karst landform of Guangxi, the distribution range of *B. luzhaiensis* falls within (II) the Fenglin plain and (III) the Fengcong depression, absent from (I) the karst plain (Fig. 2). *Begonia luzhaiensis* is easily recognizable by its distinct variegation and maculation on leaves (Cui and Guan 2013); however, considerable variation exists across its distribution range (Figs. 1e–h and 2). For example, compared with other populations of *B. luzhaiensis*, plants from the Luzhai (LZ) and Yangshuo (YS) usually have large and hairy leaves, while leaves of plants from Donglan (DL) are darker and less hairy (Fig. 2). Although *B. luzhaiensis* is the only species in *Begonia* sect. *Coelocentrum* s.str. that has a relatively wide distribution range, its population sizes vary greatly among localities, ranging from a few individuals to several hundred plants per site. Based on herbarium records at HAST, IBK, and PE, a total of 165 individuals of *B. luzhaiensis* were collected from 15 localities (Table 1; Fig. 2), covering almost all known populations of the species. One plant collected from Yizhou (YZ) in 2002 and cultivated in the greenhouse of Guangxi Institute of Botany was also included (Fig. 2). Our attempt to sample the Yizhou population failed due to habitat destruction. Fresh leaves were collected and preserved in silica-gel. Voucher specimens were deposited at the Herbarium (HAST) of Biodiversity Research Museum, Academia Sinica, Taiwan (Table 1). *Begonia asteropyrifolia* Y.M.Shui & W.H.Chen, a species closely related to *B. luzhaiensis*, was chosen as an outgroup.

Microsatellite genotyping and DNA sequencing

To overcome difficulty of extracting quality DNA of *Begonia* species, a modified CTAB protocol optimized by Kopperud and Einset (1995) was followed (e.g., Chung et al. 2014). Samples from 15 populations were genotyped by 16 EST-SSR markers using primer pairs developed for *B. luzhaiensis* (Tseng et al. 2017). Amplification and genotyping conditions followed the protocol described in Tseng et al. (2017). Based on our preliminary tests for marker variability as well as efficiency of PCR amplification, the *trnC-ycf6* intergenic spacer of the chloroplast genome was chosen for phylogeographic analysis. Total genomic DNA was extracted from the silica gel-dried leaves using the protocols outlined in Chung et al. (2014). The PCR reaction was performed in 20 μ L of final volumes with 10 μ L of 2 \times Master Mix Red (Ampliqon, Odense, Denmark), 1 μ L of each specific primer (10 μ M each), and 30 ng template. Primers for the *trnC-ycf6* region amplification were “trnC” (5’- CCAGTTCAAATCTGGGTG TC-3’) (Demesure et al. 1995) and “petN1r” (5’-CCC AAGCAAGACTTACTATATCC-3’) (Lee and Wen



2004). The PCR amplifications were initiated with set at: 94 °C for 5 min, 32 cycles of 94 °C for 40 s, 53 °C for 35 s and 72 °C for 1 min and a final extension in 72 °C for 5 min. PCR products were purified using QIAquick PCR purification Kit (Qiagen, Valencia, California, U.S.A.) and sequenced using an ABI PRISM dye terminator cycle sequencer, model 3700 (Applied Biosystems, Foster City, California, U.S.A.). All sequences were deposited in GenBank with accession numbers MG063292–MG063424.

Analysis of EST-SSR

The allelic richness (A_R), gene diversity (H_S), observed heterozygosity (H_O), expected heterozygosity (H_E), and inbreeding coefficient (F_{IS}) of EST-SSR markers were computed using FSTAT 2.9.3 (Goudet 2001). Tests for deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) for each EST-SSR loci were estimated by GENEPOP 4.2 (Rousset 2008). For each locus, the FDIST method (Beaumont and Nichols 1996) implemented in LOSITAN (Antao et al. 2008) was used

Table 1 Voucher information for samples used and GenBank accession numbers of cpDNA sequences

Population (Code)	Voucher (HAST specimen ID)	Collection locality	Geographic coordinates	GenBank accession number
<i>Begonia luzhaiensis</i> T.C.Ku				
Bama (BM)	S.-M. Ku 2022 (HAST 142361)	Bana Village, Xishan Township, Bama County, Guangxi, China	24°14'17"N 107°15'50"E	MG063376
Donglan (DL)	S.-M. Ku 2002 (HAST 142357)	Dawei Village, Donglan Town, Donglan County, Guangxi, China	24°30'33.0"N 107°19'13.0"E	MG063379, MG063380
Fengshan (FS)	S.-M. Ku 2016 (HAST 142358)	Paoli Township, Fengshan County, Guangxi, China	24°24'14.0"N 107°04'00.0"E	MG063378
Fuli (FL)	H.-Y. Huang et al. 6 (HAST 142351)	Fuli Town, Yangshuo County, China	24°47'08.2"N, 110°31'22.5"E	MG063293–MG063297
Gaotian (GT)	K.-F. Chung et al. 1855 (HAST 142364)	Paitou Village, Gaotian Town, Yanshuo County, Guangxi, China	24°44'25.8"N 110°25'04.9"E	MG063369–MG063373
Lingui (LG)	H.-Y. Huang et al. 31 (HAST 142353)	Sishan Village, Huixian Town, Lingui County, Guangxi, China	25°02'27.9"N 110°11'59.4"E	MG063323–MG063327
Lingyun (LY)	C.-I Peng et al. 24344 (HAST 138438)	Luolou Town, Lingyun County, Guangxi, China	24°25'48"N, 106°44'24"E	MG063381–MG063415
Luzhai (LZ)	H.-Y. Huang et al. 18 (HAST 142354)	Zhongdu Town, Luzhai County, Guangxi, China	24°43'37.9"N 109°39'50.6"E	MG063311–MG063322
Paoli (PL)	S.-M. Ku 2018 (HAST 142359) & 2019 (HAST 142360)	Paoli Township, Fengshan County, Guangxi, China	24°25'43.0"N 107°05'46.0"E	MG063374, MG063375
Rongshui (RS)	H.-Y. Huang et al. 33 (HAST 142355)	Xidong Scenic Area, Rongshui County, Guangxi, China	25°05'30.2"N 109°12'10.7"E	MG063334–MG063347
Tian'e (TE)	K.-F. Chung et al. 1825 (HAST 142363)	Dulou Village, Bamu Township, Tian'e County, Guangxi, China	24°52'21.2"N 107°11'31.0"E	MG063366–MG063368
Xincheng (XC)	H.-Y. Huang et al. 102 (HAST 142356)	Chengguan Town, Xincheng County, Guangxi, China	24°03'49.5"N 108°40'08.9"E	MG063348–MG063365
Xingping (XP)	K.-F. Chung et al. 3017 (HAST 142365)	Xingping Town, Yangshuo County, Guangxi, China	24°54'48.3"N 110°30'02.9"E	MG063416–MG063424
Yongfu (YF)	C.-I Peng et al. 24374 (HAST 138467)	Luojin Township, Yongfu County, Guangxi, China	25°01'12.0"N 110°09'02.0"E	MG063328–MG063333
Yangshuo (YS)	H.-Y. Huang et al. 8 (HAST 142352)	Yanshuo Town, Yanshuo County, Guangxi, China	24°46'12.9"N, 110°27'43.4"E	MG063298–MG063310
Yizhou (YZ)	Y.-M. Shui 9010 (HAST 94959)	Yankou Village, Yizhou District, Guangxi, China	24°29'24.0"N, 108°34'12.0"E	MG063377
<i>Begonia asteropyrifolia</i> Y.M.Shui & W.H.Chen				
(Outgroup)	C.-I Peng et al. 21062 (HAST 117107)	Donglan Town, Donglan County, Guangxi, China	–	MG063292

All voucher specimens are deposited in HAST. The HAST specimen ID is encrypted with the hyperlink to its digital image

to test for selection. Haplotype-based genetic differentiation (F_{ST}) between populations was calculated using permutations in ARLEQUIN 3.5.1.2 (Excoffier and Lischer 2010).

The population structure was evaluated by the software STRUCTURE 2.3.4 (<https://web.stanford.edu/group/pritchardlab/structure.html>) (Pritchard et al. 2000) and principal coordinate analysis (PCoA) using GENALEX 6.5 (Peakall and Smouse 2012). In STRUCTURE, the possible number of clusters (K) was set to vary from 1 to 15 with 10 independent runs. The analysis was performed assuming the admixture model and correlated allele frequencies among clusters. Each run consisted of 1,000,000 steps Markov chain Monte

Carlo (MCMC) replicates after a 100,000 steps burn-in (Evanno et al. 2005). The best grouping was evaluated by ΔK (Evanno et al. 2005) in STRUCTURE HARVESTER v 0.6.94 (Earl and Vonholdt 2012). Alignment of cluster assignments across replicate analyses was conducted in CLUMPP (Jakobsson and Rosenberg 2007). In PCoA, Nei 's pairwise genetic distances were calculated and visualized by PCoA using GENALEX (Peakall and Smouse 2012).

To calculate genetic variation, populations of *B. luzhaiensis* were grouped into East (FL, GT, LG, XP, YF and YS), Central (LZ and RS), and West (BM, DL, FS, LY, PL, TE, XC and YZ) regions based on the grouping results of STRUCTURE and PCoA (Fig. 3). The analysis

of molecular variance (AMOVA) was used to test genetic variation between areas (groupings based on the results of Structure and PCoA) and among and within populations. A non-grouped analysis was also conducted. The AMOVA was also assessed by ARLEQUIN using R -statistics with significance tests by 10,000 non-parametric random permutations (Excoffier and Lischer 2010).

To test for the assumption of isolation by distance (IBD), a Mantel test was performed using genetic and

geographic distances. Pairwise F_{ST} between populations used as the genetic distance metric was calculated based on Nei's approach (Nei et al. 1977) using GENALEX. Geographic distances between location pairs were calculated using Geographic Distance Matrix Generator v1.2.3 (Ersts 2011). Mantel tests and Mantel correlogram test were carried out by regressing $F_{ST}/(1 - F_{ST})$ against the natural logarithm of geographical distance using

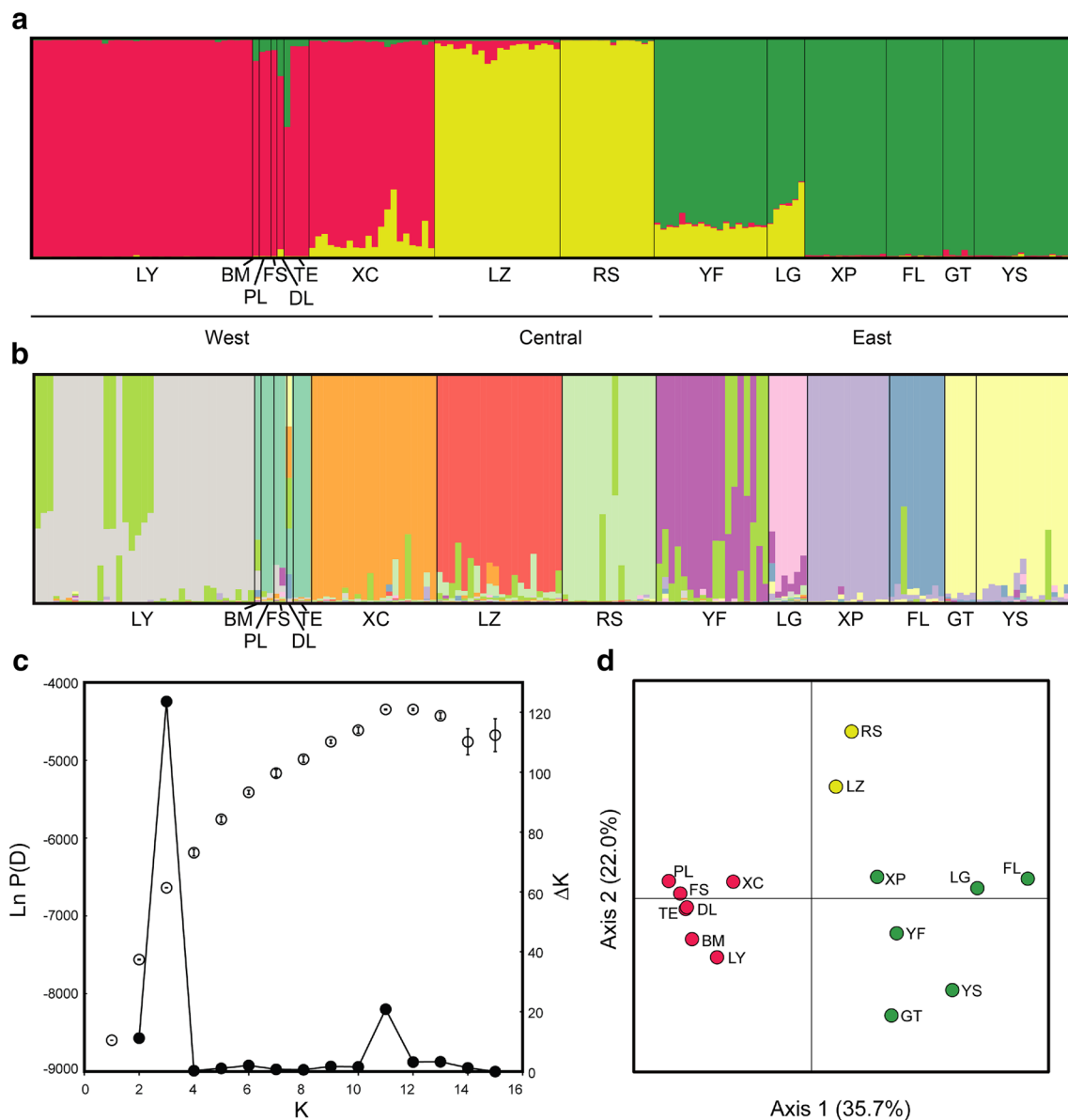


Fig. 3 The results of STRUCTURE & PCoA. **a** STRUCTURE bar plot representing $K=3$ of the Bayesian assignment of 165 individuals from 15 populations of *Begonia luzhaiensis* based on genetic variation of 16 EST-SSR loci. Each vertical bar represents an individual plant is represented by a single vertical line grouped by populations. **b** STRUCTURE bar plot representing $K=11$. **c** The ΔK plot shows that $K=3$ is the highest ΔK value. **d** PCoA analysis is based on EST-SSR loci using GENALEX

the vegan package in R (Oksanen et al. 2010) with 9999 permutations.

Analysis of *trnC-ycf6* spacer data

Raw sequences of *trnC-ycf6* spacer were edited and assembled by SeqMan™ II (DNASTAR, USA) with subsequent manual adjustments. DNA sequences were aligned by MUSCLE v3.6 (Edgar 2004) implemented in MEGA 5 (Tamura et al. 2011) using a first round of multiple alignment and posterior rounds of refinement under default setting. Alignments were subjected to manual adjustments in Mesquite v.3.0.3 (Maddison and Maddison 2018). Indels were treated as the fifth character when calculating the indices of genetic diversity. The program DnaSP v5.0 (Librado and Rozas 2009) was used to calculate the degree of polymorphism (π), haplotype diversity (h), and θ estimated by segregating sites (S) for populations with at least three samples (Table 2).

Haplotype-based genetic differentiation (F_{ST}) between populations was calculated and its significance was tested using permutations in ARLEQUIN (Excoffier and Lischer 2010). To determine whether the genetic distance between populations was determined by geographic distance, isolation by distance analysis was performed

using GENALEX (Peakall and Smouse 2012) as described above for EST-SSR. The analysis of molecular variance (AMOVA) was performed by ARLEQUIN to estimate the partitioning of genetic variation among areas (groupings based on STRUCTURE and PCoA's results) within populations and among populations.

To estimate the dynamics of population size fluctuations over time, the Bayesian Skyline Plot (BSP) method, a coalescent-based approach utilized standard MCMC integration to evaluate the posterior probability distribution of effective population size, was performed to estimate the historical demographic change of *B. luzhaiensis* using BEAST v1.8.4 (Drummond et al. 2012). Some recent studies (Chikhi et al. 2010; Heller et al. 2013) showed that population structure tends to produce a false signal of population bottleneck with some sampling schemes in demographic history inferences, but different sampling strategies can help to overcome this structure effect (Chikhi et al. 2010; Heller et al. 2013; Mazet et al. 2015, 2016). Therefore, we used four sampling strategies based on Heller et al. (2013) and Sgarlata et al. (2018) to perform the historical demographic change in BSP, including (1) local sampling: all samples from a spatial cluster (eastern, central, western regions based on the grouping results of PCoA and STRUCTURE), (2) pooled

Table 2 Genetic diversity and demographic characteristic of *Begonia luzhaiensis* based on EST-SSR markers and *trnC-ycf6* spacer sequence

Population code	<i>n</i> (SSR/cpDNA)	SSR					cpDNA			
		A_R	H_s	F_{IS}	H_o	H_e	π	h	θ_w	Haplotype (<i>n</i>)
East Guangxi										
Fuli (FL)	9/5	1.552	0.553	0.039	0.531	0.518	0.000	0.000	0.000	H2 (5)
Gaotian (GT)	5/5	1.478	0.472	-0.106	0.522	0.429	0.000	0.000	0.000	H4 (5)
Lingui (LG)	6/5	1.371	0.383	0.228	0.296	0.336	0.171	0.600	0.137	H3 (2), H5 (3)
Xingping (XP)	13/9	1.266	0.266	0.080	0.245	0.255	0.000	0.000	0.000	H1 (9)
Yongfu (YF)	18/6	1.261	0.264	0.212	0.208	0.251	0.000	0.000	0.000	H3 (6)
Yangshuo (YS)	15/13	1.475	0.476	0.073	0.441	0.459	0.000	0.000	0.000	H2 (13)
Central Guangxi										
Luzhai (LZ)	20/12	1.482	0.487	0.187	0.396	0.455	0.000	0.000	0.000	H6 (12)
Rongshui (RS)	15/14	1.326	0.328	0.142	0.282	0.313	0.000	0.000	0.000	H7 (14)
West Guangxi										
Bama (BM)	1/1	–	–	–	–	–	0.000	0.000	0.000	H8 (1)
Donglan (DL)	1/2	–	–	–	–	–	0.142	1.000	0.142	H6 (1), H9 (1)
Fengshan (FS)	1/1	–	–	–	–	–	0.000	0.000	0.000	H9 (1)
Lingyun (LY)	35/35	1.236	0.237	0.201	0.189	0.232	0.000	0.000	0.000	H8 (35)
Paoli (PL)	2/2	1.073	0.067	-0.500	0.094	0.055	0.000	0.000	0.000	H9 (2)
Tian'e (TE)	4/3	1.245	0.266	0.569	0.115	0.212	0.000	0.000	0.000	H8 (3)
Xincheng (XC)	20/18	1.294	0.295	0.101	0.265	0.285	0.000	0.000	0.000	H6 (18)
Yizhou (YZ)	-/1	–	–	–	–	–	0.000	0.000	0.000	H6 (1)

n (SSR/cpDNA), number of individuals sampled for SSR/cpDNA; A_R : mean allelic richness corrected for sample size; H_s : gene diversity; F_{IS} : inbreeding coefficient; H_o : observed heterozygosity; H_e : expected heterozygosity; π : nucleotide diversity; h : haplotype diversity

sampling: three to five randomly picked samples from each population, excluding populations with only one or two individuals, (3) scattered sampling: one randomly picked sample from each population, and (4) structured sampling: all samples. For each sampling strategy, the substitution rate of *trnC-ycf6* spacer was set as 1.01×10^9 (per site per year) referring to previous reports of cpDNA for herbaceous plants with a normal distribution (Graur and Li 2000). Based on Akaike information criteria (AIC), the J-Model test (Posada 2008), calculated through the CIPRES web-server (<http://www.phylo.org/index.php/portal/>), selected HKY as the most appropriate substitution model for *trnC-ycf6*. Prior distribution for Bayesian skyline, population parameter sizes were uniform, with an initial value of 1×10^5 , a lower value of 100 and upper value of 2×10^6 . Two independent MCMC runs were conducted in the analyses, each with 100 million generations and the trees were sampled every 1000 generations, with first 25% discard as burn-in. Time series plots of all parameters were analyzed in Tracer v.1.5 (Rambaut et al. 2018) to check for effective sample size values that were larger than 200.

Demographic history was explored using two different methods. First, Tajima's *D* (Tajima 1989) and Fu's *F_s* (Fu 1997) were calculated using ARLEQUIN to infer demography by testing deviations from the neutral equilibrium condition. Second, the mismatch distribution was implemented to identify the population expansion model using ARLEQUIN. We estimated demographic expansion factor (τ), mutation parameter (θ), and number of migrants (*M*) at 95% confidence intervals. The sum of squared deviations (*SSD*) was estimated to test the goodness-of-fit of the observed and expected mismatch distributions (Schneider and Excoffier 1999) and significance was assessed after 10,000 permutations. The expansion time (*T*) was computed with the equation $T = \tau/2u$, where *u* is the mutation rate for the entire sequence. The value of *u* is derived from $u = \mu m$, where the μ is the mutation rate per site per generation, and *m* is the number of investigated nucleotides. The generation time of *B. luzhaiensis* was assumed to be 2 years based on our observation in the experimental greenhouse of the Academia Sinica.

To test the null hypothesis of no association between geography and haplotype variation, nested clade phylogeographic analysis (NCPA) was performed (Templeton et al. 1995, 2009; Buhay et al. 2009; Santos et al. 2017; Bekker et al. 2018), with cautions taken from criticisms against NCPA (e.g., Knowles 2008). Haplotype network was reconstructed using TCS v.1.21 with the 95% parsimony criterion (Clement et al. 2000) using *B. asteropyrifolia* as an outgroup. A series of multi-level nested clade designs was defined following the rules of Templeton

and Sing (1993) based on the network constructed by TCS. Subsequently, the program GeoDis (Posada et al. 2000) was applied to calculate the clade distance (*D_c*), the nested clade distance (*D_n*), and interior-tip statistic (I-T) to test the null hypothesis of no geographic association of clades and nested clades, with 95% confidence level and 10,000 permutations. These measures were interpreted by referring to the latest inference key from GeoDis website (<http://darwin.uvigo.es/software/geodis.html>).

Result

Sample collection

A total of 165 plants from 15 populations were successfully genotyped using EST-SSR markers and 131 alleles were detected across the 16 loci (Table 2).

Analyses of EST-SSR

The mean number of alleles per locus ranged from 1.1 (BM, FS, and PL) to 3.9 (YS) with a mean of 2.3 ± 0.9 . Within populations, the allelic richness (*A_p*) ranged from 1.073 (PL) to 1.552 (FL) and gene diversity (*H_s*) from 0.067 (PL) to 0.553 (FL). Expected heterozygosity was highest in population FL (0.518) and lowest in population PL (0.055), with a mean of 0.317 ± 0.131 (Table 1). The global *F_{ST}* based on EST-SSR loci was 0.527 ($P < 0.001$) (Additional file 1: Table S1) and pairwise *F_{ST}* between populations were also high (0.315–0.676) (Additional file 1: Table S3).

Prior to estimating further population genetic parameters, neutrality tests for outlier analysis for EST-SSR was conducted by FDI_{ST}. Because the null hypothesis of neutrality was not rejected by the result of outlier analysis, the EST-SSR markers used in this study were used for further population genetic assessment (Additional file 1: Table S1).

The ΔK statistic (Evanno et al. 2005) of the STRUCTURE analyses had a peak at *K*=3, with a second and much lower peak at *K*=11 (Fig. 3c). At *K*=3, STRUCTURE identified clusters corresponding well to East Guangxi (YS, GT, FL, XP, LG, and YF), Central Guangxi (RS and LZ) and West Guangxi (XC, TE, DL, FS, PL, BM, and LY), with some admixtures detected in LG, YF, and LZ, and XC (Fig. 3a). This geographic clustering also corresponded well to the PCoA plot (Fig. 3d), in which the first and second axes extracted 35.7% and 22.0% of the total genetic variation, respectively. At *K*=11, STRUCTURE identified 10 clusters, breaking the East Guangxi into five clusters (YS+GT, FL, XP, LG, and YF), Central Guangxi into two clusters (RS and LZ), and the West Guangxi into three clusters (XC, TE+DL+FS+PL+BM, and LY) with most clusters showing various degrees of admixture (Fig. 3b).

Analysis of *trnC-ycf6* spacer data

For chloroplast *trnC-ycf6* spacer, DNA sequences of 132 plants from 16 populations were collected. The length of sequences ranged from 567 to 572 bp, with the aligned length of 601 bp. The sequence length of the outgroup species *B. asteropyrifolia* was 567 bp. Among all the populations examined, nucleotide diversity ranged from 0 (FL, YS, GT, YF, XP, LZ, RS, FS, PL, BM, TE, XC, YZ, LY) to 0.002 (LG, DL) and haplotype diversity (*h*) from 0 (FL, YS, GT, YF, XP, LZ, RS, FS, PL, BM, TE, XC, YZ, and LY) to 0.6 (LG) (Table 2). Pairwise F_{ST} values between populations were low (0–0.009, Additional file 1: Table S3).

Nine haplotypes of *trnC-ycf6* spacer (H1 to H9) in *B. luzhaiensis* were identified from 16 populations (Table 1; Fig. 4). Most populations have only one haplotype except for DL and LG that possess two shared haplotypes (DL: H6 and H9; LG: H3 and H5). Haplotype H1, H4 and H7 were distributed exclusively in population XP, GT and RS, respectively. In the statistical parsimony network,

haplotypes were inter-connected by one or two mutations steps. Haplotype H6, identified as the ancestral haplotype for its highest outgroup probability (outgroup weight=0.338) and connection with the outgroup *B. asteropyrifolia*, is distributed in central area (DL, LZ, XC, and YZ) of the distribution range of *B. luzhaiensis*, connected to the haplotypes of western (H8, H9) and eastern (H1 to H5) populations and the haplotype H7 in RS to the north (Fig. 4).

Population structure and phylogeographic analyses

The hierarchical analysis of molecular variance (AMOVA) in non-grouping analysis revealed the higher variations among populations than within population both in EST-SSR data (56.28%, $P < 0.001$; 43.72%, $P < 0.001$) and cpDNA sequence (63.46%, $P < 0.001$; 36.54%, $P < 0.001$). However, in the grouping analysis, the highest proportion of molecular variance in the EST-SSR data was found among grouping

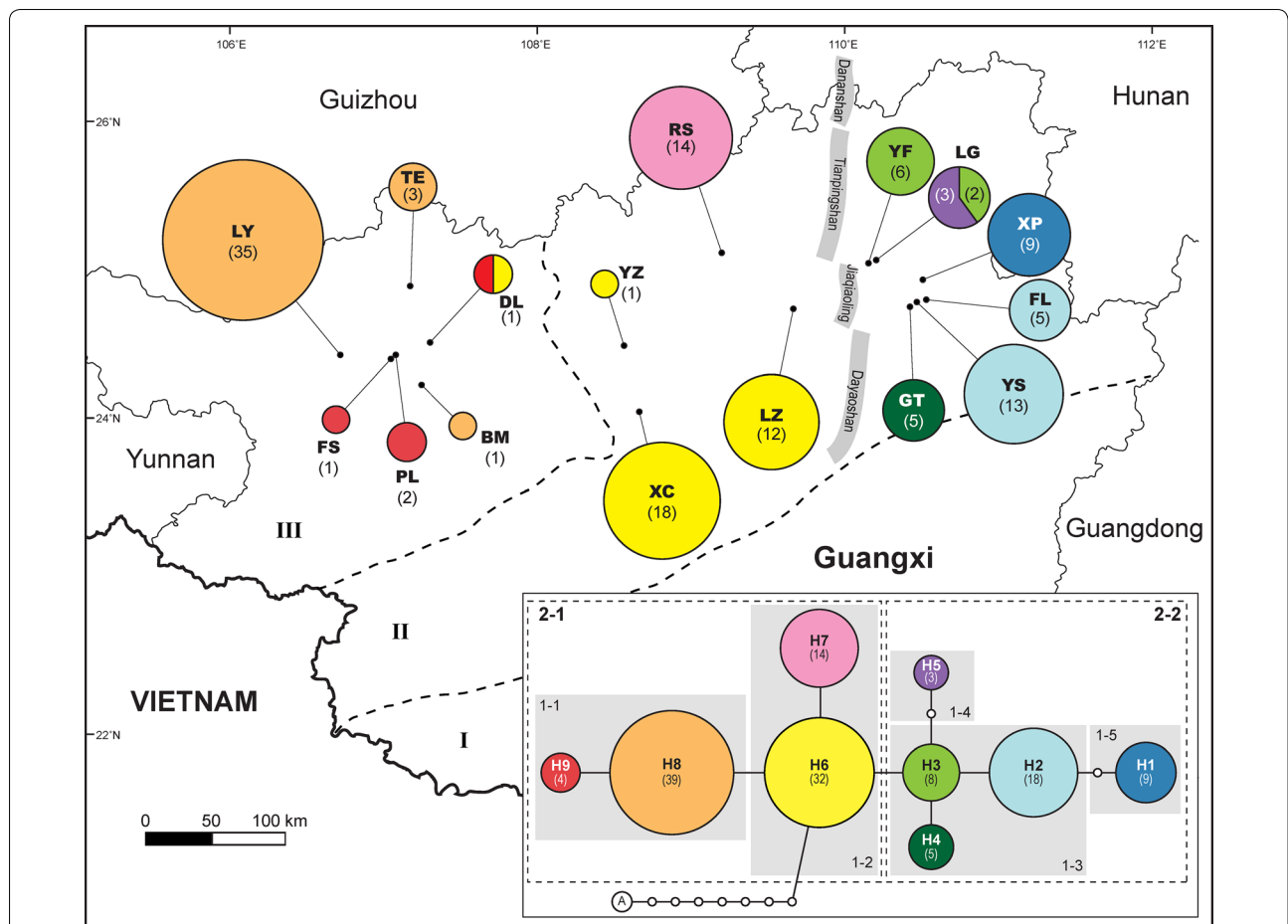


Fig. 4 Distribution of *trnC-ycf6* haplotypes, haplotype network, and nested cladistic design in *B. luzhaiensis*. Statistical parsimony network was generated by TCS 1.21 rooted by *B. asteropyrifolia* (A). The size of circles is proportional to haplotype frequency, which is also indicated in the circles. Dash lines demarcate three types (I, II, and III in Fig. 2) of karst landforms in Guangxi based of Liu et al. (2001)'s classification

(48.32%, $P < 0.001$), while the highest variance in cpDNA sequence was among population (51.89%, $P < 0.001$) (Table 3).

The value of Tajima's D and Fu's F were not significantly different from 0 in the total sample and for each region (Table 4). The mismatch distribution analysis revealed population expansion events in *B. luzhaiensis* based on the non-significant SSD (0.002, $P = 0.758$). Given the expansion parameter ($\tau = 1.161$) and equation $T = \tau/2u$, populations of *B. luzhaiensis* were estimated to expand 2.047 Mya (95% confidence interval = 0.596–6.072 Mya) (Table 5). A similar scenario was observed in the unimodal distribution of pairwise nucleotide differences (Fig. 5a), which also suggested that populations experienced demographic expansions (Additional file 2: Fig. S1).

For four of the six sampling strategies [structured (Fig. 5b), east in local (Additional file 3: Fig. S2A), pooled (Additional file 3: Fig. S2D), and scattered (Additional file 3: Fig. S2E)], the BSP results showed that populations of *B. luzhaiensis* experienced a long history of constant population size since the late Pleistocene until ca. 75 Kya when slight decrease was detected, with continuous and further decline until the present time (Fig. 5b and Additional file 3: Fig. S2D). For western and central regions in local sampling, the results showed a constant population size within the last 400 years (Additional file 3: Figs. S2B and C). Mantel test showed that there was an overall significant positive correlation between genetic (Additional file 1: Tables S3 and S4) and geographical distances (Additional file 1: Table S5) in both cpDNA and EST-SSR, and the signal was stronger for SSR ($R_{xy} = 0.589$, $P < 0.001$) than for cpDNA ($R_{xy} = 0.38$, $P < 0.01$) (Table 6).

The statistics of the nested clade phylogeographic analyses (NCPA) is summarized in Table 7. Based on NCPA, the null hypothesis of no geographic association of haplotype variation was rejected in all six clades with geographic and genetic variation (Fig. 4; Table 7). By applying the inference key, historical events appear to be the major mechanisms shaping the phylogeography of *Begonia luzhaiensis*, with allopatric fragmentation detected in Clades 1-2, 1-3, and the total cladogram, and past fragmentation and/or long distance colonization in Clades 2-1 and 2-2. In Clade 1-1, restricted gene flow with isolation by distance is inferred (Table 7).

Discussion

Our analyses using EST-SSR and cpDNA *trnC-ycf6* spacer reveal marked population genetic structure and phylogeographic pattern in *Begonia luzhaiensis*, highly

concordant with those predicted by Chung et al. (2014). Analyses of EST-SSR (Tables 2 and 3) and cpDNA sequences (Tables 2, 3 and 4) indicated a high level of genetic differentiation and apparent substructure in *B. luzhaiensis* (Fig. 3), suggesting restricted gene flow among isolated populations, as documented in previous studies of *Begonia* (Matolweni et al. 2000; Hughes and Hollingsworth 2008; Nakamura et al. 2014; Twyford et al. 2014; Chan et al. 2018; Li et al. 2018).

Population diversification in *Begonia luzhaiensis*

Hughes and Hollingsworth (2008) suggest that limited gene flow in *Begonia* could have resulted in several geographic patterns, including geographical restrictions of monophyletic groups, frequent occurrences of narrow endemic species, and relatively few widespread species. Additionally, limited gene flow may greatly enhance effects of random genetic drift, a major mechanism promoting diversification in other large plant genera such as the legume genus *Astragalus* L. (Sanderson and Wojciechowski 1996). Moreover, widespread *Begonia* species show strong population substructure (Matolweni et al. 2000; Hughes and Hollingsworth 2008; Twyford et al. 2013; Chan et al. 2018) and decreased pollen viability in offspring of artificial crossing between isolated populations within species (Twyford et al. 2013, 2014). Here, our observations give further lines of support for this model of population diversification in *Begonia*. *Begonia luzhaiensis*, one of the few widespread species of Clade SVLB, can be grouped into three subpopulations in western, eastern and central Guangxi based on EST-SSR loci (Fig. 3). Given the highly fragmented nature of the cave and cave-like microhabitats (Fig. 1c) across the Sino-Vietnamese limestone karsts inhabited by *B. luzhaiensis*, our results demonstrate that geographic isolation could have played a pivotal role in generating the strong population differentiation of the Clade SVLB as proposed by Chung et al. (2014).

Although the dispersal biology of *Begonia luzhaiensis* has not been investigated, the life history trait and habitat preference of the species fits well with the typical *Begonia*, which is generally a poor seed disperser in sheltered conditions of the forest floor (De Lange and Bouman 1999a, b; Dewitte et al. 2011). Species of *Begonia* sect. *Coelocentrum* inhabit the limestone caves and cave-like microhabitats, which are isolated, closed and have high humidity (Liang et al. 2010). Such habitat preference indicates a great disadvantage for dispersal success in sect. *Coelocentrum*, further limiting effective gene flow.

Given the low dispersal ability and strong niche conservatism of the limestone *Begonia* in the highly

Table 3 Analysis of molecular variance (AMOVA) for *Begonia luzhaiensis* based on EST-SSR loci and *trnC-ycf6* spacer sequence (cpDNA)

Source of variation	EST-SSR				cpDNA					
	d.f.	Sum of squares	Variance components	Fixation indices	Percentage of variation	d.f.	Sum of squares	Variance components	Fixation indices	Percentage of variation
(a) Non-grouping										
Among populations	2	196.28	0.932		56.288*	2	65.36	0.782		63.46**
Within populations	321	232.25	0.724	$F_{st} = 0.563$	43.72**	128	57.66	0.450	$F_{st} = 0.635$	36.54**
Total	323	429.53	1.656			130	123.02	1.232		
(b) Grouping^a										
Among groups	2	196.28	0.798	$F_{ct} = 0.483$	48.32**	2	65.36	0.563	$F_{ct} = 0.457$	44.73**
Among populations within groups	9	81.32	0.369	$F_{sc} = 0.433$	22.38**	13	54.26	0.640	$F_{sc} = 0.956$	51.89**
Within populations	312	150.94	0.484	$F_{st} = 0.707$	29.30*	116	3.40	0.029	$F_{st} = 0.976$	2.38**
Total	323	428.53	1.651			130	123.02	1.233		

F_{st} : fixation index within population; F_{sc} : fixation index among populations within groups; F_{ct} : fixation index among groups

* $P < 0.05$, ** $P < 0.001$

^a Grouping based on Structure and PCoA's results

fragmented landscape of SVLK, Chung et al. (2014) also predicted restricted gene flow constrained by isolation due to distance in SVLB, a model documented in other low vagility organisms such as amphibians (Kozak et al. 2006; Fouquet et al. 2007). In the current study, Mantel tests of both EST-SSR and cpDNA *trnC-ycf6* spacer sequences detect significant positive correlation between genetic and geographic distances (Table 6), congruent with the expectation of isolation by distance in *B. luzhaiensis*. Similarly, NCPA also infers restricted gene flow with isolation by distance as the major mechanism in shaping the phylogeographic structure in populations of western Guangxi (Clade 1-1) (Table 7).

Causes of phylogeographic patterns

NCPA inferred that phylogeography of *Begonia luzhaiensis* has been shaped predominately by historical fragmentation events, with the exception of Clade 1-1 that “restricted gene flow with isolation by distance” was inferred (Table 7). Liu et al. (2001) distinguishes three karst landforms in Guangxi—karst plain, Fenglin

plain, and Fengcong depression (Figs. 2 and 4)—that represent different stages of karstification. Karst plain (I in Figs. 2 and 4) is characterized by scattered residual limestone hills, symbolizing the oldest and last stage of karstification. Fenglin plain (II in Figs. 2 and 4), characterized by isolated, steep-side towers and peaks raising from the plain covered by alluvium and rice paddies, represent the mid stage of Guangxi’s karstification. Fengcong depression (III in Figs. 2 and 4) is the youngest of Guangxi’s karst landform, characterized by conical hills separated by deep, closed depressions. Referring to Liu et al. (2001)’s classification of Guangxi’s karst landforms, haplotypes and clades with inferences of historical fragmentation are predominately distributed in Fenglin plain (II in Fig. 4) where limestone peaks are isolated and largely fragmented, while Clade 1-1 (“restricted gene flow with isolation by distance” inferred by NCPA) is found in Fengcong depression (III in Fig. 4) where depressions separating limestone hills could still function as migration corridors for gene flow. On the other hand, *B. luzhaiensis*

Table 4 Tajima’s D and Fu’s F_s statistics for *Begonia luzhaiensis* based on *trnC-ycf6* spacer sequence

	East Guangxi	North Guangxi	West Guangxi	Total
Tajima’s D (P value)	0.403 (0.706)	1.591 (0.977)	− 0.540 (0.244)	0.485 (0.642)
Fu’s F_s (P -value)	1.167 (0.737)	1.619 (0.745)	0.728 (0.634)	1.172 (0.705)

Table 5 Estimated spatial expansion parameters of τ , θ and M for *Begonia luzhaiensis* based on *trnC-ycf6* spacer sequence

τ	θ	M	SSD (P -value)	Time of expansion (Ma)
1.161 (0.338–3.443)	0.361 (0.001–1.333)	6.625 (0.333–1047.893)	0.002 (0.758)	2.047 (0.596–6.072)

These parameters were obtained from the mismatch distribution under the spatial expansion model. Value in parentheses indicate 95% confidence intervals
SSD: Sum of squared deviation

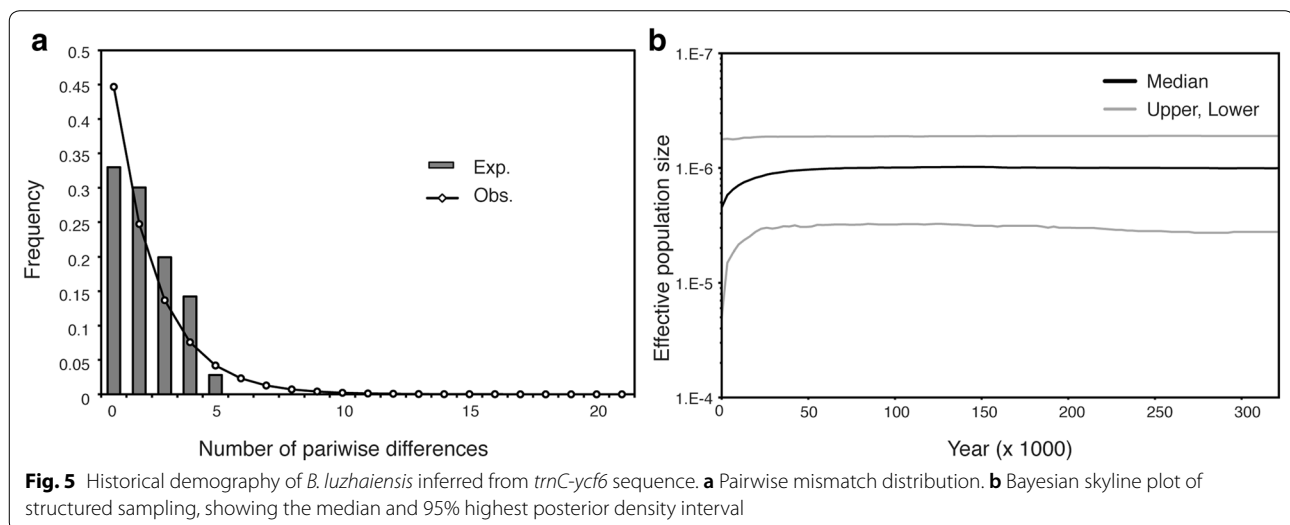


Table 6 Summary for Mantel tests

	R_{xy}	P
EST-SSR	0.589	0.001
cpDNA	0.38	0.001

is absent from karst plains where limestone hills are highly degraded. Taken together, the phylogeography of *B. luzhaiensis* appears to be correlated well with the karst landform in Guangxi, suggesting the importance of the geohistory of karstification in determining the population genetic structure of *B. luzhaiensis* and thus supporting Chung et al. (2014)'s proposition. While continuous karstification has apparently been crucial in determining the phylogeographic structure of *B. luzhaiensis*, the allopatric fragmentation inferred for the total cladogram is also in line with the mountain chain [i.e., Dananshan, Tianpingshan, Jiaqiaoling, and Dayaoshan (Fig. 4)] that separates Clades 2-1 and 2-2, suggesting that these none-limestone mountains could have created further geographic barriers for gene flow among the fragmented karst landscapes.

Implications for SVLK plant speciation

Within *B. luzhaiensis*, populations from different caves or cave-like microhabitats have very similar flowers, and yet they display a considerable range of variation in shapes, sizes and variegation in leaves across the species' distributional range (Figs. 1e–h and 2). This pattern is congruent with other widespread *Begonia* species

that all show high levels of morphological differentiation in different populations (Matolweni et al. 2000; Hughes and Hollingsworth 2008; Twyford et al. 2013, 2014; Chan et al. 2018). Dewitte et al. (2011) proposed that the high morphological diversity in *Begonia* could have been explained by adaptation to a variety of ecological conditions. However, although shapes and sizes of each cave or cave-like microhabitat seem quite different from one another across the SVLK landscape, major ecological parameters such as temperate, light intensity, humidity, and soil substrate appear to be highly homogeneous (Chung et al. 2014; Monro et al. 2018), characterized by constant temperature, high humidity, and indirect and low light (Liang et al. 2010; Ren et al. 2010; Monro et al. 2018). Additionally, recent studies also show no apparent adaptive difference between *Begonia* species that show different leaf variegation, including several species of Clade SVLB (Zhang et al. 2009; Sheue et al. 2012). Given the marked population genetic structure and phylogeographic pattern, our data also suggest that genetic drift alone could have been sufficient in generating intraspecific leaf variation in *B. luzhaiensis* as shown in other *Begonia* species (Matolweni et al. 2000; Hughes and Hollingsworth 2008; Twyford et al. 2013, 2014). Studies of *Primulina eburnea* (Hance) Yin Z.Wang, a gesnerid species restricted to island-like limestone karsts of South China, has also shown that the marked phylogeographic pattern is mainly driven by genetic drift resulted from fragmented karst landscape (Gao et al. 2015). As the process of karstification in SVLK continues and additional geographic isolation in Clade SVLB is expected, further

Table 7 The result of nested clad phylogeographic analysis of *Begonia luzhaiensis* based on *trnC-ycf6* haplotypes

Clade	Significant clades and I-T	Significant parameter	Phylogeographic inferences
1-1	H8 (I) H9 (T) I-T	$D_c > ; D_n >$ $D_c > ; D_n >$	1 → 2 → 3 → 4 → No: Restricted gene flow with isolation by distance
1-2	H6 (I) H7 (T) I-T	$D_n >$ $D_c < ; D_n <$ $D_c > ; D_n >$	1 → 19 → No: Allopatric fragmentation
1-3	H2 (I) H3 (I) H4 (T)	$D_c < ; D_n <$ $D_c < ; D_n >$ $D_c <$	1 → 19 → No: Allopatric fragmentation
2-1	1-1 (T) 1-2 (I) I-T	$D_c < ; D_n <$ $D_c < ; D_n >$ $D_c > ; D_n >$	1 → 2 → 3 → 5 → 15 → No → 21 → No: past gradual range expansion followed by fragmentation or a past larger range followed by extinction in intermediate areas
2-2	1-3 (I) 1-4 (T) 1-5 (T) I-T	$D_n >$ $D_c < ; D_n <$ $D_c >$	1 → 2 → 3 → 5 → 15 → No → 21 → No: past gradual range expansion followed by fragmentation or a past larger range followed by extinction in intermediate areas
Total	2-1 (T) 2-2 (T)	$D_c < ; D_n <$ $D_c < ; D_n >$	1 → 19 → No: Allopatric fragmentation

Phylogeographic inferences are inferred by applying the inference key of Posada et al. (2000) (<http://darwin.uvigo.es/software/geodis.html>)

(T): tip; (I): interior; D_c : clade distance; D_n : nested clade distance; I-T: interior vs. tip clades

population differentiation could eventually lead to speciation (Kozak et al. 2006; Fouquet et al. 2007; Gao et al. 2015), resulting in a steady accumulation of species manifesting as non-adaptive species radiation (Gittenberger 1991; Rundell and Price 2009; Chung et al. 2014).

Anthropogenic effects?

Given that landscape fragmentation triggered by continuous karstification have been crucial in determining the phylogeography and population genetic structure of *Begonia luzhaiensis*, the geohistory of the Sino-Vietnamese limestone karst should have also had a strong influence on the phylogeography of SVLB and other limestone plants. Values of Tajima's *D* and Fu's *F_s* did not significant deviate from 0 in the total sample or for each region, suggesting long-term constant population size. Based on results of mismatch distribution and BSP analyses, populations of *B. luzhaiensis* have experienced a stable population size since the species' expansion from ca. 2.047 Mya (95% confident interval = 0.596–6.072 Mya) until ca. 75,000 years ago (75 Kya) when its effective population size declined, with further population reduction until the present time (Table 5; Fig. 5b).

The prolonged period of stable population size in *B. luzhaiensis* since the late Pliocene to early Pleistocene (Fig. 5b) is congruent with the intensive and accelerated karstification of Sino-Vietnamese limestone karsts ca. 10 to 3 Mya since the late Miocene (Liu 1997; Liu et al. 2001), during which karst landscapes developed, creating abundant caves and cave-like microhabitats for population expansion in *B. luzhaiensis*, most likely from central Guangxi as suggested by the haplotype network (Fig. 4). Not surprisingly, the onset of the last major Northern Hemisphere glacial period ca 20–110 Kya, during which summer temperatures and precipitation reduced drastically due to the weakened summer monsoon (An et al. 2001), jointly retarded the process of karstification in South China (Zhang 1989), likely resulting in a decline of population size in *B. luzhaiensis* 75 Kya (Fig. 5b). However, our data also indicate a continuous and further population declination until present time even though the return of summer monsoons since the end of the last glacial maximum (LGM) should have also resumed the process of karstification, suggesting an alternative cause responsible for the population decline in *B. luzhaiensis* in more recent time.

Paleolithic *Homo sapiens* fossils had been excavated from a large number of sites in China (Wu and Lin 1985; Wu et al. 2012) and Guangxi is one of the most important centers of China's paleoanthropology, with at least 12 sites reported from Guangxi (Yuan and Wei 2015). Thus far, Guangxi's Paleolithic human fossils had been excavated from areas including Guilin (Wang et al. 1981), Laibin (Jia and Wu 1959), Lipu (Wu et al. 1962), and Du'an (Zhao

et al. 1981) where *B. luzhaiensis* of current study sampled (Fig. 2). All these sites of Guangxi, as well as a majority of paleoanthropological sites in China, are all limestone caves, indicating that caves are not only ideal for the preservation of archaeological remains (Wu et al. 2012; Li et al. 2014; Yuan and Wei 2015) but should have also been important for human activities in the ancient times. However, limestone caves are extremely fragile and highly vulnerable to anthropogenic disturbance (Van Beynen and Townsend 2005; Cao et al. 2015). Our own observations during the field work indicate that *B. luzhaiensis* as well as many other limestone cave plants could only be found in caves with minimal human disturbance (e.g., Fig. 1c).

In South China, the transition from Paleolithic to Neolithic ca. 20–10 Kya marks the change from hunter gathering to food-producing societies (Yuan et al. 1995), with earliest evidence of rice growing in Guangxi dated to ca. 2500–2000 B.C. (Zhang and Hung 2010). Because farmable lands for agriculture are limited in Guangxi's extensive limestone karsts, the majority of lands between fenglins and fengcongs suitable for rice growing have been transformed into farmlands (Fig. 1b, d). Prior to development, the lands between fenglins and fengcongs should have been heavily forested and contained microhabitats that could have functioned as dispersal corridors for SVLK plants such as *B. luzhaiensis*, a conjecture based on observations in residual forests escaping from development. Taken together, the increased anthropogenic activity since the early Holocene ca. 10 Kya, firstly utilizing limestone caves and secondarily transforming dispersal corridor into farmlands, could explain the further population declination in *B. luzhaiensis* until the present time (Table 5; Fig. 5b).

Conservation of limestone karst and cave flora

Based on EST-SSR and cpDNA sequences, our study suggests both natural and likely anthropogenic causes for the marked population genetic structure in *Begonia luzhaiensis*. To our knowledge, this is the first phylogeographic study providing evidence for the anthropogenic causes for the population declination of a Sino-Vietnamese limestone plant species. Limestone karsts of Southeast Asia and South China have experienced unprecedented destruction, and biodiversity in the region is severely threatened by uncontrolled increases in tourism (Sang et al. 2011), unsustainable limestone quarries (Clements et al. 2006), regional forest loss and degradation (Sodhi et al. 2004; Laurance 2007), and irrational and intensive land use (Wang et al. 2004). Even for species such as *Begonia luzhaiensis* that is not regarded as rare, our followed up survey found that the DL population had disappeared due to cattle grazing. The population in Laibin, where type specimen of *Primulina cardaminifolia* Yan

Liu and W.B.Xu was also collected, has vanished due to mud flooding (Xu et al. 2013). Considering these threats to the karst ecosystem, conservation of limestone species is urgently needed (Monro et al. 2018).

Consulting Nekola (1999)'s study, Chung et al. (2014) proposed that cave and cave-like microhabitats are "paleorefugia" of limestone cave flora, representing relics of previously much more widespread microhabitats. Our study reveals a marked genetic structure and declined population size in *B. luzhaiensis*, suggesting that current populations of the species represent remnants of a previously much bigger population, consistent with the idea of paleorefugia. Given this, our study provides important information for developing conservation strategies for limestone species. The genetic isolation and the lack of gene flow among populations of *B. luzhaiensis* suggest that we must pay special attention to small isolated areas with genetic uniqueness because they deserve high conservation priority. Therefore, conservation strategies in karst regions should focus not only on species but also the integrity of their genetic diversity. For example, populations of RS, XP, and GT are distinct genetic groups with no genetic connection with other populations based on cpDNA sequences (Fig. 4), so future management plans should focus on conserving the genetic diversity of these distinct populations. By viewing caves and cave-like microhabitats as paleorefugia, our study urges further investigations into phylogeography and population genetic structures of other SVLK flora to design suitable conservation strategies based on genetic data.

Conclusions

The microevolution patterns of *B. luzhaiensis* indicate that genetic drift enhanced by limited gene flow could have been the major factor promoting population diversification in *Begonia*, strongly suggesting the importance of extensive karstification in shaping the species diversity of the Sino-Vietnamese limestone karsts. Our phylogeographic analyses also identify the mountain chain composed of Dananshan, Tianpingshan, Jiaqiaoling, and Dayaoshan could further hinder gene flow among fragmented limestone karst habitats. The appearance of the Neolithic human beings and subsequent development of agriculture are shown to be responsible for further destruction of suitable habitats of limestone species, resulting in drastic population declines in more recent time. As most caves and cave-like microhabitats are not under protection, a majority of the SVLK plants are highly threatened by increased and unregulated human activities in the region. We urge prompt and comprehensive studies of SVLK biodiversity because they are essential for the effective conservation of unique and precious flora.

Additional files

Additional file 1: Table S1. Descriptive statistics of EST-SSR loci. **Table S2.** Genetic differentiation (pairwise F_{ST}) based on EST-SSR between populations of *Begonia luzhaiensis*. All F_{ST} values are significant ($P < 0.001$).

Table S3. Genetic differentiation (pairwise F_{ST}) *trnC-ycf6* spacer sequence between populations of *Begonia luzhaiensis*. **Table S4.** Pairwise geographical distance [$\ln(\text{km}+1)$] between populations of *Begonia luzhaiensis*.

Additional file 2: Figure S1. Plot of F_{ST} versus heterozygosity (H_e) to identify potential loci subject to selection. The dot in the red zone is the candidate under positive selection, in yellow zone for balancing selection, and in grey zone is neutral. Significance was evaluated at the 5% level.

Additional file 3: Figure S2. Demographic history of *Begonia luzhaiensis*. The BSP results from three sampling strategies, including (A) local: east; (B) local: central; (C) local: west; (D) pooled; (E) scattered.

Abbreviations

AIC: akaike information criteria; AMOVA: analysis of molecular variance; BSP: Bayesian Skyline Plot; cpDNA: chloroplast DNA; EST-SSR: expressed sequence tag-simple sequence repeat; HWE: Hardy-Weinberg equilibrium; IBD: isolation by distance; LD: linkage disequilibrium; MCMC: Markov chain Monte Carlo; NCPA: nested clade phylogeographic analysis; PCoA: principal coordinate analysis; SSD: sum of squared deviations; SVLB: Sino-Vietnamese limestone *Begonia*; SVLK: Sino-Vietnamese limestone karsts.

Acknowledgements

We thank Noah Last for improving the writing and Cheng-Tao Lin for making Fig. 2. This work was supported by grants from the Ministry of Science and Technology, Taiwan to Ching-I Peng (NSC 101-2621-B-001-003) and Kuo-Fang Chung (MOST 103-2621-B-002004-MY3) and the National Natural Science Foundation of China (Grant no. 41161011) to Yan Liu.

Authors' contributions

K-FC conceived idea, C-IP, YL, and K-FC designed the project and secured funding, H-YH, W-BX, H-AY, and K-FC collected plants, Y-HT and H-YH collected genetic data and performed data analyses, Y-HT and K-FC wrote the manuscript, and all authors read and confirmed the manuscript. All authors read and approved the final manuscript.

Funding

Supported by Ministry of Science and Technology, Taiwan [NSC 101-2621-B-001-003 (to C-IP and K-FC), MOST 103-2621-B-002004-MY3 (to K-FC)] and the National Natural Science Foundation of China [Grant no. 41161011 (to YL)].

Availability of data and materials

All DNA sequences generated in this study have registered in GenBank.

Ethics approval and consent to participate

Guangxi Key Laboratory of Plant Conservation and Restoration Ecology in Karst Terrain, Guangxi Institute of Botany has permits to conduct fieldtrips in Guangxi (to W-BX and YL).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Research Museum and Herbarium (HAST), Biodiversity Research Center, Academia Sinica, Taipei, Taiwan. ² School of Forestry and Resource Conservation, National Taiwan University, Taipei, Taiwan. ³ Guangxi Key Laboratory of Plant Conservation and Restoration Ecology in Karst Terrain, Guangxi Institute of Botany, Guangxi Zhuang Autonomous Region and Chinese Academy of Sciences, Guilin, Guangxi, China.

Received: 5 March 2019 Accepted: 8 August 2019

Published online: 06 September 2019

References

- An Z-S, Kutzbach JE, Prell WL, Porter SC (2001) Evolution of Asian monsoons and phased uplift of the Himalayan-Tibetan plateau since Late Miocene times. *Nature* 411(6833):62–66. <https://doi.org/10.1038/35075035>
- Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a F_{st} -outlier method. *BMC Bioinformatics* 9:e323. <https://doi.org/10.1186/1471-2105-9-323>
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc B Biol Sci* 263(1377):1619–1626. <https://doi.org/10.1098/rspb.1996.0237>
- Bekker EI, Karabanov DP, Galimov YR, Haag CR, Neretina TV, Kotov AA (2018) Phylogeography of *Daphnia magna* Straus (Crustacea: Cladocera) in Northern Eurasia: Evidence for a deep longitudinal split between mitochondrial lineages. *PLoS ONE* 13(3):e0194045. <https://doi.org/10.1371/journal.pone.0194045>
- Buhay JE, Crandall KA, Posada D (2009) Nested clade phylogeographic analysis for conservation genetics. In: Rizzoli A, Vernesi C, Bertorello G, Hauffe HC, Bruford MW (eds) *Population genetics for animal conservation*. Cambridge University Press, Cambridge, pp 80–103. <https://doi.org/10.1017/CBO9780511626920.005>
- Cao J-H, Yuan D-X, Tong L-Q, Mallik A, Yang H, Huang F (2015) An overview of karst ecosystem in Southwest China: current state and future management. *J Resour Ecol* 6(4):247–256. <https://doi.org/10.5814/j.issn.1674-764x.2015.04.008>
- Chan YM, Tnah LH, Lee SL, Bhasu S, Lee CT, Chua LSL (2018) Limited dispersal and geographic barriers cause population differentiation and structuring in *Begonia maxwelliana* at both large and small scales. *Plant Ecol Divers* 11(1):69–83. <https://doi.org/10.1080/17550874.2018.1471625>
- Chen W-H, Radbouchoom S, Nguyen HQ, Nguyen HT, Nguyen KS, Shui Y-M (2018) Seven new species of *Begonia* (Begoniaceae) in Northern Vietnam and Southern China. *PhytoKeys* 94:65–85. <https://doi.org/10.3897/phytokeys.94.23248>
- Chikhi L, Sousa VC, Luisi P, Goossens B, Beaumont MA (2010) The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics* 186(3):983–995. <https://doi.org/10.1534/genetics.110.118661>
- Chung K-F, Leong W-C, Rubite RR, Repin R, Kiew R, Liu Y, Peng C-I (2014) Phylogenetic analyses of *Begonia* sect. *Coelocentrum* and allied limestone species of China shed light on the evolution of Sino-Vietnamese karst flora. *Bot Stud* 55:e1. <https://doi.org/10.1186/1999-3110-55-1>
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9(10):1657–1659. <https://doi.org/10.1046/j.1365-294x.2000.01020.x>
- Clements R, Sodhi NS, Schilthuizen M, Ng PKL (2006) Limestone karsts of Southeast Asia: imperiled arks of biodiversity. *BioScience* 56(9):733–742. [https://doi.org/10.1641/0006-3568\(2006\)56%5b707:WTFa%5d2.0.CO;2](https://doi.org/10.1641/0006-3568(2006)56%5b707:WTFa%5d2.0.CO;2)
- Comes HP, Tribsch A, Bittkau C (2008) Plant speciation in continental island floras as exemplified by *Nigella* in the Aegean Archipelago. *Philos Trans R Soc B* 363(1506):3083–3096. <https://doi.org/10.1098/rstb.2008.0063>
- Cui W-H, Guan K-Y (2013) Diversity of leaf variegation in Chinese begonias. *PI Divers Resour* 35(2):119–127. <https://doi.org/10.7677/ynzwyj201312118>
- Day M, Ulrich P (2000) An assessment of protected karst landscapes in Southeast Asia. *Cave Karst Sci* 27(2):61–70
- de Lange A, Bouman F (1999a) Seed micromorphology of the genus *Begonia* in Africa: taxonomic and ecological implications. In de Wilde JJFE (ed) *Studies in Begoniaceae III* Wageningen Agricultural University 91–94. pp. 1–82
- de Lange A, Bouman F (1999b) Seed micromorphology of neotropical begonias. *Smithsonian Contribution to Botany* 90:1–49
- Demesure B, Sodhi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. *Mol Ecol* 4(1):129–131. <https://doi.org/10.1111/j.1365-294X.1995.tb00201.x>
- Dewitte A, Twyford AD, Thomas DC, Kidner CA, Huylensbroeck JV (2011) The origin of diversity in *Begonia*: genome dynamism, population processes and phylogenetic patterns. In: Grillo O, Venora G (eds) *The dynamical processes of biodiversity—case studies of evolution and spatial distribution*. InTech Open Access Publisher, England, pp 27–52
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29(8):1969–1973. <https://doi.org/10.1093/molbev/mss075>
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4(2):359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Ersts PJ (2011) Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. http://biodiversityinformatics.amnh.org/open_source/gdmg. Accessed 31 May 2018
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14(8):2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10(3):564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Ford D, Williams P (2007) *Karst Hydrogeology and Geomorphology*. Wiley, Chichester
- Fouquet A, Gilles A, Vences M, Marty C, Blanc M, Gemmill NJ (2007) Underestimation of species richness in neotropical frogs revealed by mtDNA analyses. *PLoS ONE* 2(10):e1109. <https://doi.org/10.1371/journal.pone.0001109>
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147(2):915–925
- Gao Y, Ai B, Kong H, Kang M, Huang H (2015) Geographical pattern of isolation and diversification in karst habitat islands: a case study in the *Primulina eburnea* complex. *J Biogeogr* 42(11):2131–2144. <https://doi.org/10.1111/jbi.12576>
- Gittenberger E (1991) What about non-adaptive radiation? *Biol J Linn Soc* 43:263–272
- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Lausanne University, Lausanne
- Graur D, Li W-H (2000) *Fundamentals of molecular evolution*, 2nd edn. Sinauer Associates, Sunderland
- Gu C-Z, Peng C-I, Turland NJ (2007) Begoniaceae. In: Wu Z-Y, Raven PH, Hong D-Y (eds) *Flora of China*, vol 13. Science Press and Missouri Botanical Garden, Beijing and St. Louis, pp 153–207
- Heller R, Chikhi L, Siegmund HR (2013) The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PLoS ONE* 8(5):e62992. <https://doi.org/10.1371/journal.pone.0062992>
- Hou M-F, López-Pujol J, Qin H-N, Wang L-S, Liu Y (2010) Distribution pattern and conservation priorities for vascular plants in Southern China: Guangxi Province as a case study. *Bot Stud* 51(3):377–386
- Huang Y-S, Xu W-B, Ye X-X, Liu Y (2010) *Chirita luzhaiensis*, a new species of Gesneriaceae from limestone areas in Guangxi, China. *J Trop Subtrop Bot* 18(2):137–139
- Hughes AC (2017) Understanding the drivers of Southeast Asian biodiversity loss. *Ecosphere* 8(1):e01624. <https://doi.org/10.1002/ecs2.1624>
- Hughes M, Hollingsworth PM (2008) Population genetic divergence corresponds with species-level biodiversity patterns in the large genus *Begonia*. *Mol Ecol* 17(11):2643–2651. <https://doi.org/10.1111/j.1365-294X.2008.03788.x>
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23(14):1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Jia L, Wu R (1959) Fossil human skull base of Late Paleolithic Stage from Chiling-shan, Leipin District, Kwangsi, China. *Palaeovertebrata et Paleoanthropologia* 1(1):16–18
- Knowles LL (2008) Why does a method that fails continue to be used? *Evolution* 62(11):2713–2717. <https://doi.org/10.1111/j.1558-5646.2008.00481.x>
- Kopperud C, Einset JW (1995) DNA isolation from *Begonia* leaves. *Plant Mol Biol Rep* 13(2):129–130. <https://doi.org/10.1007/BF02668783>

- Kozak KH, Weisrock DW, Larson A (2006) Rapid lineage accumulation in a non-adaptive radiation: phylogenetic analysis of diversification rates in eastern North American woodland salamanders (Plethodontidae: *Plethodon*). *Proc R Soc B Biol Sci* 273(1586):539–546. <https://doi.org/10.1098/rspb.2005.3326>
- Ku T-C (1999) Begoniaceae. In Ku T-C (ed.) *Flora Reipublicae Popularis Sinicae*. Vol. 52(1). Science Press, Beijing. pp. 126–269
- Laurance WF (2007) Forest destruction in tropical Asia. *Curr Sci India* 93(11):1544–1550
- Lee C, Wen J (2004) Phylogeny of *Panax* using chloroplast *trnC-trnD* intergenic region and the utility of *trnC-trnD* in interspecific studies of plants. *Mol Phylogenet Evol* 31(3):894–903. <https://doi.org/10.1016/j.ympev.2003.10.009>
- Li D, Wang W, Tian F, Liao W, Bae CJ (2014) The oldest bark cloth beater in southern China (Dingmo, Buning basin, Guangxi). *Quatern Int* 354:184–189. <https://doi.org/10.1016/j.quaint.2014.06.062>
- Li C, Yang LH, Tian DK, Chen Y, Wu RJ, Fu NF (2016) *Begonia leipingensis* (Begoniaceae), a new compound-leaved species with unique petiolule pattern from Guangxi of China. *Phytotaxa* 244(1):45–56. <https://doi.org/10.11646/phytotaxa.244.1.3>
- Li C, Chen J-L, Li Y-Y, Tian D-K (2018) Genetic diversity of *Begonia versicolor* (Begoniaceae), a narrow endemic species in southeast Yunnan of China. *Taiwania* 63(1):49–53. <https://doi.org/10.6165/tai.2018.63.49>
- Liang KM, Lin ZF, Liu N, Zhang QM, Wang J, Wang ZF, Guan LL (2010) Characteristics of sun- and shade-adapted populations of an endangered plant *Primulina tabacum* Hance. *Photosynthetica* 48(4):494–506. <https://doi.org/10.1007/s11099-010-0066-8>
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25(11):1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>
- Liu J-R (1997) The development history of the Guangxi tropical karst geomorphology and its sequences. *Carsologica Sin* 16(4):332–345
- Liu J-R, Huang G-B, Huang X-L, Liang Y-C (2001) Discussion on the evolution of different types of regional tropical karst landform in Guangxi. *Carsologica Sin* 20(4):247–252
- Maddison WP, Maddison DR (2018) Mesquite: a modular system for evolutionary analysis. Version 3:40
- Matolweni LO, Balkwill K, McLellan T (2000) Genetic diversity and gene flow in the morphologically variable, rare endemics *Begonia dregei* and *Begonia homonyma* (Begoniaceae). *Am J Bot* 87(3):431–439. <https://doi.org/10.2307/2656639>
- Mazet O, Rodriguez W, Chikhi L (2015) Demographic inference using genetic data from a single individual: separating population size variation from population structure. *Theor Popul Biol* 104:46–58. <https://doi.org/10.1016/j.tpb.2015.06.003>
- Mazet O, Rodriguez W, Grusea S, Boitard S, Chikhi L (2016) On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? *Heredity* 116(4):362–371. <https://doi.org/10.1038/hdy.2015.104>
- Monro AK, Bystrickova N, Fu L, Wen F, Wei Y (2018) Discovery of a diverse cave flora of China. *PLoS ONE* 13(2):e0190801. <https://doi.org/10.1371/journal.pone.0190801>
- Moonlight PW, Ardi WH, Padilla LA, Chung K-F, Fuller D, Girmansyah D, Hollands R, Jara-Muñoz A, Kiew R, Leong W-C, Liu Y, Mahardika A, Marsinghe LDK, O'Connor M, Peng C-I, Pérez AJ, Phutthai T, Pullan M, Rajbhandary S, Reynel C, Rubite RR, Sang J, Scherberich D, Shui Y-M, Tebbitt MC, Wilson HP, Zaini NH, Hughes M (2018) Dividing and conquering the fastest growing genus: towards a natural sectional classification of the mega-diverse genus *Begonia* (Begoniaceae). *Taxon* 67(2):267–323. <https://doi.org/10.12705/672.3>
- Nakamura K, Kokubugata G, Rubite RR, Huang C-J, Kono Y, Yang H-A, Lopez-Feliciano A, Labuguen ML, Yokota M, Peng C-I (2014) *In situ* glacial survival at the northern limit of tropical insular Asia by a lowland herb *Begonia fenicis* (Begoniaceae). *Bot J Linn Soc* 174(3):305–325. <https://doi.org/10.1111/boj.12138>
- Nei M, Chakravarti A, Tateno Y (1977) Mean and variance of F_{ST} in a finite number of incompletely isolated populations. *Theor Popul Biol* 11(3):291–306. [https://doi.org/10.1016/0040-5809\(77\)90013-2](https://doi.org/10.1016/0040-5809(77)90013-2)
- Nekola JC (1999) Paleoreefugia and neoreefugia: the influence of colonization history on community pattern and process. *Ecology* 80(8):2459–2473. [https://doi.org/10.1890/0012-9658\(1999\)080%5b2459:PANTIO%5d2.0.CO;2](https://doi.org/10.1890/0012-9658(1999)080%5b2459:PANTIO%5d2.0.CO;2)
- Oksanen J, Guillaume BF, Kindt R, Legendre P, O'Hara RG, Simpson GL, Solymos P, Henry M, Stevens H, Wagner H (2010) *Vegan: community ecology package*. R package version 1.17-0
- Palmer AN (1991) Origin and morphology of limestone caves. *Geol Soc Am Bull* 103(1):1–21. [https://doi.org/10.1130/0016-7606\(1991\)103%3c0001:Oamolc%3e2.3.Co;2](https://doi.org/10.1130/0016-7606(1991)103%3c0001:Oamolc%3e2.3.Co;2)
- Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19):2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Peng C-I, Kuo S-M, Yang H-A, Leong W-C, Liu Y, Nguyen TH, Kono Y, Chung K-F (2014) Two new species of *Begonia* sect. *Coelocentrum*, *B. guixiensis* and *B. longa*, from Sino-Vietnamese limestone karsts. *Bot Stud* 55:e52. <https://doi.org/10.1186/s40529-014-0052-8>
- Peng C-I, Lin C-W, Yang H-A, Kono Y, Nguyen HQ (2015a) Six new species of *Begonia* (Begoniaceae) from limestone areas in Northern Vietnam. *Bot Stud* 56:e9. <https://doi.org/10.1186/s40529-015-0089-3>
- Peng C-I, Yang H-A, Kono Y, Jung M-J, Nguyen TH (2015b) Four new species of *Begonia* (Begoniaceae) from Vietnam: *B. abbreviata*, *B. calciphila*, *B. sphenantheroides* and *B. tamdaoensis*. *Phytotaxa* 222(2):83–99. <https://doi.org/10.11646/phytotaxa.222.2.1>
- Posada D (2008) jModelTest: phylogenetic model averaging. *Mol Biol Evol* 25(7):1253–1256. <https://doi.org/10.1093/molbev/msn083>
- Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol Ecol* 9(4):487–488. <https://doi.org/10.1046/j.1365-294x.2000.00887.x>
- Poulson TL, White WB (1969) The cave environment. *Science* 165(3897):971–981
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959
- Qin H-N, Liu Y (eds) (2010) *A Checklist of Vascular Plants of Guangxi*. Science Press, Beijing
- Qin H, Yang Y, Dong S, He Q, Jia Y, Zhao L, Yu S, Liu H, Liu B, Yan Y, Xiang J, Xia N, Peng H, Li Z, Zhang Z, He X, Yin L, Lin Y, Liu Q, Hou Y, Liu Y, Liu Q, Cao W, Li J, Chen S, Jin X, Gao T, Chen W, Ma H, Geng Y, Jin X, Chang C, Jiang H, Cai L, Zang C, Wu J, Ye J, Lai Y, Liu B, Lin Q, Xue N (2017a) Threatened Species List of China's Higher Plants. *Biodiversity Science* 25(7):696–744. <https://doi.org/10.17520/biods.2017144>
- Qin Y-H, Liang Y-Y, Xu W-B, Lin C-W, Peng C-I (2017b) *Begonia ufoides* (sect. *Coelocentrum*, Begoniaceae), a new species from limestone areas in central Guangxi, China. *China. Phytotaxa* 316(3):279–284
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67(5):901–904. <https://doi.org/10.1093/sysbio/syy032>
- Ren H, Zhang QM, Wang ZF, Guo QF, Wang J, Liu N, Liang KM (2010) Conservation and possible reintroduction of an endangered plant based on an analysis of community ecology: a case study of *Primulina tabacum* Hance in China. *Plant Spec Biol* 25(1):43–50. <https://doi.org/10.1111/j.1442-1984.2009.00261.x>
- Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8(1):103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Rundell RJ, Price TD (2009) Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends Ecol Evol* 24(7):394–399. <https://doi.org/10.1016/j.tree.2009.02.007>
- Sanderson MJ, Wojciechowski MF (1996) Diversification rates in a temperate legume clade: are there "so many species" of *Astragalus* (Fabaceae)? *Am J Bot* 83(11):1488–1502. <https://doi.org/10.1002/j.1537-2197.1996.tb13942.x>
- Sang W, Ma K, Axmacher JC (2011) Securing a future for China's wild plant resources. *Bioscience* 61(9):720–725. <https://doi.org/10.1525/bio.2011.61.9.11>
- Santos AD, Carrizo TF, Canello EM, Castro ACMCE (2017) Phylogeography of *Nasutitermes corniger* (Isoptera: Termitidae) in the Neotropical region. *BMC Evol Biol* 17:e230. <https://doi.org/10.1186/s12862-017-1079-8>
- Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152(3):1079–1089

- Sgarlata GM, Salmona J, Aleixo-Pais I, Rakotonanahary A, Sousa AP, Kun-Rodrigues C, Ralantoharijaona T, Jan F, Zaranaina R, Rasolondraibe E, Zaonarivelo JR, Andriaholinirina NV, Chikhi L (2018) Genetic differentiation and demographic history of the Northern Rufous Mouse Lemur (*Microcebus tavaratra*) across a fragmented landscape in northern Madagascar. *Int J Primatol* 39(1):65–89. <https://doi.org/10.1534/genetics.110.118661>
- Sheue C-R, Pao S-H, Chien L-F, Chesson P, Peng C-I (2012) Natural foliar variegation without costs? The case of *Begonia*. *Ann Bot Lond* 109(6):1065–1074. <https://doi.org/10.1093/aob/mcs025>
- Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations. *Evolution* 47(1):264–279. <https://doi.org/10.1111/j.1558-5646.1993.tb01215.x>
- Sodhi NS, Koh LP, Brook BW, Ng PKL (2004) Southeast Asian biodiversity: an impending disaster. *Trends Ecol Evol* 19(12):654–660. <https://doi.org/10.1016/j.tree.2004.09.006>
- Sweeting MM (1978) The karst of Kweilin, southern China. *Geol J* 144(2):199–204
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123(3):585–595
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731–2739. <https://doi.org/10.1093/molbev/msr121>
- Templeton AR (2009) Statistical hypothesis testing in intraspecific phylogeography: nested clade phylogeographical analysis vs. approximate Bayesian computation. *Mol Ecol* 18(2):319–331. <https://doi.org/10.1111/j.1365-294X.2008.04026.x>
- Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* 134(2):659–669
- Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140(2):767–782
- Tseng Y-H, Huang H-Y, Xu W-B, Yang H-A, Liu Y, Peng C-I, Chung K-F (2017) Development and characterization of EST-SSR markers for *Begonia luzhaiensis* (Begoniaceae). *Appl Plant Sci* 5(5):e1700024. <https://doi.org/10.3732/apps.1700024>
- Twyford AD, Kidner CA, Harrison N, Ennos RA (2013) Population history and seed dispersal in widespread Central American *Begonia* species (Begoniaceae) inferred from plastome-derived microsatellite markers. *Bot J Linn Soc* 171(1):260–276. <https://doi.org/10.1111/j.1095-8339.2012.01265.x>
- Twyford AD, Kidner CA, Ennos RA (2014) Genetic differentiation and species cohesion in two widespread Central American *Begonia* species. *Heredity* 112(4):382–390. <https://doi.org/10.1038/hdy.2013.116>
- van Beynen P, Townsend K (2005) A disturbance index for karst environments. *Environ Manage* 36(1):101–116. <https://doi.org/10.1007/s00267-004-0265-9>
- Waltham T (2008) Fengcong, fenglin, cone karst and tower karst. *Cave Karst Sci* 35(3):77–88
- Wang L-H, Peng S-L, Chen Y-Z (1981) On the human fossils and stone artifacts found in Baojiyan Cave, Guilin. *Acta Anthropologica Sinica* 1(1):30–35
- Wang SJ, Liu QM, Zhang DF (2004) Karst rocky desertification in southwestern China: geomorphology, landuse, impact and rehabilitation. *Land Degrad Dev* 15(2):115–121. <https://doi.org/10.1002/Ldr.592>
- Wiens JJ (2004) Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution* 58(1):193–197. <https://doi.org/10.1111/j.0014-3820.2004.tb01586.x>
- Williams P (2008) World Heritage Caves and Karst. IUCN (International Union for Conservation of Nature), Gland
- Wu R-K, Lin S-L (1985) Chinese paleoanthropology: retrospect and prospect. In: Wu R-K, Olsen JW (eds) *Palaeoanthropology and Palaeolithic Archaeology in the People's Republic of China*. Academic Press, New York, pp 1–28
- Wu X-Z, Zhao Z-K, Yuan Z-X, Shen J-Y (1962) Report on paleoanthropological expedition of the northeastern part of Kwangsi. *Vertebrata Palasiatica* 6(4):408–413
- Wu X, Zhang C, Goldberg P, Cohen D, Pan Y, Arpin T, Bar-Yosef O (2012) Early pottery at 20,000 years ago in Xianrendong Cave, China. *Science* 336(6089):1696–1700. <https://doi.org/10.1126/science.1218643>
- Xu W-B, Pan B, Liu Y, Peng C-I, Chung K-F (2012) Two new species, *Primulina multifida* and *P. pseudomollifolia* (Gesneriaceae), from karst caves in Guangxi, China. *Bot Stud* 53(1):165–175
- Xu W-B, Liu Y, Kono Y, Chang H, Peng C-I, Chung K-F (2013) *Primulina cardaminifolia* (Gesneriaceae), a rare new species from limestone areas in Guangxi, China. *Bot Stud* 24:e19. <https://doi.org/10.1186/1999-3110-54-19>
- Xu W-B, Guo J, Pan B, Han M-Q, Liu Y, Chung K-F (2017) Three new species of *Paraboea* (Gesneriaceae) from limestone karsts of China based on morphological and molecular evidence. *Bot Stud* 58:e56. <https://doi.org/10.1186/s40529-017-0207-5>
- Yuan JJ, Wei X (2015) New progress in paleoanthropology of Guangxi. *J Guangxi Nor Univ Philos Soc Sci* 51(3):21–27. <https://doi.org/10.16088/j.jissn.1001-6597.2015.03.004>
- Yuan S-X, Zhou G-X, Guo Z-Y, Zhang Z-M, Gao S-J, Li K, Wang J-J, Liu K-X, Li B, Lu X-Y (1995) ¹⁴C AMS dating the transition from the Paleolithic to the Neolithic in South China. *Radiocarbon* 37(2):245–249. <https://doi.org/10.1017/S0033822200030708>
- Zhang L-Y (1989) Some special geomorphic processes of the monsoon area in East China. *CATENA* 16:121–134. [https://doi.org/10.1016/0341-8162\(89\)90036-2](https://doi.org/10.1016/0341-8162(89)90036-2)
- Zhang C, Hung H-C (2010) The emergence of agriculture in southern China. *Antiquity* 84(323):11–25. <https://doi.org/10.1017/S0003598X00099737>
- Zhang Y, Hayashi T, Hosokawa M, Yazawa S, Li Y-H (2009) Metallic lustre and the optical mechanism generated from the leaf surface of *Begonia rex* Putz. *Sci Hortic-Amsterdam* 121(2):213–217. <https://doi.org/10.1016/j.scienta.2009.01.030>
- Zhao Z-R, Lix X-S, Wang L-H (1981) Human fossils and associated fauna of Jiulengshan Hill, Guangxi. *Vertebrata Palasiatica* 19(1):45–54

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)