



Article Allele Surfing and Holocene Expansion of an Australian Fig (*Ficus*—Moraceae)

Brendan C. Wilde ^{1,*}, Susan Rutherford ², Jia-Yee S. Yap ¹ and Maurizio Rossetto ¹

- Research Centre for Ecosystem Resilience, Australian Institute of Botanical Science, The Royal Botanic Garden, Sydney 2000, Australia; samantha.yap@rbgsyd.nsw.gov.au (J.-Y.S.Y.); Maurizio.Rossetto@rbgsyd.nsw.gov.au (M.R.)
- ² Institute of Environment and Ecology, Academy of Environmental Health and Ecological Security, Jiangsu University, Zhenjiang 212013, China; susan_rutherford@ujs.edu.cn
- Correspondence: brendan.wilde@rbgsyd.nsw.gov.au

Abstract: The creek sandpaper fig of southeastern Australia, *Ficus coronata* Spin, is culturally significant to Australian traditional owners who made use of the leaves to smooth timber and ate the fruit. The species is thought to have a long history on the continent, with some suggesting a Gondwanan origin. However, distributional patterns and overall ecology suggest a recent expansion across suitable habitats. We used landscape genomic techniques and environmental niche modelling to reconstruct its history and explore whether the species underwent a recent and rapid expansion along the east coast of New South Wales. Genomic analysis of 178 specimens collected from 32 populations throughout the species' New South Wales distribution revealed a lack of genetic diversity and population structure. Some populations at the species' southern and western range limits displayed unexpected diversity, which appears to be the result of allele surfing. Field work and genetic evidence suggest a Holocene expansion which may have increased since European colonisation. We also present a novel method for detecting allele surfing—MAHF (minor allele at highest frequency).

Keywords: SNP; sandpaper fig; Ficus; allele surfing; range expansion

1. Introduction

Post-glacial range expansions, from low latitudes towards the poles, have been reported for multiple species globally [1–4]. During the Pleistocene, the climate oscillated between glacial and interglacial periods, repeatedly forcing species to contract into refugia (or become extinct), then expand as conditions improved [5,6]. Periods of contraction condensed populations into isolated gene pools, while expansion enabled populations to grow and diverge in response to new environments [7,8] or admix on contact with new populations [9]. These contraction and expansion cycles formed the basis from which today's species distribution and population structure developed [10].

Species' range expansion can be described as having a leading edge, with the gaps being filled as others migrate behind. As repeat founding events migrate further from the source, genetic diversity is reduced and homozygosity increases. Colonising populations consisting of individuals derived from a larger source population start with a random subset of the source population's allelic diversity, equating to a loss of allelic diversity or genetic drift. Being random, genetic drift is not caused by the environment, although selective pressures will ultimately determine which alleles remain. Genetic drift and decreasing relatedness with increasing geographic distance [11] contribute to the process of isolation by distance (IBD [12–14]). This means opposite ends of a large continuous species distribution can end up with distinct allelic variation [12]. Low allelic diversity and weak population structure can identify recent range expansion [15], while refugial populations often have high allelic diversity and strong population structure [16].



Citation: Wilde, B.C.; Rutherford, S.; Yap, J.-Y.S.; Rossetto, M. Allele Surfing and Holocene Expansion of an Australian Fig (*Ficus*—Moraceae). *Diversity* **2021**, *13*, 250. https:// doi.org/10.3390/d13060250

Academic Editors: Michael Wink and Mario A. Pagnotta

Received: 21 May 2021 Accepted: 4 June 2021 Published: 7 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). Low and uniform variation is often observed across the distribution of many Australian rainforest species with southeast Asian (Sunda) heritage, the result of recent migration and a short expansion period within the Australian continent [4,17–19]. Over time, as migrant populations increase in size and number, allelic diversity can also increase, either through mutations [20] or admixture when encountering closely related individuals [21].

The east coast of Australia has been shown to provide refugia for many rainforest species [22–26], especially in north and southeast Queensland (QLD) and northeast New South Wales (NSW). Das et al. [24] used climate models to identify potential rainforest refugia in Australia from which species expanded during the Holocene warming. Several recent studies explored the concept of Holocene range expansion in Australia. Fahey et al. [25] revealed the Holocene expansion of two sympatric *Tristaniopsis* species differed, with one rapidly expanding its range while the other maintained a more stable distribution. A study of species accumulation through time across three eastern NSW rainforest communities [26] found that the majority of species investigated had expanded into the study regions as recently as the late Pleistocene and Holocene.

Globally there are multiple reports of Holocene expansions of *Ficus* species. In South America, De Almeida Vieira et al. [27] show a southward Holocene expansion of *F. bonijesulapensis* from northern refugia. In the northern hemisphere, Nason et al. [28] discuss the northern range expansion of *F. hirta* within China from two or more southern refugia. Lin et al. [29] reported a drastic post-Last Glacial Maximum (LGM) expansion of a pollinator wasp species, potentially linked to the range extension of its host *F. spetica*, in Taiwan. Chen et al. [30], again based on pollinator wasps, concluded that the Chinese fig *F. pumila* expanded from multiple refugia. Ours, however, is the first study using multiple lines of evidence to investigate the putative Holocene expansion of an Australian *Ficus*.

Commonly referred to as the creek sandpaper fig, *F. coronata* (Figure 1) is one of ~80 species [31] in the monophyletic subgenus *Sycidium* [32], which are found in Africa, Asia and Australia [31]. The group has 42 species in the region of Malesia (including Malaysia, Indonesia, Philippines and New Guinea) and 13 species in Australia. The creek sandpaper fig, *F. coronata*, and the shiny sandpaper fig, *F. fraseri* Miq, are endemic to the east coast of NSW and QLD. Another six *Sycidium* species, *F. aculeata* A. Cunn. ex Miq, *F. carpentariensis* D.J. Dixon (= *F. aculeata x coronulata*, see [33]), *F. coronulata* Miq, *F. leptoclada* Benth, *F. podocarpifolia* Corner and *F. scobina* Benth, are endemic to tropical northern Australia. The remaining five Australian species, *F. copiosa* Steud, *F. melinocarpa* Blume, *F. opposita* Miq, *F. tinctoria* G.Forst and *F. virgata* Reinw. ex Blume, also occur outside of Australia.



Figure 1. (a) A small multitrunked *Ficus coronata* of around 7 m tall, Royal National Park, NSW, Australia; (b) Sandpaper-like foliage and young figs. Photography by B.C. Wilde.

As suggested by its common name, *F. coronata* has sandpaper-like leaves that were used by Aboriginal Australians to smooth timber; the fleshy purple fruit from female trees (the species is gynodioecious) is edible. The species name 'coronata' refers to the crown-like ring of bracts around each fig's ostiole. Fig fruits are eaten by many animals [34], including birds and bats [35,36], that can distribute the seed over long distances. Most commonly found around the NSW–QLD border (Figure 2), an area that includes multiple persistent rainforest refugia, the species is habitat-specific and highly dispersible, making it a good subject to study post-glacial expansion patterns.



Figure 2. Distribution of *Ficus coronata* based on herbarium data. Vouchers from Western Australia (WA), Northern Territory (NT) and Far North Queensland (FNQ) were found to be in error and are not shown. Atlas of Living Australia (ALA) (2020) *Ficus coronata* occurrence records. Available at: ala.org.au, accessed on 30 May 2020).

Using genome-wide reduced-representation sequencing (DArTseq) and environmental niche modelling (ENM), we investigated the landscape genomic patterns of *F. coronata* in order to ascertain a recent migration event into the southern range of its distribution. In particular, we ask: (1) can signals for rapid expansion of *F. coronata* be detected; and (2) can directionality of migration be determined?

2. Material and Methods

2.1. Sampling Strategy

At the time of writing, *F. coronata* was recorded for five Australian states and territories (Atlas of Living Australia (ALA) (2020) *Ficus coronata* occurrence records. Available at: ala.org.au, accessed on 30 May 2020). The species is most common in NSW with 4333 records, southeast QLD with 413 (classified as least concern, ALA 2021), northeastern Victoria (VIC) with 46 (listed as threatened, Flora and Fauna Guarantee Act 1988—Threatened List November 2019), Northern Territory with five and Western Australia one. Images of collections with doubtful identifications for Northern Territory (NT), Western Australia (WA) and Far North Queensland (FNQ) were requested from the BRI, CANB, VIC and SA herbaria. None of the sighted specimens were *F. coronata*. One specimen, CANB 268023, is *F. coronulata*; MEL 1581926 and MEL 1581933 are possibly *F. aculeata x coronulata* [33]. Three specimens, BRI AQ1005733, BRI AQ0008071 and MEL 0274275A from northern QLD, are likely *F. opposita*. This makes *F. coronata* an endemic of south east coast Australia (Figure 2).

To investigate the hypothesis of a recent southern expansion, leaf material was collected from 178 *F. coronata* individuals from 32 sites throughout the species' NSW range (Table 1) following the Restore and Renew collection protocol [37]. Where possible, six individuals were collected from each population, though some populations had less than six individuals. Individuals at least 10 m apart and older plants were prioritised. Data for samples were collected using the Restore and Renew mobile app [38]. Samples were stored at -80 °C before freeze drying and then placed in airtight containers with silica beads.

Table 1. Locations of *Ficus coronata* populations with a count of samples used in this study. Six samples were collected from Royal National Park, Red Head and Sea Acres though not all were successfully sequenced.

Population	Location	Latitude	Longitude	Samples
01	Brunby Plains	-28.4874	152.6518	6
02	Billinudgel	-28.4976	153.5372	6
03	Montecollum	-28.5919	153.4546	6
04	Nightcap Ridge	-28.6113	153.3647	6
05	Lismore	-28.8034	153.2809	6
06	Washpool	-29.4724	152.3302	5
07	Dorrigo	-30.3947	152.7459	2
08	Mount Kaputar National Park	-30.5970	150.2961	6
09	Oxley Wild Rivers National Park	-30.7429	152.0148	4
10	Port Macquarie	-31.4597	152.9213	6
11	Sea Acres National Park	-31.4641	152.9304	4
12	Red Head	-32.0438	152.5324	5
13	Barrington Tops	-32.0662	151.6635	6
14	Cellito Beach	-32.3782	152.5283	3
15	Fern Tree Gully	-32.6595	150.0347	6
16	Putty Road	-32.7304	150.9415	6
17	Colo River	-33.4319	150.8311	6
18	Yarramundi Reserve	-33.6098	150.7142	6
19	Govetts Leap, Blue Mountains	-33.6291	150.3168	6
20	Berowra River	-33.6654	151.0811	6
21	Depot Beach	-33.7960	151.0681	6
22	Royal National Park	-34.1535	151.0286	5
23	Stanwell Tops	-34.2270	150.9805	6
24	Macquarie Pass	-34.5731	150.5731	6
25	Minnamurra Forest	-34.6349	150.7246	6
26	Kangaroo Valley	-34.7607	150.6161	6
27	Bungonia National Park	-34.7995	150.1387	6
28	Wandandian Creek	-35.0982	150.5032	6
29	Mount Agony	-35.7278	150.1973	6
30	Deua National Park	-35.9640	149.9360	6
31	Gulaga National Park	-36.3179	150.0741	6
32	Brogo	-36.5569	149.8180	6

2.2. Genome Scans and Data Analysis

A 5–10 mg section of leaf from each specimen was provided to Diversity Arrays Technology (DArT; http://www.diversityarrays.com, accessed on 30 May 2020) for extraction and sequencing. DArTseq uses restriction enzymes targeting low-copy sequences (mainly nuclear) and next-generation sequencing (NGS) to provide thousands of markers across the genome. DArT supplied genetic data in the form of 109,377 single-nucleotide polymorphisms (SNPs) for each sample using a proprietary analytical pipeline. The dataset was pre-processed with an in-house R package RRtools v. 1.0 (unpublished; Jason Bragg, Royal Botanic Garden Sydney) as per the Restore and Renew workflow described in [37], removing low quality (reproducibility ≤ 0.98 and missingness ≥ 0.05) and duplicate SNPs, resulting in 7492 SNPs that were used in the analyses.

Genetic structure and admixture analyses were performed with LEA, an R package for landscape and ecological association [39], and mapped with the R library oz [40] and mapplots [41]. The best value for K was determined using snmf from LEA by plotting cross entropy for K = 2 to 10 and selecting the most suitable K from the average of 10 runs. Data for populations and individuals were run separately to check for conformity.

Genetic data were converted to genlight format and Principal Component Analysis (PCA) was performed using the function glPca (Principal Component Analysis for genlight objects) of the R package adegenet [42,43]. Population genetic measures were obtained to explore the distribution of genetic variation and connectivity between populations. The fixation index (F_{ST}) was estimated using R package SNPrelate [44] using W&H02—relative beta estimator [45]. This was also used in the F_{ST} by distance Mantel created with the R package Vegan [46]. Estimates for population allelic richness (A_R), expected heterozygosity (H_E), observed heterozygosity (H_O) and inbreeding coefficient (F_{IS}) were calculated with the R package diveRsity [47]. Private allele counts were performed with the R package poppr [48].

Range expansion analysis was performed with the R package rangeExpansion [49]. Genetic data were converted to snapp format and all regions were tested by supplying a NULL region list. Falistocco [50] notes that figs are commonly diploid so the ploidy was assumed to be two.

With the rangeExpansion results lacking sensitivity to a recently expanded population such as *F. coronata* and minor allele frequency (MAF) known to strongly affect inferred population structure [51], we experimented with using MAF to detect allele surfing. A table of minor alleles for each site was created with the R package RRtools. We then counted the occurrences of each minor allele for each population. An allele was regarded as MAHF (minor allele at highest frequency) for the population with the highest occurrence count. Each population can have 0 or more MAHF but an allele can only be MAHF for one population. All MAHF were counted for each population and graphed. This provides a basic visual of minor alleles being lost (through processes such as bottlenecking) and gained (stochastically, adaptively or through allele surfing).

Environmental niche modelling (ENM) was used to estimate possible past and future regions with suitable climate for this species, and uncover possible refugia and expansion directionality. The R library Kuenm [52] was used to create ENMs following the steps outlined in https://github.com/marlonecobos/kuenm (accessed on 5 January 2021). The package automates the creation of multiple Maxent models [53] using various regularization parameters, combinations of feature classes and environmental predictors. The package can also test all models it creates and select an optimal model and transfer this to different environmental layers or eras.

Collection records for ENM creation were downloaded from the Atlas of Living Australia (ALA; http://ala.org.au, accessed on 30 May 2020) and filtered to remove duplicates, records before 1950 and records with location accuracy greater than 100 m. The remaining records were spatially thinned with the R package spThin [54] set to 20 km, removing multiple records in close proximity and the possible bias caused to the modelling. Four eras of environment layers were downloaded using the R function getData() from the R package raster [55]: The Last Glacial Maximum (LGM, paleoclim('lgm', '2.5 m') [56] and Mid Holocene (MH, paleoclim('mh', '2.5 m') [57,58], from http://www.paleoclim.org, accessed on 5 January 2021; the current era ('worldclim', var = 'bio', res = 2.5), from worldclim (https://www.worldclim.org, accessed on 5 January 2021); and future era ('CMIP5', var = 'bio', res = 2.5, rcp = 85, model = 'HE', year = 50) (Coupled Model Intercomparison Project Phase 5) (https://pcmdi.llnl.gov/mips/cmip5/, accessed on 5 January 2021).

Layer names for environment sets LGM and MH were renamed and ordered to match those in the current and future sets. All environmental sets were cropped to include Papua New Guinea and Eastern Australia. Four of the 19 environmental layers were removed because of their possible bias effect [59], while others were removed after tests with maxent showed they had no effect of the model. The remaining layers were placed into three different sets of reduced size for use in kuenm. Set 1 included bio layers: 1, 2, 5, 6, 7, 11, 12, 13, 14, 15, 16, 17; Set 2: 1, 5, 6, 11, 12, 14, 15, 17; Set 3: 1, 12, 15, 17. Kuenm created 1581 candidate models before selecting the final model from which the 4 different eras were created with no extrapolation. On completion, kuenm creates a Final_Model directory containing a .bat file for each model tested, in this case LGM, MH, Current and Future. Images of the models were generated by copying the Java code contained in the .bat files and running it in a bash terminal window.

Data visuals were combined and edited in Adobe inDesign (Adobe Inc., San Jose, CA, USA, 2019. Adobe InDesign, Available at: https://adobe.com/products/indesign, accessed on 10 June 2020).

3. Results

Genetic Structure

Results for population genetic structure derived from LEA admixture results (Figure 3) show an overall north to south latitudinal gradient (population 01—Brunby Plains to 32— Brogo), with the western extreme population (08—Mount Kaputar National Park) and most southerly (32—Brogo) resolving as distinct from the other sites.

The Mantel test (F_{ST} by distance Figure 4; r = 0.3181, P = 0.001) showed a strong association between geographic and genetic distances, excluding populations 08 and 32, (r = 0.3966, P = 0.0001). Populations 08 and 32 show variation probably not explained by isolation by distance (see Appendix A Figure 1) or admixture (see Appendix B Figure A2), with no other sandpaper fig species nearby. Pairwise comparison of allele frequencies (see Appendix B Table A1) shows populations 08 and 32 have the highest levels of differentiation. The highest pairwise differentiation (F_{ST} = 0.395) was between populations 08 and 32 and the smallest was between populations 01 and 05 (F_{ST} = 0.013).

Principal component analysis (PCA, Figure 5) of genetic data for all 178 individuals shows very little clustering into their respective populations, with most individuals forming one continuous cluster in a broad north to south pattern. The western population (08) and the most southern population (32) are separated from all other populations, though overall variation displayed is very low (PC1 = 3.9%, PC3 = 2.9%; Figure 5).

MAHF (minor allele at highest frequency, Figure 6) analysis showed an overall loss of minor alleles from north to south (population 01 to 32). There was also a loss of alleles in an east to west direction (population 01 to 08). The MAHF count is almost zero before a large peak at the western extreme (population 08), and reduces again before another smaller peak for the most southern (population 32).



Figure 3. Sampling sites are numbered from most northern (01—Brunby Plains) to most southern (32—Brogo); see Table 1 for details. (**A**) Geographic location and population mean LEA genetic admixture (K = 4); (**B**) LEA genetic admixture (K = 4) for each individual.



Figure 4. Mantel test (F_{ST} by distance), pairwise comparisons involving population 08—Mount Kaputar National Park (western extreme), are highlighted in blue and population 32—Brogo (southern extreme), in red. Individuals from all other populations are shown in black.



Figure 5. Principal component analysis (PCA) of genetic data for all 178 individuals in this study. Population 08—Mount Kaputar National Park (western extreme) is highlighted in blue and population 32—Brogo (southern extreme), in red.



Figure 6. MAHF (minor allele at highest frequency). Populations are numbered from most northern (01—Brunby Plains) to most southern (32—Brogo). Expansion to western extreme (08—Mount Kaputar National Park) is highlighted in blue and to southern extreme (32—Brogo), in red.

Range expansion analysis (Figure 7A) highlighted all of the northern distribution as a possible origin of expansion and the southern limits as being non-origin, suggesting a general southerly expansion pattern.

	2				1 0
Population	A_R	H_E	H_O	F _{IS}	P_A
01	1.458	0.274	0.265	^H 0.018	^H 154
02	1.449	0.271	^L 0.260	^H 0.020	^H 222
03	1.451	0.268	0.266	-0.006	^H 150
04	1.430	0.257	0.265	-0.038	115
05	^H 1.470	^H 0.276	0.274	-0.006	^H 149
06	1.455	0.271	0.266	0.001	149
07	* ^L 1.368	* ^L 0.208	0.266	* ^L -0.290	* ^L 28
08	^L 1.347	^L 0.196	^L 0.236	^L -0.183	51
09	1.446	0.261	0.269	-0.047	69
10	1.452	0.264	0.273	-0.039	146
11	1.451	0.261	^H 0.277	-0.078	95
12	^H 1.472	^H 0.275	^H 0.276	-0.020	71
13	^H 1.474	^H 0.278	^H 0.276	-0.006	61
14	1.417	^L 0.236	0.272	^L -0.159	67
15	1.421	0.242	0.269	^L -0.108	59
16	1.466	0.273	0.276	-0.022	101
17	1.460	0.271	0.266	0.000	73
18	1.465	0.269	0.273	-0.028	87
19	1.438	0.248	0.271	-0.086	52
20	1.438	0.256	0.275	-0.065	69
21	^H 1.470	^H 0.275	0.270	^H 0.005	67
22	1.461	0.271	0.274	-0.028	56
23	1.465	0.274	0.275	-0.017	77
24	1.460	0.271	0.269	-0.004	54
25	1.464	0.272	0.268	-0.001	69
26	1.462	0.273	0.264	^H 0.008	58
27	^L 1.415	0.244	^L 0.243	-0.012	46
28	1.463	0.266	^H 0.277	-0.045	L 39
29	^L 1.413	^L 0.234	^L 0.261	-0.105	L 33
30	1.442	0.258	0.267	-0.043	^L 44
31	1.446	0.261	0.261	-0.012	51
32	^L 1.329	^L 0.177	^L 0.259	^L -0.34	L 30

Table 2. Population genetic diversity estimates for *Ficus coronata* collected from 32 NSW locations, showing population allelic richness (A_R), expected (H_E) and observed (H_O) heterozygosity, inbreeding (F_{IS}), and private alleles (P_A). Prefix of ^L shows the 4 lowest results for each variable and ^H the 4 highest. See Figure 7 for mapped highs and lows and Table 1 for site population details of populations. * Possible biased results with only 2 individuals collected, this site has been excluded from the maps in Figure 7.



Figure 7. (**A**) Range expansion analysis; blue shows possible origin of expansion and red shows non-origin regions. (**B**–**D**) Population genetic diversity estimates (Table 2) mapped for *Ficus coronata*. The four highest results are shown in blue, the four lowest in red. (**B**) Population allelic richness (A_R). (**C**) Heterozygosity observed (H_O). * Population 02 in red at the northeast boundary is unexpectantly low, though the difference is only 0.001 between it and populations 29 and 30 at the southern boundary. (**D**) Private alleles (P_A); the third highest value is the same for two sites.

Genetic diversity estimates for the 32 NSW populations in this study (Table 2 and Figure 7) show allelic richness (A_R ; Figure 7B) is lowest for the western extreme (08, $A_R = 1.347$) and most southern population (32, $A_R = 1.329$). A_R is highest at two central coast populations (12—Red Head, A_R = 1.472 and 13—Barrington Tops, A_R = 1.474), though there is very little difference over the whole study region. Observed heterozygosity (H_{O} , Figure 7C) was found to be lowest in a southern population (27—Bungonia National Park, $H_O = 0.243$) and western population (08, $H_O = 0.236$). H_O is highest in a central population (11—Sea Acres National Park, $H_0 = 0.277$) and in a southern population (28—Wandandian Creek, $H_0 = 0.277$) though another two central populations, 12 and 13, are only slightly lower ($H_O = 0.276$). The highest number of private alleles (P_A , Figure 7D) occurs in the northern populations ((01, $P_A = 222$) and (02—Billinudgel, $P_A = 154$)). Population 07 (Dorrigo) has the lowest amount of private alleles P_A = 28; only two samples were sequenced for the site, so this may be a biased result. Excluding population 07 leaves the southern populations 29 (Mount Agony) and 30 (Deua National Park) with the lowest P_A . Inbreeding coefficient was highest for the two most northern populations (01, F_{IS} = 0.018 and 02, F_{IS} = 0.020) and lowest for the most southern population (32, F_{IS} = -0.34) and population 07 ($F_{IS} = -0.290$).

Environmental niche modelling (Figure 8) for LGM, MH, the current era and the future (2050) show a north to south movement of suitable environment for the species, with an overall decline over time. LGM shows the largest area of suitable environment (in red) centred around southeastern QLD and northeastern NSW. The MH prediction shows suitable areas reduced by about half and located more southernly than during the LGM. The current era maintains similar levels of suitability to the MH and a continued southward shift of suitable habit. The future predication shows the species losing more suitable habitat, with only small patches of highly suitable environment shifting south. At the northern extent, suitable patches have moved westwards.



Figure 8. Environmental niche modelling (ENM) for *Ficus coronata* during Last Glacial Maximum (LGM 22,000 ybp), Mid Holocene (MH 6000 ybp), Current era and Future (2050). Red shows the most favourable habitat and blue the least.

4. Discussion

The results from our genetic analysis indicate a rapid and recent southern range expansion for *F. coronata*. Very little genetic structure was detected over the complete NSW range (Figure 5), a recognised signal of expansion [60]. All analytical approaches (Figures 3–6; and Appendix A Figure 1) identified the extreme western (08) and southern (32) populations as more genetically distinct than other populations over similar distances. Our range expansion analysis (Figure 7A) defines most of the northern distribution of this species as the origin of expansion and the tip of the southern distribution as non-origin. Although the analysis provided no clear population of origin, a consequence of very low differentiation across the species distribution, it does support a north to south expansion. Private allele counts (PA, Figure 7D) also support a north to south expansion, with the highest PA in the north and lowest in the south. Environmental niche models (Figure 8 also support a southern shift of suitable habitat from the LGM to now, and the MAHF analysis (Figure 6 supports a recent expansion with a gradual loss of MAHF in a north to south and east to west direction.

The increased frequency of MAHF for populations 08 and 32 are a likely signature of allele surfing [61–63]. Allele surfing occurs when alleles 'surf' through populations on the 'wave' or leading edge of migration. This causes alleles that are uncommon across the species distribution (low minor allele frequency—MAF) to be represented at high frequencies in populations at the migration front (minor allele at highest frequency— MAHF). Declines in population densities towards the expansion front can intensify the effects of drift [62] and the chance for alleles to surf. Mutations that develop near the wave front have a higher probability of being caught by the wave [64] and all alleles that have surfed can likely be traced to an ancestor that lived on a previous wave front [65]. As a population expands and loses alleles through genetic drift, the remaining variants can be found at higher frequency in the new front than in any other population. Surfing alleles can become fixed in a newly occupied habitat at the species' range boundaries [66], leading to populations with allelic combinations found nowhere else in the range.

The dramatic difference between MAHF (Figure 6) for populations 07 and 08 may indicate that shorter distances between origin and leading-edge populations increase the

chance of surfing. The more distant southern population 32 has a much smaller increase in MAHF; however, this may be because it is not the current migration front, which is located 150 km to the south in Victoria. The Victorian population may also show a dramatic increase in MAHF.

The idea that a newly expanded population has unusual allelic frequencies appears counter intuitive and may lead to an inference of local adaptation [67], historical persistence within refugia, or even cryptic speciation. Peischl et al. [64] suggested that some studies on human adaption to local environments, e.g., [68,69], may be misinterpreted allele surfing events.

Environmental niche modelling (Figure 8), based on climatic data, supports a post-LGM southern expansion of habitat suitable to *F. coronata*, a southerly shift that will potentially continue into the future. The period with the greatest availability of suitable habitats appears to be the LGM (Figure 8) especially around northern NSW and southern QLD. The genetic signal of rapid expansion displayed by *F. coronata* is similar to that seen in rapidly expanding Sunda-derived taxa [4,19,70].

5. Conclusions

Our results, including the novel MAHF analysis, highlight *F. coronata* as an example of the little explored topic of allele surfing and another example of a rainforest species rapidly expanding distribution during the Holocene. This study has implications for conservation and management of this species, which is currently regarded as threatened in Victoria (Flora and Fauna Guarantee Act 1988—Threatened List November 2019). Although not sampled in this study (we have included collections from Brogo, 150 km north), it appears likely the *F. coronata* has recently expanded into eastern Victoria, possibly in the last century. The earliest herbarium collections from the region are by Wakefield in 1950 and only 46 collections have been made to date, all around Mallacoota. This study suggests the small Victorian population is more likely the product of on-going southern expansion than a relictual population in decline. The widespread practice of planting this species in bush regeneration may be introducing the species to new areas, a process that needs greater consideration in view of possible range shifts as a consequence of anthropogenic climate change.

Author Contributions: Conceptualization, B.C.W., S.R., J.-Y.S.Y., M.R.; methodology, B.C.W., S.R., J.-Y.S.Y., M.R.; analysis, B.C.W., S.R., J.-Y.S.Y.; investigation, B.C.W., S.R., J.-Y.S.Y., M.R.; data curation, B.C.W., J.-Y.S.Y.; writing—original draft preparation, B.C.W.; writing—review and editing, B.C.W., S.R., J.-Y.S.Y., M.R.; visualization, B.C.W., J.-Y.S.Y.; supervision, M.R.; project administration, B.C.W., J.-Y.S.Y.; funding acquisition, M.R. All authors have read and agreed to the published version of the manuscript.

Funding: Genomic sequencing was provided by Restore and Renew, an initiative of the Research Centre for Ecosystem Resilience at the Royal Botanic Garden Sydney to sample and examine the genetics of plant species within New South Wales.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Single row SNP genomic data: https://doi.org/10.6084/m9.figshare. 14633010.v1.

Acknowledgments: Collecting permit (SL 100569) and funding for sequencing were provided by Restore and Renew, an initiative of the Research Centre for Ecosystem Resilience at the Royal Botanic Garden Sydney to sample and examine the genetics of plant species within New South Wales. Thanks to Jason Bragg, Royal Botanic Gardens Sydney, for providing code to analyse SNP data and his invaluable advice. Brendan Lepschi, Kirsten Cowley and Emma Toms at Canberra Herbarium, Pina Milne and Angharad Johnson at Victorian Herbarium and Andrew Franks at Brisbane Herbarium for supplying images of herbarium vouchers with disputed identifications. Thanks also to Cécile Truchot for assistance with the *F*_{ST} analysis and Joel Cohen, Patricia Lu-Irving and Kit King for their support.

Conflicts of Interest: The authors declare no conflict of interest.



Appendix A. IBD and FST Pairwise Heatmaps

Figure 1. (A) Heatmap showing pairwise comparison of population geographic distance or genetic isolation by distance (IBD). Scale from white to blue; no genetic difference to high genetic difference. (B) Heatmap of pairwise comparison of the fixation index (F_{ST}). Scale from white to red; no difference or total gene flow with paired population to high differentiation or no gene flow with paired population.



Appendix B. PCA for Ficus coronata and Two Closely Related Species

Figure A2. Principal component analysis showing *Ficus coronata* (Red) 178 individuals, and the 2 sandpaper species that can co-occur, *Ficus aculeata* (Blue)-4 individuals and *Ficus fraseri* (Orange)-4 individuals. No evidence was found for hybridisation within the study specimens.

Table A1. Individuals of *Ficus aculeata* and *Ficus fraseri* used in this study.

Voucher	Species	Latitude	Longitude
D0274341	Ficus aculeata	-15.6463	136.7555
D0272678	Ficus aculeata	-17.7832	137.8992
D0232704	Ficus aculeata	-16.4274	129.6194
D0264264	Ficus aculeata	-14.3511	129.8836
NSW1025125	Ficus fraseri	-33.0026	151.7286
NSW1031234	Ficus fraseri	-32.0763	152.5320
NSW1031390	Ficus fraseri	-32.4393	152.5372
NSW1031387	Ficus fraseri	-32.0764	152.5312

References

- 1. Davis, M.B.; Shaw, R.G. Range shifts and adaptive responses to quaternary climate change. Science 2001, 292, 673–679. [CrossRef]
- Baird, A.H.; Sommer, B.; Madin, J.S. Pole-ward range expansion of Acropora spp. along the east coast of Australia. *Coral Reefs* 2012, 31, 1063. [CrossRef]
- 3. Normand, S.; Ricklefs, R.E.; Skov, F.; Bladt, J.; Tackenberg, O.; Svenning, J.C. Postglacial migration supplements climate in determining plant species ranges in Europe. *Proc. R. Soc. B Biol. Sci.* **2011**, *278*, 3644–3653. [CrossRef]
- 4. Yap, J.Y.S.; van der Merwe, M.; Ford, A.J.; Henry, R.J.; Rossetto, M. Biotic exchange leaves detectable genomic patterns in the Australian rain forest flora. *Biotropica* 2020, *52*, 627–635. [CrossRef]
- 5. Hewitt, G.M. The genetic legacy of the Quaternary ice ages. Nature 2000, 405, 907–913. [CrossRef] [PubMed]
- 6. Bryant, L.M.; Krosch, M.N. Lines in the land: A review of evidence for eastern Australia's major biogeographical barriers to closed forest taxa. *Biol. J. Linn. Soc.* 2016, *119*, 238–264. [CrossRef]

- 7. Hewitt, G.M. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* **1996**, 58, 247–276. [CrossRef]
- 8. Buckley, J.; Bridle, J.R.; Pomiankowski, A. Novel variation associated with species range expansion. *BMC Evol. Biol.* **2010**, *10*, 382. [CrossRef]
- 9. Mastrantonio, V.; Porretta, D.; Urbanelli, S.; Crasta, G.; Nascetti, G. Dynamics of mtDNA introgression during species range expansion: Insights from an experimental longitudinal study. *Sci. Rep.* **2016**, *6*, 1–8. [CrossRef]
- 10. Petit, R.J.; Aguinagalde, I.; De Beaulieu, J.L.; Bittkau, C.; Brewer, S.; Cheddadi, R.; Ennos, R.; Fineschi, S.; Grivet, D.; Lascoux, M.; et al. Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* **2003**, *300*, 1563–1565. [CrossRef]
- 11. Hardy, O.J. Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Mol. Ecol.* 2003, *12*, 1577–1588. [CrossRef]
- 12. Wright, S. Isolation by distance. Genetics 1943, 28, 114–138. [CrossRef]
- 13. Ishida, Y. Sewall wright and gustave malécot on isolation by distance. Philos. Sci. 2009, 76, 784–796. [CrossRef]
- 14. Wright, S. Isolation by distance under diverse systems of mating. Genetics 1946, 31, 39–59. [CrossRef]
- 15. Yap, J.Y.S.; Rossetto, M.; Costion, C.; Crayn, D.; Kooyman, R.M.; Richardson, J.; Henry, R. Filters of floristic exchange: How traits and climate shape the rain forest invasion of Sahul from Sunda. *J. Biogeogr.* **2018**, *45*, 838–847. [CrossRef]
- Binks, R.M.; Gibson, N.; Ottewell, K.M.; Macdonald, B.; Byrne, M. Predicting contemporary range-wide genomic variation using climatic, phylogeographic and morphological knowledge in an ancient, unglaciated landscape. *J. Biogeogr.* 2019, 46, 503–514. [CrossRef]
- McPherson, H.; van der Merwe, M.; Delaney, S.K.; Edwards, M.A.; Henry, R.J.; McIntosh, E.; Rymer, P.D.; Milner, M.L.; Siow, J.; Rossetto, M. Capturing chloroplast variation for molecular ecology studies: A simple next generation sequencing approach applied to a rainforest tree. *BMC Ecol.* 2013, *13*, 8. [CrossRef]
- Rossetto, M.; Mcpherson, H.; Siow, J.; Kooyman, R.; van der Merwe, M.; Wilson, P.D. Where did all the trees come from? A novel multispecies approach reveals the impacts of biogeographical history and functional diversity on rain forest assembly. *J. Biogeogr.* 2015, 42, 2172–2186. [CrossRef]
- Joyce, E.M.; Pannell, C.M.; Rossetto, M.; Yap, J.Y.S.; Thiele, K.R.; Wilson, P.D.; Crayn, D.M. Molecular phylogeography reveals two geographically and temporally separated floristic exchange tracks between Southeast Asia and Northern Australia. *J. Biogeogr.* 2021, 48, 1213–1227. [CrossRef]
- 20. Nei, M.; Maruyama, T.; Chakraborty, R. The Bottleneck Effect and Genetic Variability in Populations. *Evolution* **2012**, *29*, 1–10. [CrossRef]
- 21. Dodd, R.S.; Afzal-Rafii, Z. Selection and dispersal in a multispecies oak hybrid zone. *Evolution* **2004**, *58*, 261–269. [CrossRef] [PubMed]
- 22. Graham, C.H.; Moritz, C.; Williams, S.E. Habitat history improves prediction of biodiversity in rainforest fauna. *Proc. Natl. Acad. Sci. USA* 2006, *103*, 632–636. [CrossRef]
- Nevill, P.G.; Bossinger, G.; Ades, P.K. Phylogeography of the world's tallest angiosperm, Eucalyptus regnans: Evidence for multiple isolated Quaternary refugia. J. Biogeogr. 2010, 37, 179–192. [CrossRef]
- Das, S.; Baumgartner, J.B.; Esperon-Rodriguez, M.; Wilson, P.D.; Yap, J.Y.S.; Rossetto, M.; Beaumont, L.J. Identifying climate refugia for 30 Australian rainforest plant species, from the last glacial maximum to 2070. *Landsc. Ecol.* 2019, 34, 2883–2896. [CrossRef]
- 25. Fahey, M.; Rossetto, M.; Wilson, P.D.; Ho, S.Y.W. Habitat preference differentiates the Holocene range dynamics but not barrier effects on two sympatric, congeneric trees (*Tristaniopsis*, Myrtaceae). *Heredity* **2019**, *123*, 532–548. [CrossRef]
- van der Merwe, M.; Yap, J.Y.S.; Bragg, J.G.; Cristofolini, C.; Foster, C.S.P.; Ho, S.Y.W.; Rossetto, M. Assemblage accumulation curves: A framework for resolving species accumulation in biological communities using DNA sequences. *Methods Ecol. Evol.* 2019, 10, 971–981. [CrossRef]
- 27. De Almeida Vieira, F.; Novaes, R.M.L.; Fajardo, C.G.; Dos Santos, R.M.; De Souza Almeida, H.; De Carvalho, D.; Lovato, M.B. Holocene southward expansion in seasonally dry tropical forests in South America: Phylogeography of *Ficus bonijesulapensis* (Moraceae). *Bot. J. Linn. Soc.* **2015**, *177*, 189–201. [CrossRef]
- 28. Yu, H.; Nason, J.D. Nuclear and chloroplast DNA phylogeography of *Ficus hirta*: Obligate pollination mutualism and constraints on range expansion in response to climate change. *New Phytol.* **2013**, *197*, 276–289. [CrossRef]
- 29. Lin, R.C.; Yeung, C.K.L.; Li, S.H. Drastic post-LGM expansion and lack of historical genetic structure of a subtropical fig-pollinating wasp (*Ceratosolen* sp. 1) of *Ficus septica* in Taiwan. *Mol. Ecol.* **2008**, *17*, 5008–5022. [CrossRef]
- Chen, Y.; Compton, S.G.; Liu, M.; Chen, X.Y. Fig trees at the northern limit of their range: The distributions of cryptic pollinators indicate multiple glacial refugia. *Mol. Ecol.* 2012, *21*, 1687–1701. [CrossRef] [PubMed]
- 31. Berg, C.C.; Corner, E.J.H. *Flora Malesiana, Series I, Volume 17/Part 2: Moraceae (Ficus)*; National Herbarium of the Netherlands: Leiden, The Netherlands, 2005.
- 32. Rønsted, N.; Weiblen, G.D.; Clement, W.L.; Zerega, N.J.C.; Savolainen, V. Reconstructing the phylogeny of figs (*Ficus*, Moraceae) to reveal the history of the fig pollination mutualism. *Symbiosis* **2008**, *45*, 45–55.
- Wilde, B.C.; Rutherford, S.; Merwe, M.; Murray, M.L.; Rossetto, M. First example of hybridisation between two Australian figs (Moraceae). Aust. Syst. Bot. 2020, 33, 436–445. [CrossRef]

- Shanahan, M.; Harrison, R.; Yamuna, R.; Boen, W.; Thornton, I. Colonization of an island volcano, Long Island, Papua New Guinea, and an emergent island, Motmot, in its caldera lake. V. Colonization by figs (*Ficus* spp.), their dispersers and pollinators. *J. Biogeogr.* 2003, 28, 1365–1377. [CrossRef]
- 35. Parry-Jones, K.; Augee, M.L. Food selection by grey-headed flying foxes (pteropus poliocephalus) occupying a summer colony site near gosford, New South Wales. *Wildl. Res.* **1991**, *18*, 111–124. [CrossRef]
- Roberts, B.J.; Catterall, C.P.; Eby, P.; Kanowski, J. Latitudinal range shifts in Australian flying-foxes: A re-evaluation. *Austral Ecol.* 2012, 37, 12–22. [CrossRef]
- 37. Rossetto, M.; Bragg, J.; Kilian, A.; McPherson, H.; van der Merwe, M.; Wilson, P.D. Restore and Renew: A genomics-era framework for species provenance delimitation. *Restor. Ecol.* **2019**, *27*, 538–548. [CrossRef]
- 38. Wilde, B.C. Restore and Renew: A Specimen Collecting App for Android and IOS Version 1.19. Available online: https://apps.apple.com/au/app/restore-renew/id1481538802 (accessed on 17 November 2020).
- Frichot, E.; François, O. LEA: An R package for landscape and ecological association studies. *Methods Ecol. Evol.* 2015, 6, 925–929.
 [CrossRef]
- 40. Venables, B.; Hornik, K. Oz: Plot the Australian Coastline and States. R Package Version 1.0–21. 2016. Available online: https://cran.r-project.org/web/packages/oz/index.html (accessed on 17 November 2020).
- Gerritsen, H. Mapplots: Data Visualisation on Maps. R Package Version 1.5.1. 2018. Available online: https://CRAN.R-project. org/package=mapplots (accessed on 17 November 2020).
- 42. Jombart, T. Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics 2008, 24, 1403–1405. [CrossRef]
- 43. Jombart, T.; Ahmed, I. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **2011**, *27*, 3070–3071. [CrossRef]
- 44. Weir, B.S.; Hill, W.G. Estimating F statistics. Annu. Rev. Genet. 2002, 36, 721–750. [CrossRef]
- 45. Zheng, X.; Levine, D.; Shen, J.; Gogarten, S.M.; Laurie, C.; Weir, B.S. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **2012**, *28*, 3326–3328. [CrossRef]
- 46. Oksanen, A.J.; Blanchet, F.G.; Kindt, R.; Legen-, P.; Minchin, P.R.; Hara, R.B.O.; Simpson, G.L.; Solymos, P.; Stevens, M.H.H. Community Ecology Package. *Ecol. Packag.* **2013**, *2*, 263.
- 47. Keenan, K.; Mcginnity, P.; Cross, T.F.; Crozier, W.W.; Prodöhl, P.A. DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol. Evol.* **2013**, *4*, 782–788. [CrossRef]
- 48. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2014, 2014, 1–14. [CrossRef]
- 49. Peter, B.M.; Slatkin, M. Detecting range expansions from genetic data. Evolution 2013, 67, 3274–3289. [CrossRef]
- 50. Falistocco, E. Presence of triploid cytotypes in the common fig (*Ficus carica* L.). Genome 2009, 52, 919–925. [CrossRef]
- Linck, E.; Battey, C.J. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Mol. Ecol. Resour.* 2019, 19, 639–647. [CrossRef]
- Cobos, M.E.; Townsend Peterson, A.; Barve, N.; Osorio-Olvera, L. Kuenm: An R package for detailed development of ecological niche models using Maxent. *PeerJ* 2019, 7, 1–15. [CrossRef]
- 53. Phillips, S.J.; Dudík, M.; Schapire, R.E. A maximum entropy approach to species distribution modeling. In Proceedings of the Twenty-First International Conference on Machine Learning, Banff, AB, Canada, 4–8 July 2004; pp. 655–662. [CrossRef]
- Aiello-Lammens, M.E.; Boria, R.A.; Radosavljevic, A.; Vilela, B.; Anderson, R.P. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* 2015, 38, 541–545. [CrossRef]
- Hijmans, R.J. Raster: Geographic Data Analysis and Modeling. R Package Version 3.4–5. Available online: https://CRAN.R-project.org/package=raster (accessed on 15 November 2020).
- 56. Karger, D.N.; Conrad, O.; Böhner, J.; Kawohl, T.; Kreft, H.; Soria-Auza, R.W.; Zimmermann, N.E.; Linder, H.P.; Kessler, M. Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **2017**, *4*, 1–20. [CrossRef]
- 57. Fordham, D.A.; Saltré, F.; Haythorne, S.; Wigley, T.M.L.; Otto-Bliesner, B.L.; Chan, K.C.; Brook, B.W. PaleoView: A tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. *Ecography* 2017, 40, 1348–1358. [CrossRef]
- 58. Brown, J.; Hill, D.; Dolan, A.; Carnaval, A.; Haywood, A. PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Sci. Data* 2018, *5*, 180254. [CrossRef]
- Escobar, L.E.; Lira-Noriega, A.; Medina-Vogel, G.; Townsend Peterson, A. Potential for spread of the white-nose fungus (Pseudogymnoascus destructans) in the Americas: Use of Maxent and NicheA to assure strict model transference. *Geospat. Health* 2014, 9, 221–229. [CrossRef]
- Zhou, H.P.; Chen, J. Spatial genetic structure in an understorey dioecious fig species: The roles of seed rain, seed and pollenmediated gene flow, and local selection. *J. Ecol.* 2010, *98*, 1168–1177. [CrossRef]
- 61. Klopfstein, S.; Currat, M.; Excoffier, L. The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* **2006**, 23, 482–490. [CrossRef]
- 62. Hallatschek, O.; Hersen, P.; Ramanathan, S.; Nelson, D.R. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19926–19930. [CrossRef]
- 63. Excoffier, L.; Foll, M.; Petit, R.J. Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* 2009, 40, 481–501. [CrossRef]

- 64. Peischl, S.; Dupanloup, I.; Bosshard, L.; Excoffier, L. Genetic surfing in human populations: From genes to genomes. *Curr. Opin. Genet. Dev.* **2016**, *41*, 53–61. [CrossRef]
- Barton, N.H.; Etheridge, A.M.; Kelleher, J.; Véber, A. Genetic hitchhiking in spatially extended populations. *Theor. Popul. Biol.* 2013, 87, 75–89. [CrossRef]
- 66. Hofer, T.; Ray, N.; Wegmann, D.; Excoffier, L. Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Ann. Hum. Genet.* **2009**, *73*, 95–108. [CrossRef]
- Streicher, J.W.; McEntee, J.P.; Drzich, L.C.; Card, D.C.; Schield, D.R.; Smart, U.; Parkinson, C.L.; Jezkova, T.; Smith, E.N.; Castoe, T.A. Genetic surfing, not allopatric divergence, explains spatial sorting of mitochondrial haplotypes in venomous coralsnakes. *Evolution* 2016, 70, 1435–1449. [CrossRef] [PubMed]
- Evans, P.D.; Gilbert, S.L.; Mekel-Bobrov, N.; Vallender, E.J.; Anderson, J.R.; Vaez-Azizi, L.M.; Tishkoff, S.A.; Hudson, R.R.; Lahn, B.T. Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans. *Science* 2005, 309, 1717–1720. [CrossRef]
- Pickrell, J.K.; Coop, G.; Novembre, J.; Kudaravalli, S.; Li, J.Z.; Absher, D.; Srinivasan, B.S.; Barsh, G.S.; Myers, R.M.; Feldman, M.W.; et al. Signals of recent positive selection in a worldwide sample of human populations. *Genome Res.* 2009, 19, 826–837. [CrossRef]
- 70. Van der Merwe, M.; Mcpherson, H.; Siow, J.; Rossetto, M. Next-Gen phylogeography of rainforest trees: Exploring landscape-level cpDNA variation from whole-genome sequencing. *Mol. Ecol. Resour.* **2014**, *14*, 199–208. [CrossRef] [PubMed]