Molecular Phylogenetics and Evolution 68 (2013) 498–501

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Short Communication

Regional population expansion in *Eucalyptus globulus*

Suat Hui Yeoh [a,b,*], Simon Y.W. Ho [a,c], Andrew H. Thornhill [a], William J. Foley [a]

[a] Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia
[b] Institute of Biological Sciences, Faculty of Science, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia
[c] School of Biological Sciences, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 9 October 2012
Revised 11 March 2013
Accepted 19 April 2013
Available online 3 May 2013

Keywords:
Bayesian skyline plot
Demographic history
Eucalyptus globulus

ABSTRACT

Foundation tree species define the structure of forest habitat and influence their ecosystem dynamics. However, there is limited understanding of both the patterns and timing of population fluctuations in foundation trees and how they vary among geographical regions. We have reconstructed the demographic history of five genetically distinct populations of the Tasmanian blue gum (*Eucalyptus globulus* ssp. *globulus*) at the species and regional levels, using three nuclear loci sequenced from 104 individuals. Analysis using a Bayesian skyline plot indicated that the species experienced two periods of expansion, commencing in the Pliocene. Regional analyses showed that island populations expanded earlier, but that the rate of expansion was relatively slow when compared to that of the mainland group. This highlights the need for local demographic history to be taken into account when inferring local adaptation for candidate genes. Population growth throughout the Quaternary signals the ability of the species to persist and thrive under the predominantly harsh conditions of this period.

1. Introduction

Past climate change has strongly influenced the spatial diversity and distribution of the Australian biota (Markgraf et al., 1995). During times of adverse climatic conditions, organisms contracted to coastal refugia or migrated to higher elevations, recolonizing areas when conditions became favorable again (Macphail, 1979). The signature of these past population dynamics persists in the DNA of long-lived forest trees, which are important drivers of the whole forest community. However, our understanding of the timing and patterns of population fluctuation through time is still lacking for tree species (Pautasso, 2009). Even poorer is an understanding of demographic history at the local scale, which defines the intraspecific genetic diversity and provides a foundation for the study of local adaptation (e.g., Ross-Ibarra et al., 2008). This is particularly important for foundation tree species that are subject to heterogeneous demographic processes across their distribution. A reliable estimate of demographic history is crucial for understanding historical survival range, climatic tolerance, and constraints on species distributions.

The Tasmanian blue gum (*Eucalyptus globulus* ssp. *globulus*, hereafter referred to as *E. globulus*) is an economically and ecologically important, long-lived lowland forest tree species. It occurs naturally in southeastern Australia and Tasmania, a region that also encompasses the Bass Strait islands, and has been planted extensively worldwide. Many species-wide studies conducted using molecular markers have found evidence of phylogeographic structure (e.g., Yeoh et al., 2012). Despite extensive studies, there remains a limited understanding of how the sizes of different *E. globulus* populations have changed through time. Elucidating the temporal and spatial demographic patterns of *E. globulus* will improve our knowledge of the evolutionary processes driving the genetic diversity of populations through time. Studying foundation tree species such as *E. globulus* is important because of the impact that their genetic variation has on wider ecosystem processes (Barbour et al., 2009). Genetic variation in foundation tree species can determine the evolution and structure of the forest community.

The application of coalescent theory has led to the development of methods for inferring complex population histories from DNA sequences. These include a group of methods known as skyline plots (for a recent review, see Ho and Shapiro, 2011), which take advantage of the relationship between the genealogical and demographic histories of a population (Pybus et al., 2000). Heled and Drummond (2008) developed the extended Bayesian skyline plot (EBSP), which enables the co-estimation of the genealogy, timescale, and population history using multiple loci. Simultaneous analysis of multiple loci can substantially reduce estimation error in demographic reconstruction.

To understand the influence of fluctuations in regional population sizes on the demographic history of *E. globulus*, we used

* Corresponding author at: Institute of Biological Sciences, Faculty of Science, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia. Fax: +60 3 79675908.
E-mail addresses: suathui_yeoh@um.edu.my (S.H. Yeoh), simon.ho@sydney.edu.au (S.Y.W. Ho), Andrew.Thornhill@anu.edu.au (A.H. Thornhill), William.Foley@anu.edu.au (W.J. Foley).

1055-7903/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2013.04.020

extended Bayesian skyline plots to infer the demographic history of the species and of its genetically distinct populations. We also estimated the timing of notable demographic events and studied the confounding effects of selection.

2. Methods

The plant material used in our study is a subset of that used by Yeoh et al. (2012). This subset comprised 108 *E. globulus* individuals distributed across 11 populations, with about 10 individuals per population (Table 1; Fig. 1). These populations represent five genetically homogeneous regions detected using microsatellite markers (Yeoh et al., 2012). Homologous sequences from *E. nitens*, *E. camaldulensis*, and *E. loxophleba* were used for estimating the evolutionary timescale of reconstructed demographic history (see below). Genomic DNA from all samples was extracted according to the method described by Glaubitz et al. (2001).

The nuclear genes used in this study are 1-deoxy-D-xylulose-5-phosphate reductoisomerase (*dxr*) and two copies of 1-deoxy-D-xylulose-5-phosphate synthase (*dxs1* and *dxs2*). The nuclear genes are located on different chromosomes of the 8X *Eucalyptus grandis* genome (http://www.phytozome.net/eucalyptus.php); 1-deoxy-D-xylulose-5-phosphate reductoisomerase (*dxr*) is located on chromosome 2 (Egrandis_v1_0.011328m.g (4528455 – 4533481 bp)), 1-deoxy-D-xylulose-5-phosphate synthase 1 (*dxs1*) on chromosome 8 (Egrandis_v1_0.004251m.g (53734930 – 53738726 bp)) and *dxs2* on chromosome 9 (Egrandis_v1_0.004068m.g (11629268 - 11632658 bp)). We obtained the full-length reference sequence of these genes and amplified them for each of 108 samples of *E. globulus* separately. Amplicons were pooled according to individuals, sheared into smaller fragments, and barcoded using the parallel tagged sequencing method described by Meyer et al. (2008) before being sequenced using the GS FLX System (454 Life Sciences, Branford, CT, USA). Consensus sequences for each individual were generated from trimmed sequence alignments, sorted according to reference sequence, and manually edited. Owing to the low number of variants found in the sequences, each individual was represented by only a single sequence per locus, with the assumption that a random copy was sampled from each individual. Details of the genes, PCR parameters, and barcode and sequence-assembly methods are provided in Supplementary data 1. We resequenced the homologous genes for *E. nitens*, *E. camaldulensis*, and *E. loxophleba* on an ABI 3100 Genetic Analyzer to enable estimation of the evolutionary timescale.

Both *dxs* and *dxr* loci code for enzymes in the terpene biosynthesis pathway. To estimate the linkage disequilibrium between pairs of segregating sites, we analyzed the three loci using TASSEL version 3 (Bradbury et al., 2007). Linkage disequilibrium was estimated for the full dataset and for each genetically distinct population separately. Linkage disequilibrium for segregating sites that were monomorphic in certain genetically distinct populations was not estimated for the respective populations and uninformative results were discarded.

To examine the effect of selection on the demographic signal, we conducted separate analyses of (i) first and second codon sites and (ii) third codon sites and introns. To infer demographic history, alignments from the three loci were analyzed simultaneously using the extended Bayesian skyline plot (EBSP; Heled and Drummond, 2008), implemented in the phylogenetic software BEAST version 1.6 (Drummond and Rambaut, 2007). Given that the presence of population structure in a species can produce biases in the reconstruction of demographic history (Pannell, 2003), we also conducted a separate EBSP analysis for each of the five genetically distinct populations of *E. globulus*. The data and details of the analyses are given in Supplementary materials.

The EBSP produces a single plot of population size through time, with the *x*-axis measured in mutations per site. To convert this into a timescale, the units on the *x*-axis were divided by $\sum \bar{d}_i l_i/(2t\sum l_i)$, where $\bar{d}_i$ is the mean pairwise genetic distance between two taxa for locus *i*, corrected using the substitution model selected using the Bayesian information criterion, l_i is the alignment length of locus *i*, and *t* is the time since the divergence of the two taxa. We performed this calculation using pairwise comparisons of *E. globulus* with *E. nitens*, *E. camaldulensis*, and *E. loxophleba*, and computed the arithmetic mean. We estimated *t* using a molecular-clock analysis of two chloroplast genes and a nuclear spacer region (Supplementary data 1). It should be noted that our approach precludes a straightforward quantification of estimation error.

3. Results

The averaged correlation between alleles at two loci (r^2) was 0.052 in *dxr*, 0.023 in *dxs1*, and 0.029 in *dxs2*. Of the pairwise comparisons, approximately 11% were significant (<0.05) in *dxr*, 6.3% in *dxs1*, and 9.0% in *dxs2*. When genetically distinct populations were considered separately, within-region average r^2 values ranged from 0.138 to 0.192 for *dxr*, 0.080 to 0.116 for *dxs1*, and 0.064 to 0.106 for *dxs2*. The r^2 value decayed quickly with distance for all three loci. Low linkage disequilibrium indicated the absence of recent selective sweep in these sequences.

The population history of *E. globulus*, estimated from the third codon sites and introns of three nuclear loci, is shown in Fig. 2a. The skyline plot indicates that there were two phases of population expansion in this species, commencing around 4.5 million years ago (Mya) and 1.3 Mya, respectively. The number of changes in population size was estimated to be 2.22 (95% credibility interval: 2–3). In contrast, the population history inferred from the first and second codon sites of the three loci was far less informative

Table 1
The number of individuals and location of the 11 populations from which those individuals were derived across five genetically homogeneous regions of *Eucalyptus globulus*.

Region	Population	No. of individuals	Longitude (°E)[a]	Latitude (°S)[a]	Altitude (m)
Otway	1. Parker Spur	10	143.59	38.82	60–200
	2. Lorne	11	143.95	38.53	100–280
Gippsland	3. Jeeralang North	9	146.53	38.35	220–460
	4. Hedley	8	146.50	38.62	20–200
Furneaux	5. Central Flinders Island	10	148.05	40.05	140–240
	6. West Cape Barren	10	148.07	40.40	20–220
Eastern Tasmania	7. St. Helens	8	148.30	41.27	120
	8. Jericho	9	147.27	42.42	500
	9. Moogara	10	146.91	42.78	430–500
Western Tasmania & King Island	10. Little Henty River	9	145.20	41.93	10
	11. Central King Island	10	144.00	40.00	20–100
	Total	104			

[a] Coordinates for each population were chosen based on the averages of natural stands recorded by Gardiner and Crawford (1987, 1988).