

Research Article

Does Large Genome Size Limit Speciation in Endemic Island Floras?

Maxim V. Kapralov and Dmitry A. Filatov

Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

Correspondence should be addressed to Maxim V. Kapralov, maxim.kapralov@plants.ox.ac.uk

Received 27 May 2011; Accepted 28 August 2011

Academic Editor: Andrea Polle

Copyright © 2011 M. V. Kapralov and D. A. Filatov. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Genome sizes in plants vary by several orders of magnitude, and this diversity may have evolutionary consequences. Large genomes contain mainly noncoding DNA that may impose high energy and metabolic costs for their bearers. Here we test the large genome constraint hypothesis, which assumes that plant lineages with large genomes are diversifying more slowly Knight et al. (2005), using endemic floras of the oceanic archipelagos of the Canaries, Hawaii, and Marquesas Islands. In line with this hypothesis, the number of endemic species per genus is negatively correlated with genus-average genome size for island radiations on Hawaiian and Marquesas archipelagos. However, we do not find this correlation on the Canaries, which are close to the continent and therefore have higher immigration rate and lower endemism compared to Hawaii. Further work on a larger number of floras is required to test the generality of the large genome constraint hypothesis.

1. Introduction

The DNA content of one nonreplicated holoploid genome with the chromosome number n , referred as $1C$ -value [1], varies nearly 2400-fold across angiosperms [2], from $1C = 0.0648$ pg in *Genlisea margaretae* (Lentibulariaceae) [3] to $1C = 152.23$ pg in *Paris japonica* (Melanthiaceae) [4]. However, gene numbers per genome in angiosperms do not vary so greatly [5] despite gargantuan variation in DNA content. For example, genome size of *Zea mays* is nearly 8-fold bigger than one of another grass species, *Brachypodium distachyon*, but difference in gene numbers between these species is only 22%, the discrepancy which is mainly explained by different retrotransposon content in two genomes [5]. Number of eukaryotic genes is relatively stable and makes up a small fraction of total DNA while much of the variation in genome size is due to noncoding DNA [6], which may be energetically and metabolically costly for their bearers [7]. Vinogradov [8] and Knight et al. [9] found negative correlation between the genus-level diversity and the genus-average genome size in plants, suggesting a genome size constraint on capacity for diversification. Further Knight et al. [9] proposed the

large genome constraint hypothesis (LGCH) that suggests that species with large genomes are less likely to generate progenitor species. The LGCH is in agreement with the general observation that most angiosperm species have small genomes, with a mode, median, and mean genome size ($1C$) of just 0.6, 2.6, and 6.2 pg, respectively [10]. The LGCH echoes point of view that larger genomes are maladaptive, as they may constrain growth [11], and evolved in populations with smaller effective population size and hence low efficacy of natural selection [6]. However, no relationship between effective population size and genome size was found in seed plants [12]. Some support to the LGCH comes from Suda et al. who hypothesized that “rapid insular burst of speciation is more likely to happen in angiosperms with minute nuclear DNA amounts” [13] (page 234) after finding that many island lineages of Macaronesian angiosperms which underwent adaptive radiations have very small genome sizes. Analyses by Vinogradov [8] and Knight et al. [9] do tentatively support the LGCH, but the negative correlations they found between the genus-level diversity and the genus-average genome size were quite weak, -0.11 and -0.065 , respectively, and the methods they used might be a

subject to a phylogenetic bias—more closely related species are expected to have more similar genome sizes, which was not taken into account in the previous analyses. There is a call for phylogenetic comparative analyses of genome size [14] with a complete genus-level phylogeny of plants [9]. While the complete genus-level phylogeny of flowering plants is yet to be achieved despite the current significant progress in the field [15], some of the regional floras have been studied well enough for the task.

Oceanic archipelagos have been regarded as nature's laboratories since Darwin's and Wallace's seminal works [16, 17] and may offer a particularly good opportunity to test the LGCH. Oceanic archipelagos are groups of islands with an exclusive volcanic origin that have never been connected to continents [16]. Biota of oceanic islands is composed of species that arrived via long-distance dispersal or evolved through *in situ* speciation often via "bursts" of speciation that form multiple closely related species adapted to a broad spectrum of ecological niches [18]. Usually a significant proportion of species on oceanic archipelagos are endemic [19]. If we assume that a large genome is an evolutionary handicap, then island endemic lineages with bigger genome sizes should generate fewer progenitor species compared to their relatives with smaller genomes. Here we use a phylogenetic framework to test this prediction of the LGCH in the endemic floras of the oceanic archipelagos of the Canaries, Hawaii, and Marquesas Islands. All three archipelagos possess highly diverse and intensively studied floras which are well suited to test our hypothesis [18, 20, 21].

The Canary archipelago is formed by 11 volcanically active islands and islets that are 17–24 My old and about 7447 km² in area [22]. The Canaries, located just off the northwest coast of mainland Africa, 100 km west of the border between Morocco and the Western Sahara. The archipelago possesses high ecosystem diversity, including dry semidesert vegetation of the coastal lowlands, woodlands, the laurel forest zone, pine forest, and the summit scrub. There are about 680 endemic vascular plant taxa in the Canaries accounting for over 50% of native flora [20]. *C*-values of 40% of plant species endemic to the Canaries were estimated by Suda et al. [13, 23] that makes the Canary flora the best covered regional flora from the genome size perspective.

The Hawaiian archipelago includes 8 major islands with a total land area of about 16636 km² located in the middle of the Pacific Ocean. The origin of the archipelago dates back to about 70 My ago, although most of extant islands are younger than 5 My [24]. The high, up to 3000 m, elevation of the islands creates steep climatic gradients and diverse ecosystems ranging from the dry exposed coastal cliffs, through dry, mesic, and wet forests, to alpine summit scrubs. As an isolated archipelago, Hawaii is relatively poor in species with 1009 native angiosperm species but rich in endemic taxa which constitute about 90% of the native flora [25].

The Marquesas archipelago, located in the Eastern Pacific Ocean, is comprised of 9 main islands with a total area of 1049 km². These tropical islands are subjected to frequent drought conditions due to the prevailing easterly winds formed from the dry air masses above the Humboldt Current. The Marquesas archipelago is characterized by relatively

homogeneous conditions and hence by an impoverished native flora (ca. 360 species), with a high proportion of endemics (42%; [21]).

Here we are using endemic floras of these archipelagos to test the large genome constraint hypothesis [9]. Unlike the previous studies [8, 9], we employ a phylogenetic framework to take the relatedness of species into account. The negative correlation between the number of endemic species per genus and genome size remains significant under this framework for the Pacific archipelagos, providing additional support to the LGCH.

2. Methods

2.1. Data Collection. Data on the endemic angiosperm flora of the Canary Islands were obtained from the checklist [26]. The lists of endemic angiosperm species for the floras of the Hawaiian and Marquesas Islands were obtained from the websites developed by the Smithsonian Institution [27, 28]. Only the species-level taxa were included in the analysis. Further we obtained the average genome size for genera with endemic species from the Plant DNA *C*-values database at the Royal Botanical Gardens at Kew [29] which contains *C*-values for 1.8% of all angiosperm species and 58% of angiosperm families [2]. The incompleteness of the *C*-value database and cases when large genera are represented just by one or few species contributed random gaps and noise into our analyses making it more conservative. For the Canaries only data from Suda et al. [13, 23] were used. Genome size values for the Hawaiian endemic genus *Schiedea* were obtained from [30]. This resulted in a dataset with information on numbers of endemic species per genus and genus-average genome size for 126, 67, and 17 genera from the Canary Islands, the Hawaiian Islands, and the Marquesas Islands, respectively (see Table S1 of Supplementary Material available online at doi:1.1155/2011/458684).

2.2. Data Analyses. We tested correlations between numbers of endemic species per genus and genus-average genome size using phylogenetically independent contrasts (PIC; [31]). The PIC approach is more conservative than conventional statistics; the difference in trait values is calculated at each node of the phylogeny, resulting in $n - 1$ contrasts where n is the number of species in a fully resolved tree. We conducted Felsenstein's independent contrasts in Mesquite (version 2.72) [32] using the PDAP:PDTREE module [33]. For the analyses of the endemic floras of Hawaii and the Marquesas Islands as well as for combined data set of Hawaii and Marquesas Islands, which share many genera with endemic species and belong to the same biogeographic area, we used phylogenetic trees built using *rbcl* sequences obtained from GenBank [34]. Phylogenetic trees were reconstructed with Bayesian inference using MrBayes 3.1.2 [35, 36]. Alignments were partitioned by codons, and the general time-reversible nucleotide substitution model with gamma shape parameter was used. All model parameters were optimized independently for each codon position. Two independent analyses, each with four parallel chains, were run for 4000 000 generations, sampling trees every 100

TABLE 1: Genus-average genome size and endemic diversity statistics of studied archipelagos.

Archipelago	N genera included	N endemic species	N species per genus	Genus-average C-value (1C, pg)		Felsenstein's contrasts correlation		
				mean	SD	d.f.	R	P value
The Canaries	126	480	3.81	2.05	2.81	109	-0.032	0.360
Hawaii	67	334	4.99	2.36	2.22	64	-0.251	0.020
Marquesas Islands	17	52	3.06	1.48	0.83	14	-0.084	0.375
Hawaii and Marquesas Is.	73	386	5.29	2.35	2.14	70	-0.267	0.011

generations after a burn-in period of 1000 000 generations. However, for about a half of the genera from the data set representing the Canary Islands *rbcL* sequences were not available. Thus, to build a phylogeny of Canary endemics we used the program Phylomatic 2 [37] which utilizes current knowledge of phylogenetic relations between angiosperm taxa in combination with published phylogenies [38–41] and our own tree for available genera (Supplementary Material). All resulted phylogenies (Figure 1; Figures S1–S3) matched generally accepted phylogenetic relationships.

3. Results and Discussion

The sampled 126, 67, and 17 genera from the Canary Islands, the Hawaiian Islands, and the Marquesas Islands contained 480, 334, and 52 endemic species, respectively. The joint Hawaii-Marquesas dataset contained 73 genera with 386 endemic species and had the highest number of endemic species per genus (5.29). The Canaries had fewer endemic species per genus (3.81). Average 1C values were 2.05 and 2.35 pg for the Canaries and the joint Hawaii-Marquesas dataset, respectively. Differences between Hawaii-Marquesas dataset and the Canary Islands were marginally insignificant for numbers of endemic species per genus (P value = 0.07; t -test) and not significant for genus-average C -values (P value = 0.24; t -test).

Mean C -values for all archipelagos is nearly threefold lower than the mean calculated for all available angiosperms [10] in accordance with Suda et al's conclusions for Macaronesian angiosperms [13]. Thus, relatively small genome size of island endemics compared to the mainland biota is confirmed for three oceanic archipelagos and seems to be a general rule. Smaller genomes of island endemics could be explained by either genome miniaturization during or after island speciation events or by the predominance of colonizers with small genomes. Island populations often have small effective population size due to limited resources and bottlenecks during island-hopping speciation. Hence, given increased activity of transposable elements in small populations [42], genome miniaturization might not be very common on islands but the opposite trend may prevail. Indeed, nearly threefold genome increase in younger species without a change in ploidy level was reported for the Hawaiian endemic genus *Schiedea* (Caryophyllaceae) presumably due to accumulation of transposons [30]. Despite this increase, *Schiedea* is also a good illustration of smaller

genome size of island sister taxa compared to the mainland counterpart given that this Hawaiian endemic genus has over fourfold smaller genome compared to its sister mainland genus, *Honckenya* [30]. Thus the predominance of colonizers with small genomes and/or higher naturalization potential of species with small genomes is a more likely explanation for smaller genomes of island endemics. This is in agreement with recent findings that invasive plant species have smaller genomes than their noninvasive relatives [43–45].

PIC analyses showed a negative correlation between the number of endemic species per genus and genus-average genome size for all archipelagos analyzed separately as well as for the joint Hawaii-Marquesas dataset (Table 1). However, only in the Hawaiian and joint Hawaii-Marquesas datasets correlation was significant (Table 1). The Marquesas Islands alone do not show a significant correlation, presumably because of a small sample size; however when they were combined with biogeographically similar Hawaii, it made the negative correlation stronger (Table 1).

While smaller genomes of island endemics hold for all studied archipelagos, there is a striking difference between Hawaii and the Canaries in negative correlation between numbers of endemic species per genus and genus-average genome size. This difference is not explained by the age and size of islands, total number of species sampled, or C -values, which are relatively similar. However, the average number of species used to calculate mean C -values per genus is about threefold lower for the Canaries compared to one for Hawaii (Table S1), and together with less resolved phylogeny this could make PIC analysis for the Canary Islands more conservative. Also from biogeographical point of view, the Canaries are close to Africa while Hawaii is a much more isolated archipelago with a higher proportion of endemic species. Geographical isolation and its consequences for colonization potential perhaps explain the higher number of endemic species per genus in Hawaii and might influence the relationship between genus-average genome size and endemic diversity statistics. Thus, significant negative correlation between numbers of endemic species per genus and genus-average genome size is not a universal feature of studied oceanic archipelagos and may depend on such factors as proximity to mainland or island size, with more local endemics on larger islands [46]. Hence, further work on a larger number of floras is required to test the generality of the LGCH.

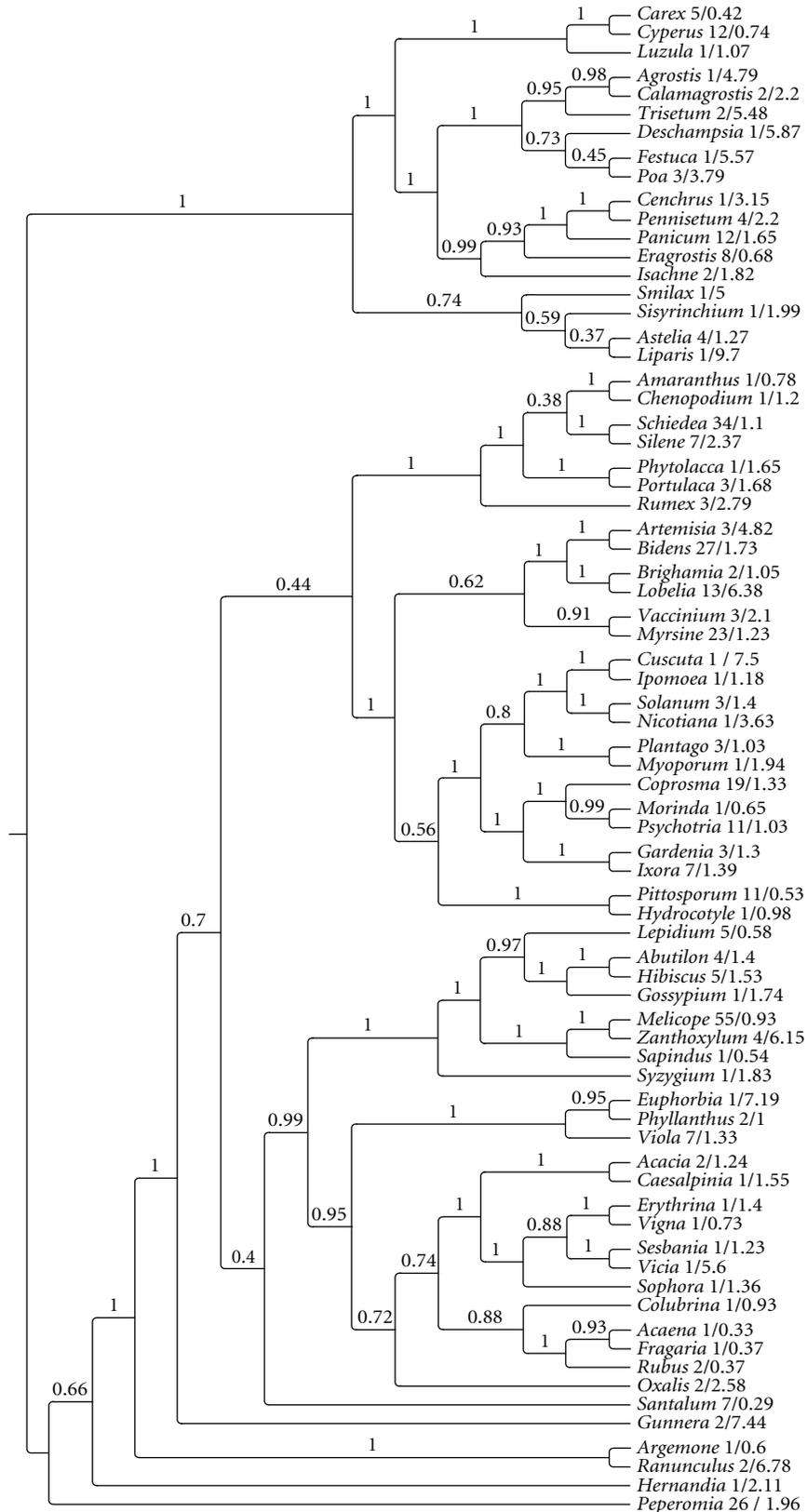


FIGURE 1: Bayesian phylogeny of the joint dataset of Hawaii and the Marquesas Islands based on *rbcl* sequences. Posterior probabilities are shown above branches; numbers of endemic species per genus and the genus-average genome size (1C, pg) are shown after genera names before and after slash, respectively. Felsenstein's contrasts correlation R between numbers of endemic species per genus and the genus-average genome size is -0.267 (P value = 0.011).

Acknowledgments

The authors thank Mark Chapman for comments on the paper and the Natural Environment Research Council, UK for funding.

References

- [1] J. Greilhuber, J. Doležel, M. A. Lysák, and M. D. Bennett, "The origin, evolution and proposed stabilization of the terms "genome size" and "C-value" to describe nuclear DNA contents," *Annals of Botany*, vol. 95, no. 1, pp. 255–260, 2005.
- [2] M. D. Bennett and I. J. Leitch, "Nuclear DNA amounts in angiosperms: targets, trends and tomorrow," *Annals of Botany*, vol. 107, no. 3, pp. 467–590, 2011.
- [3] J. Greilhuber, T. Borsch, K. Müller, A. Worberg, S. Porembski, and W. Barthlott, "Smallest angiosperm genomes found in Lentibulariaceae, with chromosomes of bacterial size," *Plant Biology*, vol. 8, no. 6, pp. 770–777, 2006.
- [4] J. Pellicer, M. F. Fay, and I. J. Leitch, "The largest eukaryotic genome of them all?" *Botanical Journal of the Linnean Society*, vol. 164, no. 1, pp. 10–15, 2010.
- [5] K. M. Devos, "Grass genome organization and evolution," *Current Opinion in Plant Biology*, vol. 13, no. 2, pp. 139–145, 2010.
- [6] M. Lynch and J. S. Conery, "The origins of genome complexity," *Science*, vol. 302, no. 5649, pp. 1401–1404, 2003.
- [7] P. SanMiguel, A. Tikhonov, Y. K. Jin et al., "Nested retrotransposons in the intergenic regions of the maize genome," *Science*, vol. 274, no. 5288, pp. 765–768, 1996.
- [8] A. E. Vinogradov, "Selfish DNA is maladaptive: evidence from the plant Red List," *Trends in Genetics*, vol. 19, no. 11, pp. 609–614, 2003.
- [9] C. A. Knight, N. A. Molinari, and D. A. Petrov, "The large genome constraint hypothesis: evolution, ecology and phenotype," *Annals of Botany*, vol. 95, no. 1, pp. 177–190, 2005.
- [10] I. J. Leitch, J. M. Beaulieu, M. W. Chase, A. R. Leitch, and M. F. Fay, "Genome size dynamics and evolution in monocots," *Journal of Botany*, vol. 2010, Article ID 862516, 18 pages, 2010.
- [11] M. D. Bennett and I. J. Leitch, "Genome size evolution in plants," in *The Evolution of the Genome*, T. R. Gregory, Ed., pp. 89–162, Elsevier, Amsterdam, The Netherlands, 2005.
- [12] K. D. Whitney, E. J. Baack, J. L. Hamrick et al., "A role for nonadaptive processes in plant genome size evolution?" *Evolution*, vol. 64, no. 7, pp. 2097–2109, 2010.
- [13] J. Suda, T. Kyncl, and V. Jarolímová, "Genome size variation in Macaronesian angiosperms: forty percent of the Canarian endemic flora completed," *Plant Systematics and Evolution*, vol. 252, no. 3–4, pp. 215–238, 2005.
- [14] B. Charlesworth and N. Barton, "Genome size: does bigger mean worse?" *Current Biology*, vol. 14, no. 6, pp. R233–R235, 2004.
- [15] B. Bremer, K. Bremer, M. W. Chase et al., "An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III," *Botanical Journal of the Linnean Society*, vol. 161, no. 2, pp. 105–121, 2009.
- [16] A. R. Wallace, *Island Life, Or, The Phenomena and Causes of Insular Faunas and Floras, Including a Revision and Attempted Solution of the Problem of Geographical Climates*, Harper Collins College, London, UK, 1881.
- [17] C. R. Darwin, *Journal of Researches into the Natural History and Geology of the Countries Visited During the Voyage of H.M.S. Beagle Round the World, under the Command of Capt. Fitz Roy, R.N.*, John Murray, London, UK, 1845.
- [18] W. L. Wagner and V. A. Funk, Eds., *Hawaiian Biogeography. Evolution on a Hot Spot Archipelago*, Smithsonian Institution Press, Washington, DC, USA, 1995.
- [19] T. F. Stuessy, G. Jakubowsky, R. S. Gómez et al., "Anagenetic evolution in island plants," *Journal of Biogeography*, vol. 33, no. 7, pp. 1259–1265, 2006.
- [20] J. A. Reyes-Betancort, A. Santos Guerra, I. R. Guma, C. J. Humphries, and M. A. Carine, "Diversity, rarity and the evolution and conservation of the Canary Islands endemic flora," *Anales del Jardín Botánico de Madrid*, vol. 65, no. 1, pp. 25–45, 2008.
- [21] J. Florence and D. H. Lorence, "Introduction to the flora and vegetation of Marquesas Islands," *Allertonia*, vol. 7, pp. 226–237, 1997.
- [22] G. Féraud, G. Giannérini, R. Campredon, and C. J. Stillman, "Geochronology of some Canarian dike swarms: contribution to the volcano-tectonic evolution of the archipelago," *Journal of Volcanology and Geothermal Research*, vol. 25, no. 1–2, pp. 29–52, 1985.
- [23] J. Suda, T. Kyncl, and R. Freiová, "Nuclear DNA amounts in Macaronesian angiosperms," *Annals of Botany*, vol. 92, no. 1, pp. 153–164, 2003.
- [24] J. P. Price and D. A. Clague, "How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence," *Proceedings of the Royal Society B*, vol. 269, no. 1508, pp. 2429–2435, 2002.
- [25] J. P. Price, "Floristic biogeography of the Hawaiian Islands: influences of area, environment and paleogeography," *Journal of Biogeography*, vol. 31, no. 3, pp. 487–500, 2004.
- [26] M. Arechavaleta, S. Rodriguez, N. Zurita, and A. Garcia, Eds., *Lista de Especies Silvestres de Canarias. Hongos, Plantas y Animales Terrestres*, Gobierno de Canarias, 2010.
- [27] W. L. Wagner, D. R. Herbst, and D. H. Lorence, Flora of the Hawaiian Islands website, 2005, <http://botany.si.edu/pacificislandbiodiversity/hawaiianflora/index.htm>.
- [28] W. L. Wagner and D. H. Lorence, Flora of the Marquesas Islands website, 2002, <http://botany.si.edu/pacificislandbiodiversity/marquesasflora/index.htm>.
- [29] M. D. Bennett and I. J. Leitch, Plant DNA C-values Database, release 5.0, 2010, <http://www.kew.org/genomesize/homepage>.
- [30] M. V. Kapralov, M. Stift, and D. A. Filatov, "Evolution of genome size in Hawaiian endemic genus *Schiedea* (Caryophyllaceae)," *Tropical Plant Biology*, vol. 2, no. 2, pp. 77–83, 2009.
- [31] J. Felsenstein, "Phylogenies and the comparative method," *American Naturalist*, vol. 125, no. 1, pp. 1–15, 1985.
- [32] W. P. Maddison and D. R. Maddison, Mesquite: a modular system for evolutionary analysis. Version 2.72, 2009.
- [33] P. E. Midford, T. Garland Jr., and W. Maddison, PDAP:PDTree package for Mesquite, version 1.12, 2008.
- [34] National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov>.
- [35] J. P. Huelsenbeck and F. Ronquist, "MRBAYES: bayesian inference of phylogenetic trees," *Bioinformatics*, vol. 17, no. 8, pp. 754–755, 2001.
- [36] F. Ronquist and J. P. Huelsenbeck, "MrBayes 3: bayesian phylogenetic inference under mixed models," *Bioinformatics*, vol. 19, no. 12, pp. 1572–1574, 2003.
- [37] C. O. Webb and M. J. Donoghue, "Phylomatic: tree assembly for applied phylogenetics," *Molecular Ecology Notes*, vol. 5, no. 1, pp. 181–183, 2005.
- [38] S. R. Downie, M. F. Watson, K. Spalik, and D. S. Katz-Downie, "Molecular systematics of Old World Apioideae

- (Apiaceae): relationships among some members of tribe Peucedaneae sensu lato, the placement of several island-endemic species, and resolution within the apioid superclade," *Canadian Journal of Botany*, vol. 78, no. 4, pp. 506–528, 2000.
- [39] S. C. Kim, M. R. McGowen, P. Lubinsky, J. C. Barber, M. E. Mort, and A. Santos-Guerra, "Timing and tempo of early and successive adaptive radiations in Macaronesia," *PLoS One*, vol. 3, no. 5, Article ID e2139, 2008.
- [40] M. E. Mort, D. E. Soltis, P. S. Soltis, J. Francisco-Ortega, and A. Santos-Guerra, "Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data," *Systematic Botany*, vol. 27, no. 2, pp. 271–288, 2002.
- [41] T. L. P. Couvreur, A. Franzke, I. A. Al-Shehbaz, F. T. Bakker, M. A. Koch, and K. Mummenhoff, "Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae)," *Molecular Biology and Evolution*, vol. 27, no. 1, pp. 55–71, 2010.
- [42] J. F. Y. Brookfield and R. M. Badge, "Population genetics models of transposable elements," *Genetica*, vol. 100, no. 1–3, pp. 281–294, 1997.
- [43] M. Kubešová, L. Moravcová, J. Suda, V. Jarošík, and P. Pyšek, "Naturalized plants have smaller genomes than their non-invading relatives: a flow cytometric analysis of the Czech alien flora," *Preslia*, vol. 82, no. 1, pp. 81–96, 2010.
- [44] S. Lavergne, N. J. Muenke, and J. Molofsky, "Genome size reduction can trigger rapid phenotypic evolution in invasive plants," *Annals of Botany*, vol. 105, no. 1, pp. 109–116, 2010.
- [45] E. Grotkopp, M. Rejmánek, M. J. Sanderson, and T. L. Rost, "Evolution of genome size in pines (*Pinus*) and its life-history correlates: supertree analyses," *Evolution*, vol. 58, no. 8, pp. 1705–1729, 2004.
- [46] A. C. Algar and J. B. Losos, "Evolutionary assembly of island faunas reverses the classic island-mainland richness difference in *Anolis* lizards," *Journal of Biogeography*, vol. 38, no. 6, pp. 1125–1137, 2011.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

