Microsatellite analysis of populations of the endangered tree Gomortega keule suggests pre-Columbian differentiation

Journal:	*New Zealand Journal of Botany*
Manuscript ID	NZJB-2016-0034.R3
Manuscript Type:	Research Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Munoz-Concha, Diego; Universidad Catolica del Maule, Escuela de Agronomia Davey, Michael; University of Nottingham, Plant and Crop Sciences Division Ribas, Gracia; University of Nottingham, Plant and Crop Sciences Division Mayes, Sean; University of Nottingham - Malaysia Campus, Crops for the Future Research Centre
Keywords:	Gomortega keule, population genetics, fragmentation, genetic structure, megafaunal extinction, relict populations

Microsatellite analysis of populations of the endangered tree *Gomortega keule* suggests pre-Columbian differentiation

[1] Diego Muñoz-Concha *
[2] Michael R Davey
[2] Gracia Ribas
[2,3] Sean Mayes

[1] Departamento de Ciencias Agrarias, Universidad Católica del Maule, Curicó, Chile
[2] Plant and Crop Sciences Division, University of Nottingham, Loughborough, LE12 5RD, UK
[3] Crops for the Future Research Centre, Jalan Broga, Semenyih, 43500 Selangor, Malaysia

* Corresponding author. Email: dmunoz@ucm.cl

Style Definition: Balloon Text: English (New Zealand)
Style Definition: Comment Text: English (New Zealand)

ABSTRACT

Temperate forests have been affected extensively by human activities, resulting in land cover changes and population fragmentation. However, these anthropogenic effects can be superimposed onto the natural history of species, making it difficult to determine which effect is more important for a particular species. *Gomortega keule* is an endangered tree that is found in one of the world's biodiversity hotspots in central-south Chile. Human activities have significantly impacted on the original habitat in this region in recent years and are commonly considered to be the main cause of the scarcity of this species. However, aspects of the natural history of this evergreen tree may also help explain its present-day genetic structure. Therefore, in this study, we undertook microsatellite genotyping of the two southernmost populations of *G. keule*, which are 7.5 km apart from each other and well isolated from other populations. We found that there was genetic differentiation between these populations, suggesting that they exhibited at least some differentiation before becoming isolated, most likely before human activities first impacted the region some two centuries ago. Molecular estimates of their divergence time supported a more ancient differentiation of the populations than would be explained by human activities alone. Therefore, it is possible that their isolation may have followed the extinction of megafaunal seed dispersers around 12,000 years before present in this region, as indicated by fruit characteristics, the absence of recruitment by seedlings and the existence of clonal trees.

Keywords: *Gomortega keule*, population genetics, fragmentation, genetic structure, megafaunal extinction, relict populations

INTRODUCTION

Temperate forests have been significantly affected by human activities, particularly agriculture and forestry, resulting in land use changes and the fragmentation of existing species. However, the life history traits and natural history of species can also be key elements that affect population isolation and thus the population genetics of trees (Bacles and Jump 2011).

The coastal forests of central-south Chile lie within one of the 25 biodiversity hotspots on Earth (Myers et al. 2000). This region served as a refugium during the Quaternary glacial period, with many species surviving as a result of the temperatures being moderated by the ocean (Villagrán & Armesto 2005). Indeed, the distributions of several woody species are currently restricted to latitudes 36º–40º, which has been attributed to this natural history (Villagrán & Armesto 2005). However, over the last two centuries, this region has also been extensively affected by fires, timber exploitation, land clearance and the introduction of exotic forestry plantations (Armesto et al. 2010), and consequently the population fragmentation and reproductive isolation that is observed in many native woody species in these forests is commonly attributed to anthropogenic disturbance.

Gomortega keule (Mol.) Baillon is an endangered species that is found in central-south Chile. It exhibits a distribution pattern of small, isolated populations over a narrow geographical area near to the coast (Fig. 1) and occurs in only one of the national protected areas in Chile. San Martín & Sánchez (2000) provided details of 22 locations where the species exists, most

of which have less than 100 individuals. Wood extraction activities and land clearance for agriculture occurred on a small scale before the arrival of Europeans in central Chile (Armesto et al. 2010), and indigenous people had some impact on the native forests, particularly in the main river basins. However, it is unlikely that *G. keule* populations were strongly affected by these activities because this evergreen tree mainly grows in mountainous areas, often in ravines and on pronounced slopes, where the impacts of indigenous human populations were restricted. Over the last two centuries, however, the areas in which *G. keule* occurs have been considerably affected by the logging of trees and especially the setting of fires for land clearance (Serra et al. 1986; Armesto et al. 2010), as well as the establishment of exotic timber plantations in more recent years, which has led to a reduction in native forest cover from 21% in AD 1975 to 7% in AD 2000 in the region where the species occurs (Armesto et al. 2010). Furthermore, individual *G. keule* trees have been found as stump shoots, indicating that the parental trees had been damaged or logged (Serra et al. 1986) and suggesting the loss of individuals in recent decades. In a genetic study of 11 populations using inter simple sequence repeat (ISSR) markers, García-Gonzáles et al. (2008) concluded that human activities caused the isolation of populations of *G. keule*. However, this assertion should be considered with caution since this is a long-lived species that occurs in a habitat that has experienced only relatively recent anthropogenic disturbance.

The ecology of *G. keule* has some interesting features. The flowers are pollinated by syrphid flies (Lander et al. 2009). Although barochory has been suggested (Le Quesne & Stark 2006), there have been no studies on seed dispersion. Interestingly, the fruit are among the largest in the Chilean flora and are edible. Janzen & Martin (1982) proposed the idea of 'evolutionary

anachronism' for several Central American species, such as *Gymnocladus dioicus* (Zaya & Howe 2009), and this could also apply to *Gomortega keule*. If this was the case, the fruit of *G. keule* would have been dispersed by animals that once existed in central Chile, such as *Paleolama*, *Antifer*, *Equus*, Milodontidae and Gomphoteriidae (Labarca & López 2006). Indeed, the edible pulp and very hard endocarp that encloses the seed (Fig. 2) appear to fit the megafaunal dispersion model (Guimarães et al. 2008). Today, the fruit of these trees rot on the forest floor in the absence of animals that are able to eat the complete fruit and transport the seed. In some regions, domestic animals (cows, pigs and goats) eat the flesh of the fruit but leave the seed. Although small seedlings can be found, they do not grow further, possibly due to foraging by domestic or wild (rabbits, hares) animals, or because of the ecological conditions that occur near settlements (humidity, light, litter alteration). Thus, with the extinction of the megafauna in South America around 12,000 years before present (yr BP) (Barnosky & Lindsey 2010), *G. keule* may have lost one class of its natural seed dispersers, causing it to become confined to areas that are close to the stands and populations where it remains today.

In this study, we investigated whether the isolation of populations of *G. keule* occurred much longer ago than could be explained by human activities in the last 200 years. To do this, we evaluated the genetic differentiation of two spatially close but distinct populations based on the notion that if these populations showed significant genetic differentiation, a process of genetic differentiation must have already been underway before any superimposed human activities in recent centuries. This could argue for a more complex situation than suggested by García-Gonzáles et al. (2008), with the megafaunal extinction around 12,000 yr BP potentially contributing to this process.

The two southernmost populations of *G. keule* were selected for this study, which are only 7.5 km apart and in the same valley, but nearly 70 km from the next nearest population. Importantly, there is a clear geographical separation between these populations today, with no *G. keule* trees between. While most populations of *G. keule* are generally of a similar size to the study populations (60–80 individuals over <1 hectare), we chose these particular populations because they occur in an area with a more limited degree of recent anthropogenic influence, being relatively distant from current human settlements. By contrast, populations near big cities (e.g. Concepción) or agricultural areas (northern area of distribution) are likely to have been affected for a longer period and more deeply by human activities.

MATERIALS AND METHODS

Sample populations

Plant materials were obtained from the Maule and Bío-Bío regions of Chile (Table 1) in AD 2007. In population A (Fig. 1; $n = 75$), six *G. keule* trees (trunk diameters >20 cm) were growing from old stumps, while the remainder were young trees or shoots (trunk diameters <10 cm) growing under the canopy in a plantation of *Pinus radiata*. These young individuals were derived from the stumps of old trees that had been cut before or at the time of plantation establishment and so would be expected to have the same genotypes as the original logged trees. No seedlings were seen during collection. By contrast, population B ($n = 63$) mainly comprised very old trees, some of which were the

oldest representatives of this species we found (trunk diameters ca. 2 m). As for population A, no seedlings were observed in association with this population. Samples were collected from all of the *G. keule* trees that could be found in each of these populations, with only one sample being collected from each group of shoots or trunks derived from an individual tree stump in population A. In addition, 14 individuals were sampled from a population in the northern area of the species' distribution (190 km away from populations A and B) as an outgroup (group C).

Sample collection and DNA extraction

Young leaves were collected from each tree or alternatively the cambium was sampled if leaves could not be found for a particular tree. The collected tissues were dried in sealable plastic jars containing self-indicating silica gel crystals, which were renewed as required, and stored at -20°C.

DNA extraction followed Lander et al. (2007), whereby 5 mg diethyldithiocarbamic sodium salt, 10 mg PVP-40000 and 5 µl β-mercaptoethanol were added per 1 ml of cetyltrimethylammonium bromide (CTAB) solution that had been pre-heated to 65ºC. This solution (500 µl) was then added to each sample for DNA extraction.

Microsatellite genotyping

Each sample was genotyped using seven microsatellite primer pairs. Five of these (Gk-1, Gk-30, Gk-31, Gk-35 and Gk-39) were developed by Lander et al. (2007), while the remaining two (CS2 and CS8) were developed by S. Mayes

following the method of Haddrill et al. (2002). The procedure to develop primers CS2 and CS8 involved the digestion and PCR amplification of genomic DNA, hybridisation to filters with artificial SSR repeat oligonucleotides, elution, and amplification. Rather than using cloning, a proportion of the enriched library was pyrosequenced using a 1/16 run (non-titanium reagents; 454 Life Sciences, Connecticut, USA).

To visualise the DNA, 1% agarose gels were prepared in 1× TAE buffer and ethidium bromide was added to a final concentration of 0.5 µg/ml before casting the gels. The PCR products with fluorescently labelled primers were then visualised on 2% agarose gels in 0.5× TBE buffer that were stained with ethidium bromide to a final concentration of 0.1 µg/ml.

The forward primers were directly labelled with D2, D3 or D4 fluorescent dye (WellRED; Sigma Aldrich, St. Louis, Missouri, USA). The optimum annealing temperature for PCR for each labelled primer was then determined in a Px2 Thermal Cycler using an annealing temperature gradient from 45°C to 60°C across the block. PCR for each sample was performed using a 20-µl reaction volume containing 2 µl template DNA, 2.5 µl buffer, 0.2 µl deoxynucleotide (dNTP), 0.05 µl bovine serum albumin (BSA; 10 µg/µl), 0.02 µl of each forward (labelled) and reverse primer (100 µM), 15 µl autoclaved H_2O, and 0.2 µl Taq polymerase. Master mixes for multiple reactions were made wherever possible. The reaction conditions were 94°C for 3 min, 30 cycles of 94°C for 1 min, 1 min at the specific annealing temperature for each primer, 72°C for 1 min, and final extension at 72°C for 20 min. The annealing temperatures were 60.9°C (Gk-1), 65.0°C (Gk-30, Gk-31 and Gk-35), 56.8°C (Gk-39) and 52.7°C°C (Gk-44, CS2 and CS8).

A total of 152 genotypes were analysed for seven loci using capillary electrophoresis (CEQ 8000 Genetic Analysis System; Beckman Coulter, California, USA). The PCR products were pooled according to their different fluorescent tags, with a maximum of three different coloured samples being analysed per well.

Microsatellite data analysis

The peaks that were produced by capillary electrophoresis were thoroughly examined for each locus and assigned to classes according to Amos et al. (2007) using a software tool developed by the same authors (FlexiBin).

For each sample, matrices were generated using allele size to score co-dominant markers at each locus. Polymorphic loci (Laurentin 2009) were used to generate unbiased estimates of heterozygosity (Nei 1978) and investigate population structure.

Population genetic analysis

Samples that were considered to have originated from the same individual were excluded from the population analysis. This included any trees that were both <10 m apart and had identical genotypes in the binary matrix In total, 19 distinct genotypes were identified among 42 samples, resulting in 23 samples (42 minus 19) being excluded. Of the remaining 129 samples that were included in the population analysis, 62 were from population A, 53 from B and 14 from C, all of which were assessed for all seven loci.

The number of alleles per population and per locus, and the allelic richness and private allelic richness per population were calculated with HP-Rare (Kalinowski 2005), using the rarefaction method to account for unequal sample sizes. The observed (H_O) and expected (H_E) heterozygosity for each locus and population were then computed with Arlequin 3.5 (Excoffier et al. 2005) using formulas from Nei (1987), and values of H_O and H_E for the combined populations were calculated with GDA using formulas from Nei (1978).

According to Laurentin (2009), co-dominant markers can be used to assess the diversity between groups with methodologies based on allele frequencies, such as Wright's F-statistics and Nei's parameters. Therefore, the inbreeding coefficient (F_{IS}) for each locus and population was computed using FSTAT (Goudet 1995) according to Weir & Cockerham (1984). The pairwise fixation index (Rho_{ST} or ρ_{ST}) was calculated for each population using Genepop 4.0.5.3 (Rousset 2008), following which a second version of the fixation index (F_{ST}) was calculated under the stepwise mutation model using Arlequin.

Exact tests for deviations from Hardy-Weinberg equilibrium (HWE) were performed with Arlequin and Genepop using the Markov chain method (forecasted chain length: 1,000,000; dememorisation steps: 100,000). Linkage disequilibrium was tested separately for each population using Arlequin and a Bonferroni correction was applied to the *P* values. The frequency of null alleles was calculated assuming simultaneous inbreeding for each locus and population with the software INEst (Chybicki & Burczyk 2009) using either the population or individual inbreeding model (1,000,000 iterations).

The similarity index of Nei & Li (1979) was used to evaluate the genetic relationship among individuals, and cluster analysis (UPGMA) and principal coordinate analysis (PCoA) with Euclidean distances were used to visualise these relationships, both of which are appropriate for this type of study (Laurentin 2009). PCoA makes no assumptions about the distribution of the data or population genetics and the Euclidean distance does not consider the common absence of an allele as a shared characteristic (Kloda et al. 2008).

Cluster analysis and PCoA were performed using the MVSP software (Multivariate Statistical Package version 3.13; Kovach Computing Services, Anglesey, UK). A scatter plot was generated with the two largest eigenvalues along the first two axes and a dendrogram was produced with UPGMA using Nei & Li's index (Nei & Li 1979) as a measure of similarity. All 152 samples, which included those samples that were considered to have originated from the same individual, were included in the PCoA and cluster analysis.

Population structure was assessed using a Bayesian cluster analysis with the software Structure 2.3.4 (Pritchard et al. 2000). Analyses were performed using the admixture model with independent allele frequencies (20 runs), with burn-in and simulation lengths of 300,000 and 1,000,000 iterations, respectively. The optimal value for K was estimated by calculating the statistic ΔK (Evanno et al. 2005). These procedures were initially followed for all samples and then for all samples except group C in an independent set of runs.

The divergence time of populations A and B was estimated with the software IMa2 (Hey & Nielsen 2007), which estimates population-genetic parameters by calculating posterior probabilities in a Bayesian sampling framework. A total of

100,000 genealogies were sampled to estimate the joint posterior probability distribution of the divergence time parameter. Generation times of 25, 100 and 175 years were used, as suggested by Jones et al. (2013) for long-lived trees, and a mutation rate of 0.000316 was used, as proposed by Tamaki et al. (2016) for trees. More than 100 runs were made to adjust the priors (values of q = 20, m = 500 and t = 2), allowing for burn-in periods of 127–340 million steps. For each generation time, 13 final estimates were made using different seed numbers, and the mean values of the divergence time (yr BP) and the lower and upper bounds of the estimated 95% highest posterior density intervals were calculated.

RESULTS

Microsatellite genotyping

The number of alleles that was found for each locus is presented in Table 2. All seven markers were polymorphic (major allele frequency <0.95) according to Laurentin (2009). The allelic richness and private allelic richness (after rarefaction) were lower in populations A (2.19 and 0.34, respectively) and B (2.13 and 0.24) than in C (2.63 and 1.61) (Table 2). Private alleles were found for all loci and their occurrence was associated with the presence of rare alleles, as many of them showed low frequencies. Although sample size can affect the number of alleles when a sampling approach is used, to the best of our knowledge populations A and B represented the entire populations.

316 The level of heterozygosity that was identified in *G. keule* (Table 3) is similar
317 to or lower than that reported for other endangered trees (Tamaki et al. 2008;
318 Finger et al. 2011; Shepherd & Perrie 2011). The largest difference between H_O
319 and H_E (>0.2) occurred at locus *Gk-31* for population A and locus *Gk-35* for
320 group C (Table 3), with exact tests for deviation from HWE being significant
321 (Bonferroni corrected $P < 0.002$) at each of these loci. The inbreeding
322 coefficient for each locus and population ranged from -0.233 to 0.874 (Table 4),
323 with the highest values (F_{IS}>0.8) being observed for locus *Gk-31* in population
324 A and locus *Gk-35* in group C.
325
326 The pairwise population fixation index ρ_{ST} was 0.045 for A–B, 0.882 for A–C
327 and 0.882 for B–C, while F_{ST} was 0.061, 0.407 and 0.382 for the same pairwise
328 comparisons, respectively. Significant linkage disequilibrium after Bonferroni
329 correction was present in one pairwise comparison for population A and one for
330 population B. The estimated frequency of null alleles was significantly different
331 from zero ($P < 0.001$) for locus *Gk-31* in population A using the population
332 inbreeding model (PIM) and the individual inbreeding model (IIM) with INEst
333 (Chybicki & Burczyk 2009). Similarly, the frequency was also significantly
334 different from zero for locus *Gk-35* in group C ($P < 0.001$ using PIM and $P <$
335 0.01 using IIM). Although the presence of null alleles could have some
336 influence on the accuracy of the statistical analyses, it was not considered likely
337 that it had a major influence on the overall results.
338
339 **Population structure**
340
341 There was some genetic differentiation among the populations studied.
342 Moderate genetic differentiation (F_{ST} = 0.051–0.150; Yeh 2000) was found

between populations A and B (F_{ST} = 0.061), while very great genetic differentiation (F_{ST} > 0.25; Yeh 2000) was found between populations A and C (F_{ST} = 0.41), and B and C (F_{ST} = 0.38).

The PCoA analysis (Fig. 3) showed that there was some differentiation between populations A and B, which is consistent with their geographical location: these populations are only 7.5 km apart (Fig. 1) and each occupies <1 hectare.

The dendrogram (Fig. 4) supported the PCoA analysis, with all samples except those from group C falling into a separate cluster, and samples from populations A and B generally being grouped into two partially overlapping clusters.

The Bayesian analysis with the software Structure (Fig. 5) appeared to confirm the groups that were identified by the PCoA and the dendrogram analysis. The graphical output of Structure for K = 2 indicated some degree of differentiation between populations A and B. When the analysis included group C, the highest ΔK corresponded to two clusters, placing populations A and B in the same cluster. However, the output for K = 3 also suggested some differentiation between populations A and B that is biologically meaningful.

The estimated divergence times for populations A and B and the 95% highest posterior density intervals were 890 (395–2908) yr BP for a generation time of 25 years, 3956 yr BP (1978–12,263) for 100 years and 6784 yr BP (3185–20,951) for 175 years.

DISCUSSION

Megafaunal dispersal and natural history of *G. keule*

It has been proposed that the seeds of *G. keule* require megafaunal dispersal (Muñoz-Concha & Davey 2011), and so it is possible that populations A and B have become increasingly isolated following the extinction of megafauna from Chile, leading to their current level of differentiation. This idea that population differentiation was already underway before human activities affected these populations is supported by a number of observations from the present study. The data suggest that populations A and B, which are geographically very close, have a moderate level of genetic differentiation but are clearly more closely related to each other than to group C, which could argue for the previous existence of a more continuous population that is now fragmented. Since populations A and B occur in the same valley, fragmentation as a result of physical barriers alone can be ruled out. The minimum divergence times calculated in the present study are greater than 200 yr BP (minimum 395 yr BP with 95% confidence for a 25-year generation time), which would argue for pre-existing differentiation before any superimposed effect of human activities.

The PCoA analysis showed a clearer separation of the populations than was found by García-Gonzáles et al. (2008), who used ISSR to analyse the genetic structure of 11 populations of *G. keule*, but only sampled some of the individuals from populations A and B (29 and 7, respectively). In addition to the larger number of individuals that were sampled in the present study, the use

of co-dominant microsatellite markers improved the resolution of a number of the analyses and also allowed heterozygosity-based indices to be calculated.

Differentiation between the examined populations of *G. keule* is demonstrated by the fixation index values, which lie within the range of those assessed with microsatellites for other endangered tree species, such as *Magnolia stellata* (Tamaki et al. 2008), *Dalbergia monticola* (Andrianoelina et al. 2009), *Medusagyne oppositifolia* (Finger et al. 2011) and *Pseudopanax ferox* (Shepherd & Perrie 2011). Furthermore, similar levels of genetic differentiation have also been observed among isolated populations of *Myrtus nivellei*, a relict species that is currently experiencing geographic range contraction (Migliore et al. 2013).

It has previously been argued that clonality plays a role in maintaining unaltered genotypes over a long period of time (even thousands of years) in other woody species with relict and isolated populations, as discussed by Backs et al. (2015), Bradbury et al. (2016) and Migliore et al. (2013). Although there have been no reported studies on the role of clonality as a mechanism to maintain heterozgosity and prevent inbreeding, and no current evidence for the age of clones or trees of *G. keule*, the ability to maintain a population through off-shoots could be a key factor that explains the current level of variation in these populations and so warrants further investigation.

The generation times of trees can be greatly extended through vegetative (clonal) reproduction. Indeed, clones of relict tree species have been dated back to 3000–11,000 yr BP in *Eucalyptus absitia* (Bradbury et al. 2016) and over 13,000 yr BP in *Quercus palmeri* (May et al. 2009). Therefore, since clones of

G. keule were observed in the present study and by Lander et al. (2010), it is possible that this species has an extended generation time, making a generation time of 25 years unlikely.

The estimated time for the divergence of populations A and B is over 800 yr BP, placing the population differentiation process well before two centuries ago. Therefore, the isolation of these two populations of *G. keule* was probably well progressed before extensive anthropogenic disturbances first impacted the area around 200 yr BP (Armesto et al. 2010), indicating that population differentiation may have started in glacial times, well before human-mediated fragmentation of the area. This situation was also discussed for the endemic tree *Nothofagus alessandrii* by Torres-Díaz et al. (2007), who concluded that there was probably no gene flow between populations of this species before they were impacted by relatively recent human activities.

A long regeneration cycle suggests that the study populations have been isolated for a long time. The rare occurrence or complete absence of successful sexual regeneration as evidenced by a lack of seedling establishment in the *G. keule* populations sampled has also been observed in other relict tree species such as *Myrtus nivellei* (Migliore et al. 2013), *Q. palmeri* (May et al. 2009), *Q. hinckleyi* (Backs et al. 2015) and *E. absita* (Bradbury et al. 2016). The failure of seedlings to establish and develop into new trees may indicate that the effective generation time is very long, but could also reflect the cessation of seedling recruitment due to the more recent introduction of grazing animals such as hares.

449 The fruit of *G. keule* matches the definition of megafaunal fruits that was
450 introduced by Guimarães et al. (2008), being 4–7 cm in diameter and yellow,
451 with a small number of large seeds (Fig. 2). The species also matches the
452 predictions of those authors, showing clumped spatial patterns, reduced
453 geographical ranges and high levels of among-population structuring. Other
454 traits that suggest that *G. keule* may have megafaunal seed dispersal
455 characteristics include the lack of modern dispersal agents, large and well
456 protected seed, edible fruit pulp and a strong vegetative propagation ability, as
457 previously discussed for *Gymnocladus dioicus* by Zaya & Howe (2009). In
458 addition, the present-day rare occurrence or complete absence of recruitment by
459 seedlings and the isolated distribution of individuals and populations further
460 support the idea of megafaunal dispersal of *G. keule*.
461
462 A number of questions still need to be answered to better understand the
463 population dynamics of *G. keule*. These include the juvenility and longevity of
464 individuals, the ecological conditions that are required for seedling survival,
465 and the minimum viable population size.
466
467 **Sampling strategy**
468
469 Some of the samples that were found to be genetically identical were clearly
470 shoots from an old tree that had since disappeared, as reported previously (San
471 Martín & Donoso 1996). However, some trees of *G. keule* that were growing
472 <2 m apart were found to be genetically different, allowing us to reject the
473 assumption that close physical proximity indicates that individuals are clonally
474 related. The intricate way in which the trees occupy space on the forest floor is
475 also well illustrated by the case where a root (10 cm diameter) that was sampled

476 from population B was found to be different from three individuals growing <7
477 m away but identical to a large tree growing 15 m away. Therefore, developing
478 an understanding of the spatial relationships within the population will be an
479 important aspect of any future conservation planning.
480
481 The findings of the present study support the idea that vegetative propagation
482 (i.e. regeneration by sprouting) is currently important in natural populations of
483 *G. keule*, as stated by San Martín & Donoso (1996). Some individuals that were
484 physically close to each other had the same genetic profile and so were
485 probably produced from a single tree. Although we did not undertake extensive
486 sampling of shoots that were very close together, the observation of shoots with
487 the same genetic profile may indicate that clonality has played an important role
488 in the recent survival of *G. keule* trees.
489
490 Clonal production is a relatively common event in trees with reduced
491 populations such as *Eucalyptus absita* (Bradbury et al. 2016) and *Myrtus*
492 *nivellei* (Migliore et al. 2013), with some remarkably old individuals being
493 found, as seen in *Quercus palmeri* (May et al. 2009) and *Q. hinckleyi* (Backs et
494 al. 2015). Since clonality effectively extends the reproductive cycle time by
495 maintaining the original genetic combinations, the estimated divergence times
496 for *G. keule* may be underestimates.
497
498 **Implications for conservation**
499
500 Although the occurrence of small and increasingly isolated populations of *G.*
501 *keule* seems to have preceded human history in the area, anthropogenic effects
502 remain very important for the future conservation status of this species. It is

apparent that there has been extensive destruction of the habitat, populations and trees of this species, as evidenced by the fact that most individuals are recently re-grown trees derived from old tree stumps with signs of destruction by humans (logging, fires).

Long-term isolation and genetic differentiation among populations may have important implications for conservation management strategies for *G. keule* and other species that occur in disjunct populations, particularly where the natural dispersal agents are no longer present. Conservation efforts should first be directed towards the protection of remnant populations and individuals, including the exclusion of cattle and hares from protected zones, and the careful management of light through canopy maintenance, particularly for individuals that are re-growing in current or former forestry plantations. Given the very limited levels of seedling survival at present, it will be difficult to maintain high genetic diversity through natural seed production. Therefore, seeds or vegetative material should be sampled from as many populations as possible to maximise the amount of genetic diversity that is captured for immediate *ex situ* conservation. However, the risks of outbreeding depression should also be evaluated.

The level of population differentiation that was observed in *G. keule* may offer a model for the conservation of other Chilean species for which the dispersal agents may be extinct, particularly *Pitavia punctata* and *Jubaea chilensis*, with a need to sample from all extreme populations of each of these. The southern populations of *G. keule* are very important from an agricultural and forestry perspective, as each isolated population contains valuable genetic resources for

the future domestication and genetic breeding of the species, and may contain genetic information that is not found among other extant populations.

ACKNOWLEDGEMENTS

We would like to thank Fernando Campos (Corporación Nacional Forestal de Chile) for field assistance, and Roberto Muñoz (Forestal Celco S.A.) and Forestal Tierra Chilena S.A. for providing permission for field visits. We also acknowledge suggestions for laboratory protocols from Dr. Tonya A. Lander, and Cristian Echeverría, Pablo San Martín and Joselyn San Juan for providing data essential for creating the map in Fig. 1. This research was part of Diego Muñoz-Concha's PhD programme, funded by Universidad Católica del Maule and Comisión Nacional de Investigación Científica y Tecnológica – Gobierno de Chile.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill AVS 2007. Automated binning of microsatellite alleles: problems and solutions. Molecular Ecology Notes 7: 10–14.

556 Andrianoelina O, Favreau B, Ramamonjisoa L, Bouvet JM 2009. Small effect
557 of fragmentation on the genetic diversity of *Dalbergia monticola*, an
558 endangered tree species of the eastern forest of Madagascar, detected
559 by chloroplast and nuclear microsatellites. Annals of Botany 104:
560 1231–1242.
561
562 Armesto JJ, Manuschevich D, Mora A, Smith-Ramírez C, Rozzi R, Abarzúa
563 AM, Marquet PA 2010. From the Holocene to the Anthropocene: a
564 historical framework for land cover change in southwestern South
565 America in the past 15000 years. Land Use Policy 27: 148–160.
566
567 Backs JR, Terry M, Klein M, Ashley MV 2015. Genetic analysis of a rare
568 isolated species: a tough little West Texas oak, *Quercus hinckleyi* C.H.
569 Mull. The Journal of the Torrey Botanical Society 142(4): 302–313.
570
571 Bacles CFE, Jump AS 2011. Taking a tree's perspective on forest
572 fragmentations genetics. Trends in Plant Science 16: 13–18.
573
574 Barnosky AD, Lindsey EL 2010. Timing of Quaternary megafaunal extinction
575 in South America in relation to human arrival and climate change.
576 Quaternary International 217: 10–29.
577
578 Bradbury D, Grayling PM, MacDonald B, Hankinson M, Byrne M 2016.
579 Clonality, interspecific hybridisation and inbreeding in a rare mallee
580 eucalypt, Eucalyptus absita (Myrtaceae), and implications for
581 conservation. Conservation Genetics 17: 193–205.
582

583 Chybicky IJ, Burczyk J 2009. Simultaneous estimation of null alleles and
584 inbreeding coefficients. Journal of Heredity 100: 106–113.
585
586 Evanno G, Regnaut S, Goudet J 2005. Detecting the number of clusters of
587 individuals using the software Structure: a simulation study.
588 Molecular Ecology 14: 2611–2620.
589
590 Excoffier L, Laval G, Schneider S 2005. Arlequin ver. 3.0: an integrated
591 software package for population genetics data analysis. Evolutionary
592 Bioinformatics Online 1: 47–50.
593
594 Finger A, Kettle CJ, Kaiser-Bunbury CN, Valentin T, Doudee D 2011. Back
595 from the brink: potential for genetic rescue in a critically endangered
596 tree. Moecularl Ecology 20: 3773–3784.
597
598 García-Gonzáles R, Carrasco B, Peñailillo P, Letelier L, Herrera R, Lavandero
599 B, Moya M, Caligari PDS 2008. Genetic variability and structure of
600 *Gomortega keule* (Molina) Baillon (Gomortegaceae) relict
601 populations: geographical and genetic fragmentation and its
602 implications for conservation. Botany 86: 1299–1310.
603
604 Goudet J 1995. Fstat version 1.2: a computer program to calculate F-statistics.
605 Journal of Heredity 86: 485–486.
606
607 Guimarães PR, Galetti M, Jordano P 2008. Seed dispersal anachronisms:
608 rethinking the fruits extinct megafauna ate. PLoS One 3: e1745 (doi:
609 10.1371/journal.pone.0001745).

Haddrill PR, Majerus MEN, Mayes S 2002. Isolation and characterization of highly polymorphic microsatellite loci in the 2-Spot Ladybird, *Adalia bipunctata*. Molecular Ecology Notes 2: 316–319.

Hey J, Nielsen R 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proceedings of the National Academy of Sciences 104: 2785–2790.

Janzen DH, Martin PS 1982. Neotropical anachronisms: the fruits the Gomphotheres ate. Science 215: 19–27.

Jones FA, Cerón-Souza I, Hardesty BD, Dick CW 2013. Genetic evidence of Quaternary demographic changes in four rain forest tree species sampled across the Isthmus of Panama. Journal of Biogeography 40: 720–731.

Kalinowski ST 2005. HP-Rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes 5: 187–189.

Kloda JM, Dean PDG, Maddren C, MacDonald DW, Mayes S 2008. Using principle component analysis to compare genetic diversity across polyploidy levels within plant complexes: an example from British Restharrows (*Ononis spinosa* and *Ononis repens*). Heredity 100: 253–260.

637 Labarca RO, López PG 2006. Los mamíferos finipleistocénicos de la formación
638 Quebrada Quereo (IV Región Chile): biogeografía, bioestratigrafía e
639 inferencias paleoambientales. Mastozoología Neotropical 13: 89–101.
640
641 Lander TA, Boshier DH, Harris SA 2007. Isolation and characterization of
642 eight polymorphic microsatellite loci for the endangered, endemic
643 Chilean tree *Gomortega keule* (Gomortegaceae) Molecular Ecology
644 Notes 7: 1332–1334.
645
646 Lander TA, Boshier DH, Harris SA 2010. Fragmented but not isolated:
647 contribution of single trees, small patches and long-distance pollen
648 flow to genetic connectivity for *Gomortega keule*, an endangered
649 Chilean tree. Biological Conservation 143: 2583–2590.
650
651 Lander TA, Harris SA, Boshier DH 2009. Flower and fruit production and
652 insect pollination of the endangered Chilean tree, *Gomortega keule* in
653 native forest, exotic pine plantation and agricultural environments.
654 Revista Chilena de Historia Natural 82: 403–412.
655
656 Laurentin H 2009. Data analysis for molecular characterization of plant genetic
657 resources. Genetic Resources and Crop Evolution 56: 277–292.
658
659 Le Quesne C, Stark D 2006. *Gomortega keule* (Mol.) Baillon. In: Donoso C ed.
660 Las especies arbóreas de los bosques templados de Chile y Argentina,
661 Autoecología. Valdivia, Marisa Cuneo Ediciones. Pp 277–284.
662

May MR, Provance MC, Sanders AC, Ellstrand NC, Ross-Ibarra J 2009. A Pleistocene clone of Palmer's Oak persisting in Southern California. PLoS ONE 4: e8346.

Migliore J, Baumel A, Juin M, Fady B, Roig A, Duong N, Médail F 2013. Surviving in mountain climate refugia: new insights from the genetic diversity and structure of the relict shrub *Myrtus nivellei* (Myrtaceae) in the Sahara Desert. PLoS ONE 8: e73795.

Muñoz-Concha D, Davey MR 2011. *Gomortega keule*, the neglected and endangered Chilean fruit tree. European Journal of Forest Research 130: 677–693.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Kent J 2000. Biodiversity hotspots for conservation priorities. Nature 403: 853–858.

Nei M 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.

Nei M 1987. Molecular evolutionary genetics. New York, Columbia University Press. 512 p.

Nei M, Li WH 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences 76: 5269–5273.

690 Pritchard JK, Stephens M, Donnelly P 2000. Inference of population structure
691 using multilocus genotype data. Genetics 155: 945–959.
692
693 Rousset F 2008. Genepop'007: a complete reimplementation of the Genepop
694 software for Windows and Linux. Molecular Ecology Resources 8:
695 103–106.
696
697 San Martín J, Donoso C 1996. Floristic structure and human impact on the
698 Maulino Forest of Chile. In: Armesto JJ, Villagrán C, Arroyo MK eds.
699 Ecología de los bosques nativos de Chile. Santiago, Editorial
700 Universitaria. Pp. 153–168.
701
702 San Martín J, Sánchez A 2000. The remnant communities of *Gomortega keule*
703 (Gomortegaceae, Magnoliopsida) in central Chile. Anales del Jardín
704 Botánico de Madrid 57: 317–326.
705
706 Serra M, Gajardo R, Cabello A 1986. Ficha técnica de especies amenazadas:
707 Gomortega keule (Mol.) Baillon "Keule" (Gomortegaceae), especie en
708 peligro. Santiago, Corporación Nacional Forestal. 18 p.
709
710 Shepherd LD, Perrie LR 2011. Microsatellite DNA analyses of a highly
711 disjunct New Zealand tree reveal strong differentiation and imply a
712 formerly more continuous distribution. Molecular Ecology 20: 1389–
713 1400.
714
715 Tamaki I, Setsuko S, Tomaru N 2008. Genetic variation and differentiation in
716 populations of a threatened tree, *Magnolia stellata*: factors influencing

the level of within-population genetic variation. Heredity 100: 415–423.

Tamaki I, Setsuko S, Tomaru N 2016. Genetic diversity and structure of remnant *Magnolia stellata* populations affected by anthropogenic pressures and a conservation strategy for maintaining their current genetic diversity. Conservation Genetics 17: 715-725.

Torres-Díaz C, Ruiz E, González F, Fuentes G, Cavieres L 2007. Genetic diversity in *Nothofagus alessandrii* (Fagaceae), an endangered endemic tree species of the coastal Maulino forest of central Chile. Annals of Botany 100: 75–82.

Villagrán C, Armesto JJ 2005. Fitogeografía histórica de la Cordillera de la Costa de Chile. In: Smith C, Armesto JJ, Valdovinos C eds. Biodiversidad y ecología de los bosques de la Cordillera de la Costa de Chile. Santiago, Editorial Universitaria. Pp 99–116.

Weir BS, Cockerham CC 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Yeh FC 2000. Population genetics. In: Young A, Boshier D, Boyle T (eds) *Forest conservation genetics*. CSIRO Publishing: Collingwood, pp 21–37.

Zaya DN, Howe HF 2009. The anomalous Kentucky coffeetree: megafaunal fruit sinking to extinction?. Oecologia 161: 221–226.

744
745
746
747

748

749 **Table 1** Geographical locations and altitudes of the *Gomortega keule*
750 populations sampled.

Population /group	Locality	Latitude South	Longitude West	Altitude (metres above sea level)	Number of individuals sampled
A	Predio Carmávida, Bosques Arauco S.A. (private forestry farm)	37º 41′	73º 18′	300	75
B	Predio Pino Huacho, Forestal Tierra Chilena Ltda. (private forestry farm)	37º 40′	73º 13′	450	63
C	Reserva Nacional Los Queules (national protected area)	35º 58′	72º 40′	500	11
	Predio Ralbún, Forestal Celco S.A. (private forestry farm)	36º 03′	72º 38′	540	2

751

752

753

754 **Table 2** Number of alleles, allelic richness and private allelic richness per locus and population/group in *Gomortega*
755 *keule*.

Locus	Allele size (bp)	Number of alleles				Allelic richness			Private allelic richness		
		A	B	C	Total	A	B	C	A	B	C
Gk-1	224–240	5	4	4	7	2.49	1.92	2.96	0.43	0.26	1.76
Gk-30	182–190	3	4	4	5	2.11	3.35	2.49	0.03	1.07	0.38
Gk-31	208–228	4	3	5	8	1.74	1.52	3.35	0.66	0.06	1.99
Gk-35	224–252	4	2	3	7	2.42	1.99	2.21	0.54	0.11	2.21
Gk-39	133–201	2	2	3	5	1.97	1.83	1.84	0.16	0.03	1.84
CS2	104–119	2	3	3	4	1.98	2.11	2.76	0.01	0.02	1.66
CS8	210–220	3	3	4	5	2.64	2.23	2.83	0.53	0.16	1.40
Mean		3.29	3.00	3.71	5.86	2.19	2.13	2.63	0.34	0.24	1.61

756

757

Table 3 Observed (H_O) and expected (H_E) heterozygosity for each locus and population/group of *Gomortega keule*.

Locus	A		B		C		All populations	
	H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E
Gk-1	0.468	0.451	0.283	0.308	0.786	0.643	0.426	0.500
Gk-30	0.339	0.391	0.735	0.730	0.500	0.585	0.519	0.616
Gk-31	0.032	0.255	0.094	0.159	0.643	0.712	0.124	0.298
Gk-35	0.468	0.512	0.566	0.503	0.071	0.500	0.465	0.614
Gk-39	0.469	0.475	0.358	0.343	0.286	0.262	0.403	0.535
CS2	0.500	0.498	0.491	0.523	0.571	0.648	0.504	0.569
CS8	0.597	0.588	0.453	0.431	0.429	0.558	0.519	0.534
Mean	0.410	0.453	0.426	0.428	0.469	0.558	0.423	0.524

764

765 **Table 4** Inbreeding coefficient (F_{IS}) for each locus and population/group of

766 *Gomortega keule*.

Locus	A	B	C	All
Gk-1	-0.036	0.083	-0.233	-0.033
Gk-30	0.134	-0.008	0.150	0.058
Gk-31	0.874	0.410	0.100	0.533
Gk-35	0.086	-0.126	0.862	0.083
Gk-39	0.014	-0.046	-0.095	-0.013
CS2	-0.005	0.063	0.122	0.039
CS8	-0.015	-0.050	0.239	0.002
All	0.095	0.006	0.164	0.069

767 F_{IS} values >0 indicate a deficit and <0 indicate an excess of heterozygotes.

768

769

Figure legends

Fig. 1 Geographic distribution of *Gomortega keule*. **A**, Map of Chile with an enlargement of the collection area (arrow). **B**, Populations of *G. keule* (green dots) and the sampled populations (yellow). The distance between populations A and B is 7.5 km.

Fig. 2 Fruit of *Gomortega keule*. **A**, Fruit showing the endocarp that encloses the seed. **B**, Representation of extinct megafauna (Gomphoteriidae) eating the fruit of *G. keule*.

Fig. 3 Principal coordinate analysis for three populations of *Gomortega keule* using microsatellites for seven loci. Populations A (n = 75) and B (n = 63) are the southernmost populations of this species, while group C (n = 14) comprises individuals found in the northern area of the species' natural distribution. The graph was generated using MVSP 3.13 with the first two axes accounting for 30.5% of the variance (Axis 1, 19.4%; Axis 2, 11.1%).

Fig. 4 Phenetic dendrogram for populations of *Gomortega keule* generated by cluster analysis with UPGMA using the index of Nei & Li (1979) and the software MVSP 3.13.

Fig. 5 Structure analysis of samples of *Gomortega keule* excluding group C showing two clusters (top); and of all samples showing two (middle) and three (bottom) clusters. A Bayesian analysis was conducted using the software Structure 2.3.4 with the admixture model.

797

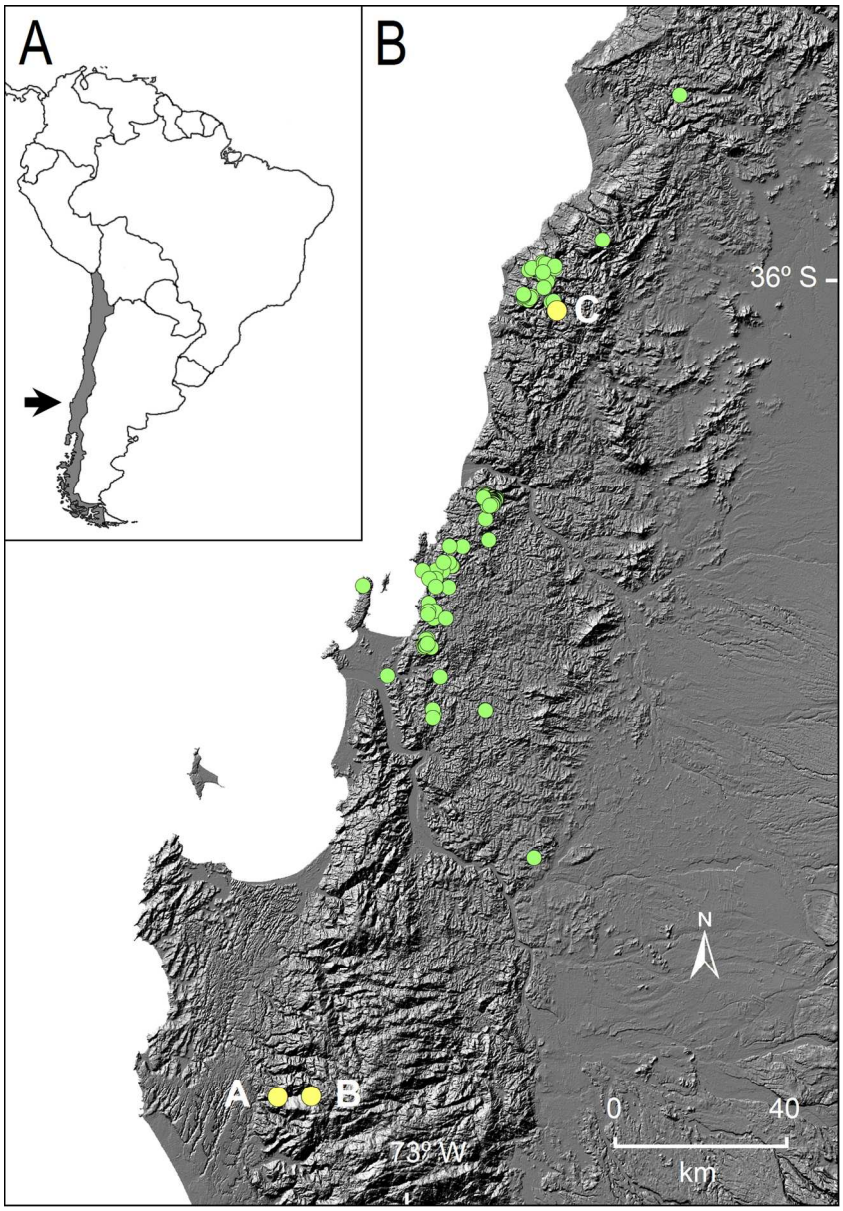

Fig. 1 Geographic distribution of Gomortega keule. A, Map of Chile with an enlargement of the collection area (arrow). B, Populations of G. keule (green dots) and the sampled populations (yellow). The distance between populations A and B is 7.5 km.

124x179mm (300 x 300 DPI)

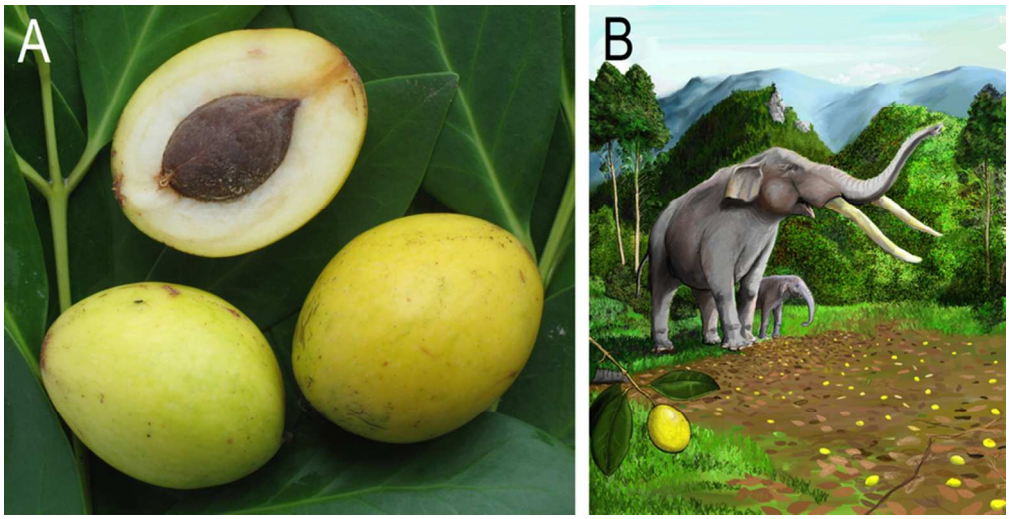

Fig. 2 Fruit of Gomortega keule. A, Fruit showing the endocarp that encloses the seed. B, Representation of extinct megafauna (Gomphoteriidae) eating the fruit of G. keule.

78x39mm (300 x 300 DPI)

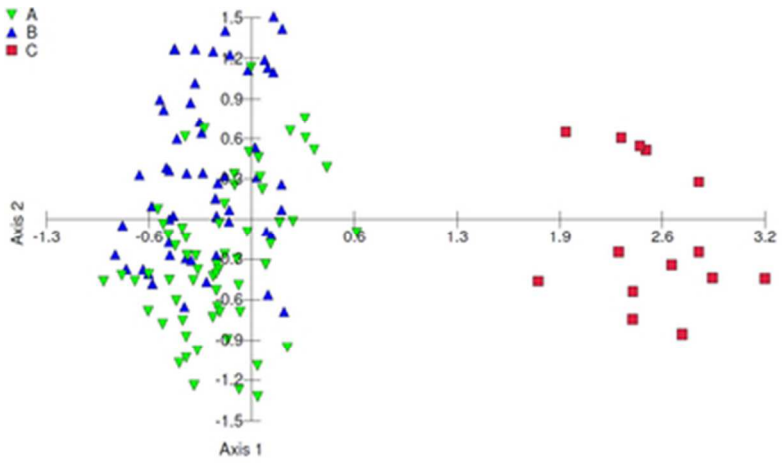

Fig. 3 Principal coordinate analysis for three populations of Gomortega keule using microsatellites for seven loci. Populations A (n = 75) and B (n = 63) are the southernmost populations of this species, while group C (n = 14) comprises individuals found in the northern area of the species' natural distribution. The graph was generated using MVSP 3.13 with the first two axes accounting for 30.5% of the variance (Axis 1, 19.4%; Axis 2, 11.1%).

33x19mm (300 x 300 DPI)

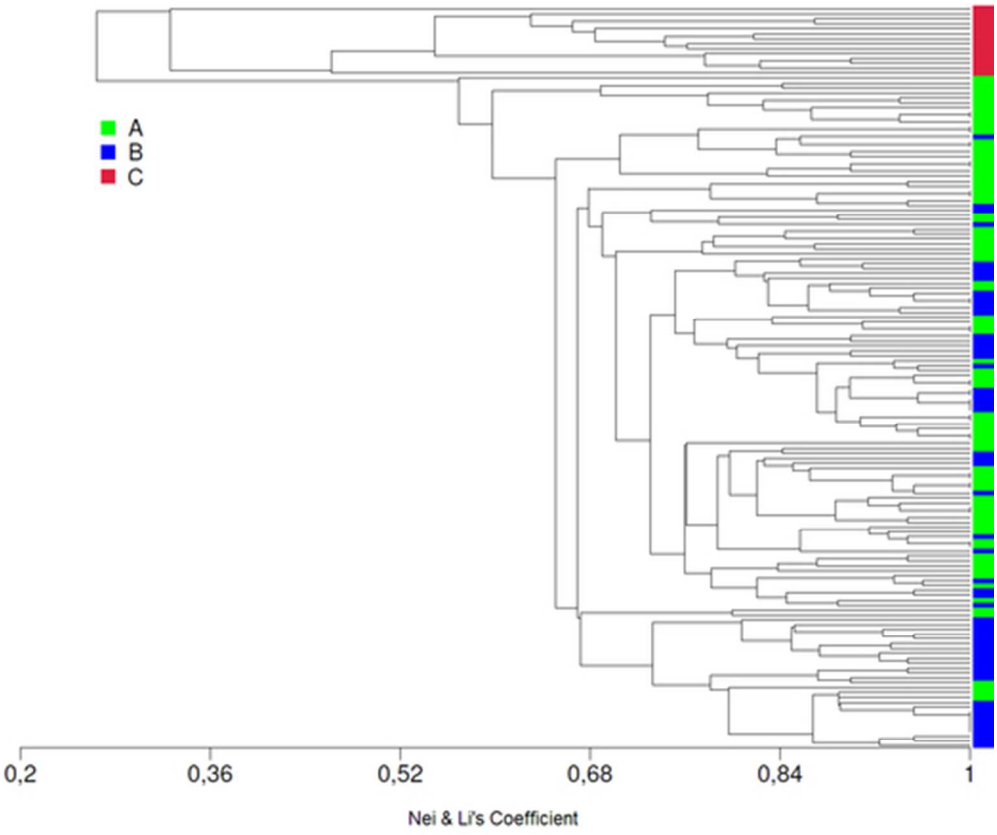

Fig. 4 Phenetic dendrogram for populations of Gomortega keule generated by cluster analysis with UPGMA using the index of Nei & Li (1979) and the software MVSP 3.13.

44x37mm (300 x 300 DPI)

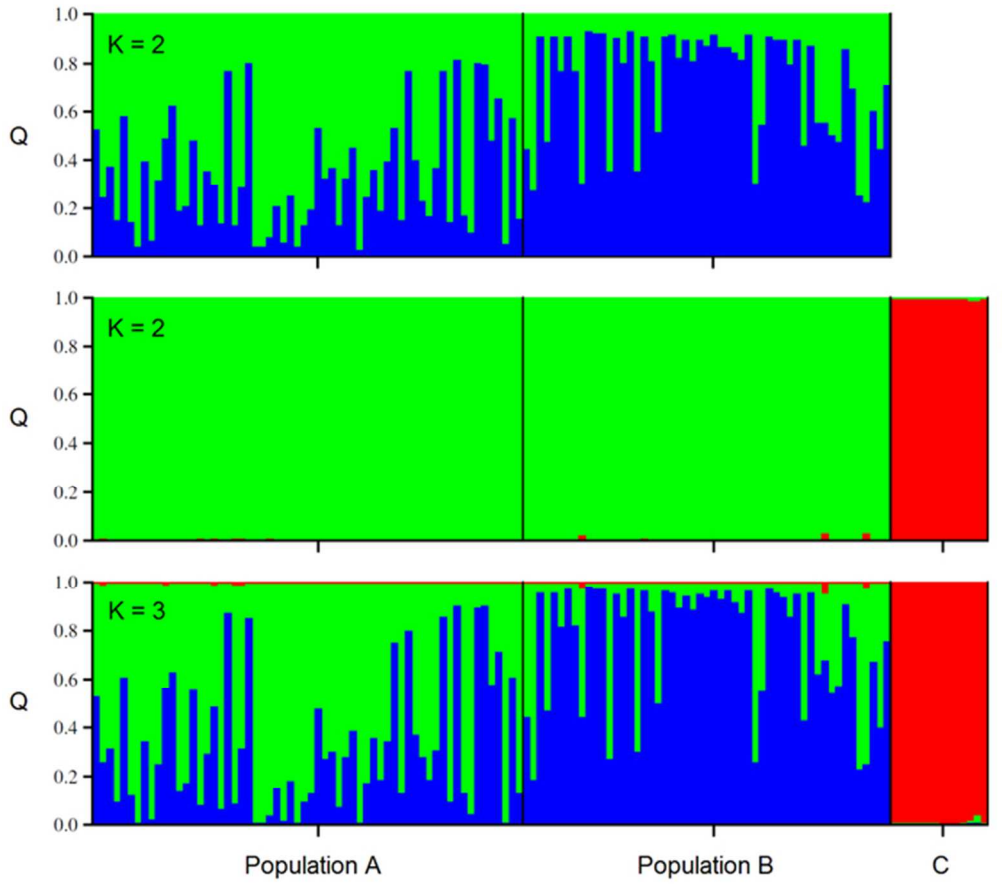

Fig. 5 Structure analysis of samples of Gomortega keule excluding group C showing two clusters (top); and of all samples showing two (middle) and three (bottom) clusters. A Bayesian analysis was conducted using the software Structure 2.3.4 with the admixture model.

58x52mm (300 x 300 DPI)