Downloaded from rspb.royalsocietypublishing.org on February 10, 2013

PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The complex history of the olive tree: from Late Quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant

G. Besnard, B. Khadari, M. Navascués, M. Fernández-Mazuecos, A. El Bakkali, N. Arrigo, D. Baali-Cherif, V. Brunini-Bronzini de Caraffa, S. Santoni, P. Vargas and V. Savolainen

Proc. R. Soc. B 2013 **280**, 20122833, published 6 February 2013

Supplementary data "Data Supplement" http://rspb.royalsocietypublishing.org/content/suppl/2013/02/05/rspb.2012.2833.DC1.html

References **This article cites 52 articles, 6 of which can be accessed free** http://rspb.royalsocietypublishing.org/content/280/1756/20122833.full.html#ref-list-1

Subject collections Articles on similar topics can be found in the following collections

environmental science (189 articles)
evolution (1390 articles)
genetics (51 articles)

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Proc. R. Soc. B* go to: **http://rspb.royalsocietypublishing.org/subscriptions**

Downloaded from rspb.royalsocietypublishing.org on February 10, 2013

PROCEEDINGS OF THE ROYAL SOCIETY B

rspb.royalsocietypublishing.org

Research

Cite this article: Besnard G, Khadari B, Navascués M, Fernández-Mazuecos M, El Bakkali A, Arrigo N, Baali-Cherif D, Brunini-Bronzini de Caraffa V, Santoni S, Vargas P, Savolainen V. 2013 The complex history of the olive tree: from Late Quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant. Proc R Soc B 280: 20122833.
http://dx.doi.org/10.1098/rspb.2012.2833

Received: 28 November 2012
Accepted: 10 January 2013

Subject Areas:
evolution, genetics, environmental science

Keywords:
coalescent-based molecular dating, crop genetic resources, long-term refugia, Mediterranean climate, plant domestication, species distribution modelling

Author for correspondence:
G. Besnard
e-mail: guillaume.besnard@univ-tlse3.fr

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2012.2833 or via http://rspb.royalsocietypublishing.org.

Royal Society Publishing
Informing the science of the future

The complex history of the olive tree: from Late Quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant

G. Besnard[1,2], B. Khadari[3], M. Navascués[4], M. Fernández-Mazuecos[5], A. El Bakkali[3], N. Arrigo[6], D. Baali-Cherif[7], V. Brunini-Bronzini de Caraffa[8], S. Santoni[3], P. Vargas[5] and V. Savolainen[2,9]

[1]CNRS-UPS-ENFA, EDB, UMR 5174, Bât. 4R1, 31062 Toulouse Cedex 9, France
[2]Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK
[3]INRA/CBNMED, UMR 1334, AGAP, 34060 Montpellier, France
[4]INRA, UMR1062 CBGP, 34988 Montferrier-sur-Lez, France
[5]Real Jardín Botánico de Madrid, CSIC (RJB-CSIC), Plaza de Murillo 2, Madrid 28014, Spain
[6]Department of Ecology and Evolutionary Biology, University of Arizona, PO Box 210088, Tucson, AZ 85721, USA
[7]Laboratoire de Recherche sur les Zones Arides, USTHB/INA El Harrach, Alger BP44, Algeria
[8]UMR CNRS 6134 SPE, LBBMV, Université de Corse, Corte 20250, France
[9]Royal Botanic Gardens, Kew, Richmond TW9 3DS, UK

The location and timing of domestication of the olive tree, a key crop in Early Mediterranean societies, remain hotly debated. Here, we unravel the history of wild olives (oleasters), and then infer the primary origins of the domesticated olive. Phylogeography and Bayesian molecular dating analyses based on plastid genome profiling of 1263 oleasters and 534 cultivated genotypes reveal three main lineages of pre-Quaternary origin. Regional hotspots of plastid diversity, species distribution modelling and macrofossils support the existence of three long-term refugia; namely the Near East (including Cyprus), the Aegean area and the Strait of Gibraltar. These ancestral wild gene pools have provided the essential foundations for cultivated olive breeding. Comparison of the geographical pattern of plastid diversity between wild and cultivated olives indicates the cradle of first domestication in the northern Levant followed by dispersals across the Mediterranean basin in parallel with the expansion of civilizations and human exchanges in this part of the world.

1. Introduction

Plant and animal domestication has been an essential step in the rise of human cultures [1], yet disentangling the history of domestication is challenging. Domestication events have generally started several millennia ago, and have been obscured by human-mediated spreads over long distances and recurrent crosses between cultivars and wild forms. Among the old crops of the Mediterranean basin, the olive tree (*Olea europaea* ssp. *europaea*) is the most iconic species owing to its ecological, economical and cultural importance [2–4]. This plant is considered one of the best biological indicators of the Mediterranean climate [3,5], and its cultivation has accompanied the emergence of Early Mediterranean civilizations [4]. The importance of the cultivated olive tree in people's lives has turned this species into a symbol in ancient sacred literature, and the origins of this crop are often subject to controversies [4]. Although early exploitation and use of wild olive trees (namely oleasters) has been documented since the Neolithic from the Near East to Spain [2,6], it is usually accepted that the domestication of the olive tree—characterized by vegetative propagation of the best cultivated genotypes that may have preceded orchard establishment—began in the Near East approximately 6000 years ago [2,4]. Several genetic

© 2013 The Author(s) Published by the Royal Society. All rights reserved.

Downloaded from rspb.royalsocietypublishing.org on February 10, 2013

studies have, however, supported multiple origins of cultivars across the Mediterranean area [7–12], but it remains unclear whether this reflects secondary diversification or multiple independent primary events. Also, a centre of domestication for the olive tree in the Near East was questioned because the eastern Mediterranean oleaster populations have low genetic diversity compared with those of the western area [10–13]. In addition, the discriminating power of the markers used until now has been poor, and a holistic approach is needed [4,14]. Current genetic patterns could also be interpreted in the light of palaeoclimatic, fossil and subfossil records.

Here, we investigated the impact of climate changes on oleaster populations during the Pleistocene, and particularly the genetic footprint in the plastid genome, and then searched for the primary origins of the domesticated olive. The most comprehensive dataset of plastid DNA polymorphisms was generated for this species in order to assess the diversification timing of oleaster populations and trace the origins of the cultivated maternal lineages. The fossil/subfossil records and modelling of suitable habitats under present and past temperature conditions helped the interpretation of genetic patterns and brought new insights on the olive persistence in space and time during the Late Quaternary and on the origins of early domesticated forms.

2. Material and methods

(a) Plant sampling and plastid genome profiling

In this study, we sampled 1846 trees (see the electronic supplementary material, tables S1 and S2), comprising 534 distinct cultivars, 1263 oleasters (from 108 locations) and 49 sub-Saharan trees belonging to the subspecies *cuspidata*, which were used as an outgroup (see below). DNA was extracted as previously described [10]. Owing to the highly fragmented and human-disturbed Mediterranean habitat, oleaster populations were usually collected from present orchards, but we cannot exclude that some trees/populations were feral (i.e. issued from cultivars).

Plastid genome polymorphism was characterized because maternally inherited haploid genomes are dispersed at shorter distance than nuclear genes, although ptDNA can be propagated at long distance with the human-mediated spread of cultivated clones. PtDNA is also more prone to stochastic events because the effective population size is half that of diploid genomes, allowing a more accurate detection of evolutionary events such as a long persistence of relict populations in refuge zones (e.g. hotspots of diversity) during previous glaciations [15]. Here, we used a profiling method to characterize variable microsatellites, indels and single nucleotide polymorphisms (SNPs) of the plastid genome [16]. Plastid DNA profiles (haplotypes) were defined by the combination of alleles from 61 loci (see the electronic supplementary material, table S3), and their distribution was represented on a map of the Mediterranean Basin.

(b) Haplotype network reconstruction and diversity analyses

Haplotype networks on the whole dataset or at lineage level were reconstructed with the reduced median method implemented in NETWORK [17]. For reconstructing networks at the lineage level, CAPS-*Xap*I was not considered owing to a high homoplasy on this locus [16]. We also computed three diversity parameters: the total diversity (defined as $H_T = 1 - \Sigma p_i^2$, with p_i the frequency of haplotype i), the haplotypic richness (A) and the rarity index (i.e. A weighted by haplotype frequencies; $R = \Sigma\, 1/p_i$) [18]. Spatial patterns of genetic diversity were inferred using a 'sliding window' [19].

(c) Molecular dating of the plastid genome diversification

Molecular dating of the Mediterranean plastid lineage diversification was performed using a coalescent-based Bayesian approach. The objective of this analysis is to estimate the time of most recent common ancestor (TMRCA) for the different plastid lineages of the Mediterranean olive tree. The first step is the estimation of the mutation rate (μ) for plastid microsatellites. This analysis was calibrated using the divergence time between *O. e.* ssp. *europaea* (lineages E1, E2 and E3) and *O. e.* ssp. *cuspidata* (African olive, lineage A), which was estimated to be 6.1 (8.3–4.0) Ma with a fossil-calibrated phylogeny [20]. Microsatellite data from lineages E1, E2, E3 and A were analysed with DIYABC [21], which implements an approximate Bayesian computation (ABC) method of inference in population genetics. Lineages E1, E2 and E3 were treated as populations to force the coalescence of individuals within the three separate lineages (as defined by SNPs and indels that were not used for this analysis) [10,16]. The demographic model (see the electronic supplementary material, figure S1) thus consisted of four populations (E1, E2, E3 and A). Populations E1, E2 and E3 diverge at the same time (T_E) from an ancestral population E (because relationships between the Mediterranean lineages were unresolved based on the complete ptDNA genome suggesting they started diverging from a common ancestor approximately at the same time) [16], and populations A and E diverge at time T_{AE}. Population size changes are allowed for each population; however, they are simultaneous (at time TD_E) and of the same magnitude (from $N_{ancestral}$ to N_E) for populations E1, E2 and E3, but independent for population A at time TD_A, from size $N_{ancestral}$ to N_A. Effective population size parameters ($N_{ancestral}$, N_E and N_A) are scaled to plastid effective population size. Prior distributions for the parameters are shown in electronic supplementary material, table S4. Prior distribution on parameter T_{AE} was chosen to incorporate the previous knowledge on the divergence between lineages E and A, and considering a generation time of 100 years. Note that this latter prior is used to scale calendar time (years) to coalescent time (generations) for the coalescent analyses, but it has little influence on final results because the estimated number of generations is finally converted into number of years for their interpretation. Microsatellite evolution was modelled with a strict stepwise mutation model.

Based on the μ estimate, a Bayesian method implemented in BATWING [22] was then used to infer the divergence time (expressed in generation number) from the most recent common ancestor of each Mediterranean ptDNA lineage (i.e. E1, E2 and E3). This second step consisted of the analysis of all plastid polymorphisms (microsatellites, indels and SNPs), but two loci that were incompatible with unique-event polymorphisms were excluded (i.e. loci CAPS-*Xap*I and 52) [16]. TMRCA was inferred for lineages E1, E2 and E3, and for sublineages with a value of posterior probability of being monophyletic higher than 0.9 (which includes sublineages defined by SNPs and indels, but also some sublineages with characteristic combination of microsatellite alleles). The demographic model consisted of an exponentially expanding population with uninformative prior distributions for effective population size at present (N) and growth rate (α); prior distribution for mutation rate was determined by the estimated distribution from the DIYABC analysis. Each of the three lineages was analysed separately. Three independent Markov chains of 5×10^9 steps (sampling every 5×10^5 steps the parameter values and every 5×10^6 steps the topology; proportion of parameter/tree

Downloaded from rspb.royalsocietypublishing.org on February 10, 2013

topology changes = 1/500) were run for each analysis, and estimates were drawn from them after removing the first 1000 samples from each chain as burn-in and combining all three chains. Unique-event polymorphisms (indels and SNPs) were used only to condition the topology of the genealogies (option 2 in BATWING). Ancestral states for SNPs and indels were considered unknown.

(d) Distribution modelling of present and past oleaster

Species distribution modelling (SDM) was performed to evaluate the potential oleaster distribution under present climatic conditions and to project it to the past. GPS coordinates of the populations analysed in our study were used as input data on the present oleaster distribution (see the electronic supplementary material, table S1). We discarded all localities where the three most common 'cultivated haplotypes' of lineage E1 (E1.1, E1.2 and E1.3; see below) were detected at a frequency superior to 70 per cent in the population to exclude putative feral populations for which the settlement may be due to human activities. Locations where only one individual was sampled (herbarium specimens) were also not considered. We used the maximum entropy algorithm, as implemented in MAXENT v. 3.3 [23], because it is appropriate for presence-only data and its performance has been shown to be consistently competitive in comparison with other methods [24]. Only temperature variables were used because olive survival is limited by frost [3,25]. In particular, the present wild olive distribution was reported to be mainly explained by the mean temperature of the coldest month [25]. Georeferenced data of 11 temperature variables under current conditions were retrieved from the WorldClim website (www.worldclim.org) [26], clipped to the extent of the Mediterranean region and used as predictors to calibrate the distribution model in MAXENT. When running the algorithm, occurrence data (88 GPS coordinates) were randomly split into training data (75%), used for model building, and test data (remaining 25%), used to evaluate model accuracy. Ten subsample replicates were performed, and fitness of the resulting model was assessed with the area under the receiver-operating characteristic curve [23]. A jackknife analysis was used to evaluate variable contributions to the model, and response curves were created to assess the extent to which the predictions were affected by variable values. To convert continuous suitability values to presence/absence, we chose the threshold obtained under the maximum training sensitivity plus specificity rule, which has been shown to produce accurate predictions [27]. The three distribution models under current conditions were projected to a period (*ca* 21 ka BP) of the last glacial maximum (LGM) using palaeoclimatic layers simulated from two general atmospheric circulation models: the Community Climate System Model [28] and the Model for Interdisciplinary Research on Climate [29]. We also projected the distribution models to the last interglacial (LIG; *ca* 120–140 ka BP) using the climatic model of Otto-Bliesner *et al.* [30]. It was assumed that the ecological requirements of *Olea* have remained stable throughout at least the last climatic cycle [31], which seems fairly reasonable given the short time scale.

3. Results

(a) Plastid genetic diversity in the Mediterranean olive

Polymorphisms were screened on the whole plastid genome (ptDNA) of 1797 trees and allowed the identification of 48 distinct profiles delineating three diverging lineages, namely E1, E2 and E3 (see the electronic supplementary material, figure S2 and table S3). A higher diversity was found in oleasters ($H_T = 0.88$, $A = 41.95$ haplotypes; figure 1) than in cultivars ($H_T = 0.35$, $A = 15$ haplotypes; electronic supplementary material, figure S3), and all cultivar haplotypes but one (L1.1) were observed in oleasters. Lineages E2 (detected in 71/108 locations) and E3 (33/108) exclusively occur in western and central Mediterranean areas, whereas lineage E1 (75/108) is currently distributed over the whole basin (figure 1).

The diversity indices support the occurrence of regional ptDNA diversity hotspots for the three lineages (see the electronic supplementary material, figure S4). In particular, a few haplotypes of lineages E2 and E3 are unique to oleaster populations from the Gibraltar Strait, whereas most E1 haplotypes are unique to the Aegean area and the Levant (figure 1). The three E1 haplotypes observed in the western area correspond to the main haplotypes of cultivars (figure 1). When haplotypes observed in the cultivated pool are excluded (see the electronic supplementary material, figure S5), the number of haplotypes shared between the Aegean area and the Levant is considerably reduced, whereas there remains only three rare, local E2 haplotypes in the central part of the Mediterranean basin.

(b) Diversification of wild olive lineages

Coalescent-based molecular dating of the plastid microsatellite variation allowed us to test for post- or pre-glacial oleaster diversification. The DIYABC estimate for the divergence time between the three Mediterranean lineages (1.66–6.17 Ma; electronic supplementary material, table S4) identified their most recent common ancestor between the Miocene–Pliocene boundary and the Mid-Pliocene, and could coincide with the establishment of the Mediterranean climatic rhythm [20,32–35]. Further investigation (BATWING dating; electronic supplementary material, tables S5, S6 and S7) then showed that the most recent common ancestor of each Mediterranean lineage dates back to the Middle or Upper Pleistocene: 139 100 BP for E1 (95% CI: 49 200–482 100), 284 300 BP for E2 (95% CI: 84 400–948 100) and 143 700 BP for E3 (95% CI: 37 100–542 700). These molecular dating analyses thus indicate that oleaster populations have diversified long before the LGM (26 500 to 19 000 BP [36]). Our molecular dating analyses also indicate that the E2 haplotypes unique to the central Mediterranean area (i.e. E2.7, E2.10) may have arisen after post-LGM colonization (see the electronic supplementary material, table S6).

(c) Oleaster distribution during the Late Quaternary

We used SDM and fossil/subfossil evidence to evaluate whether genetic diversity hotspots matched with putative refugia. The present oleaster distribution was relatively well predicted (figure 2*a*), but overflowed in some areas where oleaster populations have never been observed, such as Galicia/País Vasco (Spain) and Vendée (France). Several factors that are not accounted for by the SDM (including post-glacial dispersal limitations, human impact and interspecific competition) may explain the current absence of oleasters in these climatically suitable areas [31,37]. Distribution models fitted to current climatic conditions were then projected to the past. These SDM analyses indicate that suitable habitats for oleasters have persisted for a long period of time, and particularly during the LGM and the LIG (figure 2; electronic supplementary material, figure S6). Available macrofossils fit with the predictions for the LGM (figure 2*b*), in agreement with the persistence of oleasters in southern Europe and the Near East [3]. Putative Quaternary long-term refugia, defined

Downloaded from rspb.royalsocietypublishing.org on February 10, 2013

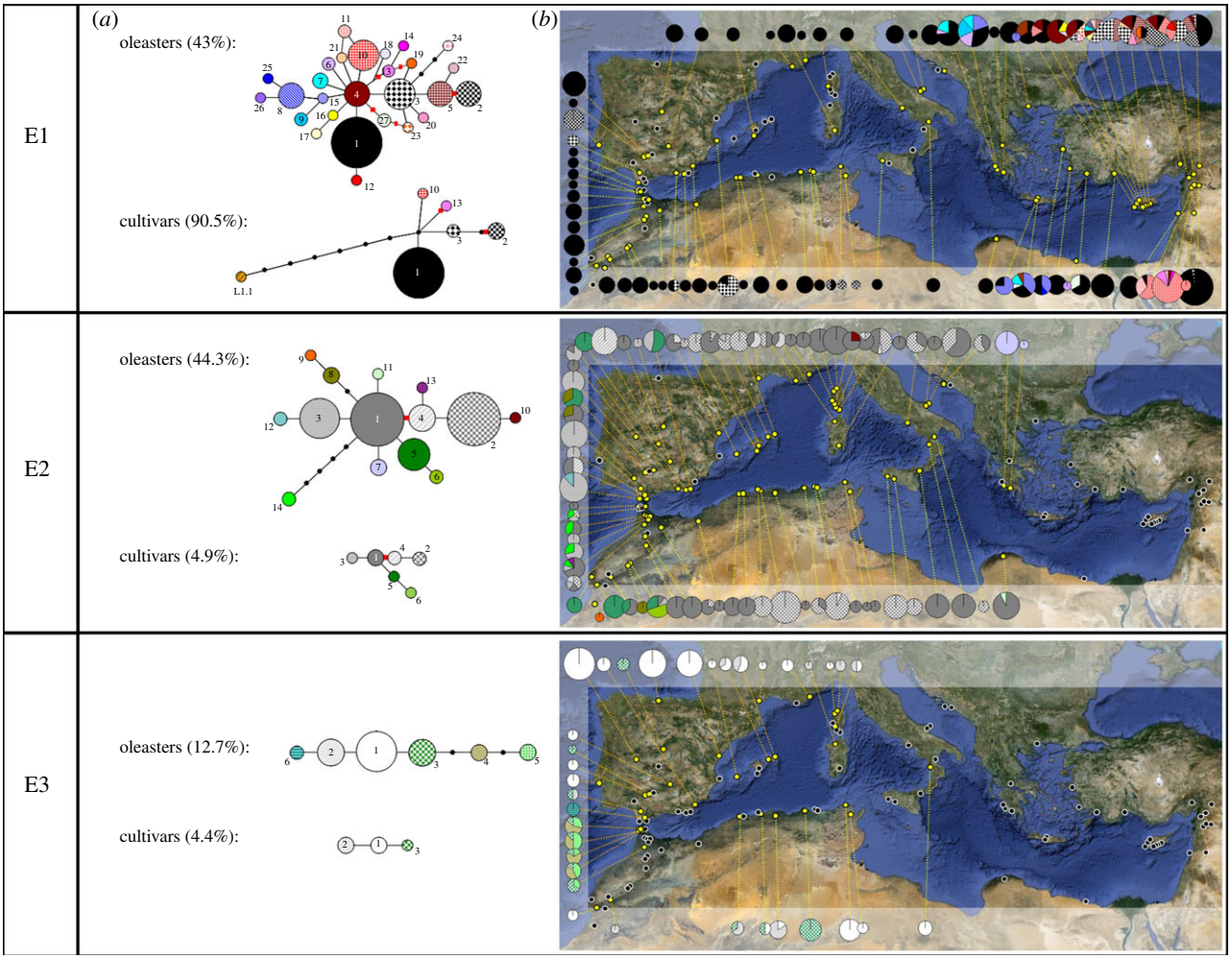


Figure 1. Diversity of the three Mediterranean olive plastid lineages [10]. (*a*) The reduced median haplotype networks [17] for each lineage and for both wild and cultivated gene pools. Each haplotype is numbered and represented by a symbol with a definite colour and/or motif. Haplotype frequencies are proportional to symbol diameter. The missing, intermediate nodes are indicated by small black points. Small red squares on the networks indicate mutational steps at the homoplasious CAPS-*Xap*I locus, which was excluded for the network analysis [16]. The frequency of each lineage in oleasters and cultivars is indicated in brackets. (*b*) The geographical distribution of haplotypes in oleaster populations. The size of pie charts is relative to the number of trees analysed per location. (Online version in colour.)

as areas with a continuous suitability for oleasters, were predicted in southern Europe (particularly in southern Iberia), as well as on a coastal band in the northern Maghreb, Cyrenaica and the Levant (figure 2*d*).

4. Discussion

The strong genetic differentiation between the western and eastern Mediterranean basin (E2/E3 versus E1) indicates olive lineages have been isolated in two distant areas, and this pattern is congruent with that of numerous other Mediterranean species of shrubs and trees [38–41]. The high diversity of lineage E1 in the eastern Mediterranean basin (see the electronic supplementary material, figure S4) suggests that this lineage originated in this area [10]. This result thus disagrees with previous statements about an impoverished ptDNA diversity in the Near East [10,11].

Patterns of Quaternary diversification of each ptDNA lineage could be explained by the succession of glaciations and interglacial periods. Marine isotopic stage 6 (MIS6; from 189 000 to 130 000 BP [42]) was a long glacial period in which oleaster populations probably persisted in southern refugia. This was followed by a long interglacial period (MIS5; from 130 000 to 74 000 BP) that is considered particularly favourable to the establishment of Mediterranean ecosystems [42]. Median age estimates suggest that diversification of present plastid lineages E1 and E3 might have started during the Eemian interglacial (MIS5e), while diversification within lineage E2 appears to have begun before MIS6.

The putative Quaternary long-term refugia correspond to a large part of the south-western Mediterranean basin (southern Iberia and Morocco), and coastal areas from the central Mediterranean basin to the Levant. Interestingly, the diversity hotspots detected for each plastid lineage coincided with some long-term persistence areas (see the electronic supplementary material, figures S4 and S5). Particularly, numerous unique haplotypes of lineages E2 and E3 are found on both sides of the Strait of Gibraltar, a region considered a major refugia for the Mediterranean flora (including olive) and fauna during the LGM [3,43–45]. Similarly, 27 haplotypes of lineage E1 are detected in the eastern Mediterranean area, most of them with a geographically restricted distribution. Two genetically differentiated areas can be identified: the Aegean region and the Levant (including Cyprus; figure 1). Natural seed-mediated gene flow between these areas appears to have been particularly limited, as already reported for *Laurus* and *Nigella* [40,46]. Indeed, a barrier to plant migration between these two

Downloaded from rspb.royalsocietypublishing.org on February 10, 2013

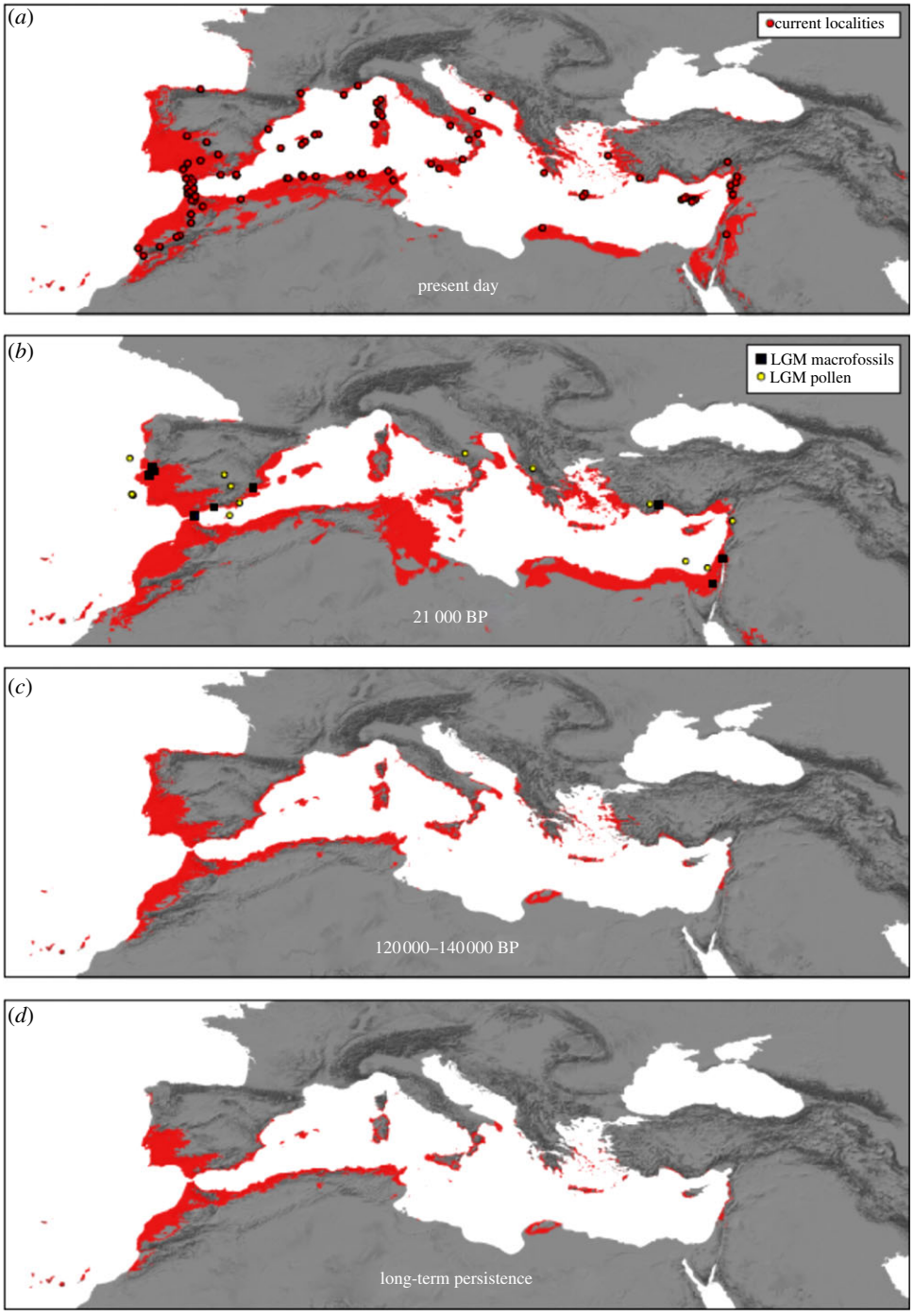


Figure 2. Distribution modelling of suitable habitats for oleasters based on temperature variables. (*a*) Distribution model fitted to current climatic conditions. (*b*) Projection to a period of the LGM (*ca* 21 000 BP). Results under the Community Climate System Model (CCSM) [28] and Model for Interdisciplinary Research on Climate (MIROC) [29] general circulation model simulations are averaged (for separate CCSM and MIROC results, see electronic supplementary material, figure S6). The locations of available olive fossil/subfossil remains (macrofossils and pollen) during the LGM are indicated (see the electronic supplementary material, table S8). Note that palynological data show low congruence with our model predictions, but this is probably due to efficient pollen dispersal by natural agents over long distances [3]. (*c*) Projection to the LIG (*ca* 120 000 – 140 000 BP) [30]. (*d*) Areas inferred as continuously suitable under LIG, LGM and current climatic conditions (putative 'long-term refugia'). In all cases, the average model of ten replicates is shown. (Online version in colour.)

phytogeographic regions, currently named the Rechinger's line [47], seems to have acted in southern Turkey since the Late Miocene [46].

The low ptDNA diversification in the central Mediterranean may be due to either rapid population turnover (maintaining low ptDNA diversity owing to stochasticity leading to genetic drift) or post-LGM colonization from major refugia located in the western and eastern areas. To date, there is no macrofossil for olive LGM persistence in the central Mediterranean basin, but just one piece of anthracological evidence for a presence during the Preboreal at Grotta dell'Uzzo (Sicily), suggesting an early colonization from putative proximal LGM refugia [3]. A similar distribution pattern of genetic diversity and LGM fossils was also reported for *Laurus nobilis* and *Myrtus communis*, two other circum-Mediterranean pre-Pliocene taxa [33,40,41]. The distribution of such species may have thus retracted and expanded in similar ways during the Quaternary glaciations and interglacial periods. In addition, phylogeographic structure within the olive fruitfly (*Bactrocera oleae*), which is considered the main

Downloaded from rspb.royalsocietypublishing.org on February 10, 2013

pest of the cultivated olive, also shows three Mediterranean lineages with a specific distribution on the Levant (including Cyprus), the Aegean area with continental Italy and the western Mediterranean basin (from southern Italy to Morocco and the Iberian Peninsula [48–50]). Pre-LGM diversification of the three lineages was also reported [49]. The geographical subdivisions in the olive fly can thus reflect three major refugia during the Quaternary glaciations that are putatively shared with its host [49].

The phylogeographic patterns observed in oleaster populations were finally used to trace back the origin of the main cultivated ptDNA haplotypes. Most present cultivars (90%) are characterized by E1 haplotypes (see the electronic supplementary material, figure S3). A human-mediated dispersal of E1.1, E1.2 and E1.3 (i.e. feral trees) from the eastern Mediterranean into western localities is strongly suggested by the absence of diversification within lineage E1 from Italy–Libya to Iberia–Morocco (figure 1). This result agrees with archaeological and historical evidence, which supports spread of cultivated olive trees from the Near East to the western Mediterranean [2,4]. Two populations from Israel and southwest Syria also exhibit exclusively major cultivated haplotypes and are therefore probably feral. The other studied wild populations from the southern Levant (southwest Syria, Jordan) have haplotypes mainly observed in oleasters (see the electronic supplementary material, figure S5), suggesting that this area was not the primary source of the domesticated olive and questioning initial expectations [2]. A co-occurrence of E1.1, E1.2 and E1.3 with intermediate or derived haplotypes unique to oleasters (i.e. E1.4, E1.5, E1.12 and E1.19; figure 1) was only observed in populations close to the Syrian–Turkish border. This observation suggests that the maternal haplotypes of olive cultivars have primarily originated in the northeastern Levant territory.

Evidence for Neolithic olive exploitation and occurrence of selection activity during the Bronze Age in the western Mediterranean (i.e. Iberian Peninsula and southern France) have also been indicated by fossil/subfossil data [6]. However, most of the present cultivars from this area are more closely related to eastern oleaster populations, according to both plastid (this study) and nuclear markers [8,11,13]. Accordingly, we conclude that the western Mediterranean was not a major primary centre of domestication of the olive tree.

The cradle of primary domestication of the olive tree is located in the northeastern Levant, where populations currently contain substantial genetic diversity, although not the highest in the Mediterranean basin (i.e. the Strait of Gibraltar [13,43]). This paradox can be explained by the fact that advanced civilizations emerged in the north Levant, such as the Pre-Pottery Neolithic B [51,52], and that they had enough genetic resources to succeed in domesticating a self-incompatible tree. The domestication of the olive tree appears to have been a long and continuous process that involved numerous genetic exchanges between the cultivated trees and wild gene pools, as already reported for other crops [53]. The first domesticated gene pool of olive was more likely to have spread with agriculture, first to the whole Levant and Cyprus [54] before being progressively disseminated to the western Mediterranean. Genetic evidence for multi-local origins of cultivars previously reported by several authors [6–12,55] may be explained by secondary domestication events involving crosses between newly introduced cultivars and local oleasters across the entire Mediterranean.

We thank K. Guschanski, C. Mammides, A. Minou, A. Moukhli, A. Ozgul, A. Papadopoulos, O. Pilia and C. Saslis-Lagoudakis for providing olive samples; H. Haouane, C. Tollon, M. Powell and S. del Vecchio for laboratory assistance; J. S. Carrión for advice on olive fossil/subfossil data; and P. A. Christin, H. Hipperson, S. Renner, C. Thébaud and an anonymous referee for valuable comments. This work was financially supported by the fellowship PIEF-GA-2008–220813, Agropolis Fondation FruitMed no. 0901–007, and the ERC, the Leverhulme Trust and the Royal Society. This work has been conducted at Silwood Park, the UMR AGAP and the laboratory EDB, part of the LABEX entitled TULIP (ANR-10-LABX-41). Herbarium specimens were provided by Kew Gardens (D. Goyder) and the British Museum of Natural History (M. Carine).

References

1. Diamond J. 2002 Evolution, consequences and future of plant and animal domestication. *Nature* **418**, 700–707. (doi:10.1038/nature01019)
2. Zohary D, Hopf M, Weiss E. 2012 *Domestication of plants in the Old World: The origin and spread of cultivated plants in Southwest Asia, Europe, and the Mediterranean basin*. Oxford, UK: Oxford University Press.
3. Carrión Y, Ntinou M, Badal E. 2010 *Olea europaea* L. in the North Mediterranean Basin during the Pleniglacial and the Early–Middle Holocene. *Q. Sci. Rev.* **29**, 952–968. (doi:10.1016/j.quascirev.2009.12.015)
4. Kaniewski D, Van Campo E, Boiy T, Terral JF, Khadari B, Besnard G, 2012 Primary domestication and early uses of the emblematic olive tree: palaeobotanical, historical and molecular evidences from the middle east. *Biol. Rev.* **87**, 885–899. (doi:10.1111/j.1469-185X.2012.00229.x)
5. Blondel JJ, Aronson J, Bodiou JY, Boeuf G. 2010 *The Mediterranean region: biological diversity in space and time*. Oxford, UK: Oxford University Press.
6. Terral JF *et al*. 2004 Historical biogeography of olive domestication (*Olea europaea* L.) as revealed by geometrical morphometry applied to biological and archaeological material. *J. Biogeogr.* **31**, 63–77. (doi:10.1046/j.0305-0270.2003.01019.x)
7. Belaj A, Trujillo I, De la Rosa R, Rallo L. 2001 Polymorphism and discriminating capacity of randomly amplified polymorphic markers in an olive germplasm bank. *J. Am. Soc. Horticulture Sci.* **126**, 64–71.
8. Besnard G, Baradat P, Breton C, Khadari B, Bervillé A. 2001 Olive domestication from structure of oleasters and cultivars using nuclear RAPDs and mitochondrial RFLPs. *Genet. Sel. Evol.* **33**, S251–S268.
9. Baldoni L, Tosti N, Ricciolini C, Belaj A, Arcioni S, Pannelli G, Germana MA, Mulas M, Porceddu A. 2006 Genetic structure of wild and cultivated olives in the central Mediterranean basin. *Ann. Bot.* **98**, 935–942. (doi:10.1093/aob/mcl178)
10. Besnard G, Rubio de Casas R, Vargas P. 2007 Plastid and nuclear DNA polymorphism reveals historical processes of isolation and reticulation in the olive tree complex (*Olea europaea* L.). *J. Biogeogr.* **34**, 736–752. (doi:10.1111/j.1365-2699.2006.01653.x)
11. Breton C, Terral JF, Pinatel C, Médail F, Bonhomme F, Bervillé A. 2009 The origins of the domestication of the olive tree. *Comptes Rendus Biol.* **332**, 1059–1064. (doi:10.1016/j.crvi.2009.08.001)
12. Belaj A *et al*. 2012 Developing a core collection of olive (*Olea europaea* L.) based on molecular markers (DArTs, SSRs, SNPs) and agronomic traits. *Tree Genet. Genomes* **8**, 365–378. (doi:10.1007/s11295-011-0447-6)
13. Lumaret R, Ouazzani N. 2001 Ancient wild olives in mediterranean forests. *Nature* **413**, 700. (doi:10.1038/35099680)
14. Zeder MA, Emshwiller E, Smith BD, Bradley DG. 2006 Documenting domestication: the intersection of genetics and archaeology. *Trends Genet.* **22**, 139–155. (doi:10.1016/j.tig.2006.01.007)
15. Petit RJ *et al*. 2002 Identification of refugia and post-glacial colonisation routes of European white

Downloaded from rspb.royalsocietypublishing.org on February 10, 2013

oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecol. Manage.* **156**, S49–S74. (doi:10.1016/S0378-1127(01)00634-X)

16. Besnard G, Hernández P, Khadari B, Dorado G, Savolainen V. 2011 Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biol.* **11**, 80. (doi:10.1186/1471-2229-11-80)
17. Bandelt HJ, Forster P, Röhl A. 1999 Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48. (doi:10.1093/oxfordjournals.molbev.a026036)
18. Ehrich D, Alsos IG, Brochmann C. 2008 Where did the northern peatland species survive the dry glacials: cloudberry (*Rubus chamaemorus*) as an example. *J. Biogeogr.* **35**, 801–814. (doi:10.1111/j.1365-2699.2007.01864.x)
19. Arrigo N, Felber F, Parisod C, Buerki S, Alvarez N, David J, Guadagnuolo R. 2010 Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat. *New Phytol.* **187**, 1170–1180. (doi:10.1111/j.1469-8137.2010.03328.x)
20. Besnard G, Rubio de Casas R, Christin PA, Vargas P. 2009 Phylogenetics of *Olea* (Oleaceae) based on plastid and nuclear ribosomal DNA sequences: tertiary climatic shifts and lineage differentiation times. *Ann. Bot.* **104**, 143–160. (doi:10.1093/aob/mcp105)
21. Cornuet JM, Ravigné V, Estoup A. 2008 Inferring population history with DiyABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24**, 2713–2719. (doi:10.1093/bioinformatics/btn514)
22. Wilson IJ, Weale ME, Balding DJ. 2003 Inferences from DNA data: population histories, evolutionary processes and forensic match probabilities. *J. R. Stat. Soc. A* **166**, 155–201. (doi:10.1111/1467-985X.00264)
23. Phillips SJ, Anderson RP, Schapire RE. 2006 Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **190**, 231–259. (doi:10.1016/j.ecolmodel.2005.03.026)
24. Elith J *et al*. 2006 Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151. (doi:10.1111/j.2006.0906-7590.04596.x)
25. Rubio R *et al*. 2002 On the historical presence of the wild olive [*Olea europaea* L. var. *sylvestris* (Miller) Lehr. (Oleaceae)] in the Eurosiberian region of the Iberian Peninsula. *Anal. J. Bot. Madrid* **59**, 342–344.
26. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005 Very high resolution interpolated climate surfaces for global land areas. *Int. J. Clim.* **25**, 1965–1978. (doi:10.1002/joc.1276)
27. Jiménez-Valverde A, Lobo JM. 2007 Threshold criteria for conversion of probability of species presence to either-or presence-absence. *Acta Oecol.* **31**, 361–369. (doi:10.1016/j.actao.2007.02.001)
28. Collins WD *et al*. 2006 The Community Climate System Model Version 3 (CCSM3). *J. Clim.* **19**, 2122–2143. (doi:10.1175/JCLI3761.1)
29. Hasumi H, Emori S. 2004 *K-1 coupled model* (*MIROC*) *description*. K-1 Technical Report 1. Tokyo, Japan: Center for Climate System Research, University of Tokyo.
30. Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A. 2006 Simulating arctic climate warmth and icefield retreat in the last interglaciation. *Science* **311**, 1751–1753. (doi:10.1126/science.1120808)
31. Nogués-Bravo D. 2009 Predicting the past distribution of species climatic niches. *Glob. Ecol. Biogeogr.* **18**, 521–531. (doi:10.1111/j.1466-8238.2009.00476.x)
32. Suc JP. 1984 Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature* **307**, 429–432. (doi:10.1038/307429a0)
33. Palamarev E. 1989 Paleobotanical evidences of the tertiary history and origin of the mediterranean sclerophyll dendroflora. *Plant Syst. Evol.* **162**, 93–107. (doi:10.1007/BF00936912)
34. Terral JF, Badal E, Heinz C, Roiron P, Thiebault S, Figueiral I. 2004 A hydraulic conductivity model points to post-neogene survival of the mediterranean olive. *Ecology* **85**, 3158–3165. (doi:10.1890/03-3081)
35. Jiménez-Moreno G, Fauquette S, Suc JP. 2010 Miocene to pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. *Rev. Palaeobot. Palynol.* **162**, 403–415. (doi:10.1016/j.revpalbo.2009.08.001)
36. Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW, McCabe AM. 2009 The last glacial maximum. *Science* **325**, 710–714. (doi:10.1126/science.1172873)
37. Calleja JA, Benito Garzón M, Sáinz Ollero H. 2009 A quaternary perspective on the conservation prospects of the tertiary relict tree *Prunus lusitanica* L. *J. Biogeogr.* **36**, 487–498. (doi:10.1111/j.1365-2699.2008.01976.x)
38. Arroyo-García R *et al*. 2006 Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp. *sativa*) based on chloroplast DNA polymorphisms. *Mol. Ecol.* **15**, 3707–3714. (doi:10.1111/j.1365-294X.2006.03049.x)
39. Désamoré A, Laenen B, Devos N, Popp M, González-Mancebo JM, Carine MA, Vanderpoorten A. 2011 Out of Africa: north-westwards Pleistocene expansions of the heather *Erica arborea*. *J. Biogeogr.* **38**, 164–176. (doi:10.1111/j.1365-2699.2010.02387.x)
40. Rodríguez-Sánchez F, Guzmán B, Valido A, Vargas P, Arroyo J. 2009 Late Neogene history of the laurel tree (*Laurus* L., Lauraceae) based on phylogeographical analyses of Mediterranean and Macaronesian populations. *J. Biogeogr.* **36**, 1270–1281. (doi:10.1111/j.1365-2699.2009.02091.x)
41. Migliore J, Baumel A, Juin A, Médail F. 2012 From Mediterranean shores to central Saharan mountains: key phylogeographical insights from the genus *Myrtus*. *J. Biogeogr.* **39**, 942–956. (doi:10.1111/j.1365-2699.2011.02646.x)
42. Martinson DG, Pisias NG, Hays JD, Imbrie J, Moore TC, Shackleton NJ. 1987 Age dating and the orbital theory of the ice ages: development of a high-resolution 0 to 300 000-year chronostratigraphy. *Q. Res.* **27**, 1–29. (doi:10.1016/0033-5894(87)90046-9)
43. Rubio de Casas R, Besnard G, Schönswetter P, Balaguer L, Vargas P. 2006 Extensive gene flow blurs phylogeographic but not phylogenetic signal in *Olea europaea* L. *Theor. Appl. Genet.* **113**, 575–583. (doi:10.1007/s00122-006-0306-2)
44. Rodríguez-Sánchez F, Pérez-Barrales R, Ojeda F, Vargas P, Arroyo J. 2008 The strait of gibraltar as a melting pot for plant biodiversity. *Q. Sci. Rev.* **27**, 2100–2117. (doi:10.1016/j.quascirev.2008.08.006)
45. Médail F, Diadema K. 2009 Glacial refugia influence plant diversity patterns in the mediterranean basin. *J. Biogeogr.* **36**, 1333–1345. (doi:10.1111/j.1365-2699.2008.02051.x)
46. Bittkau C, Comes HP. 2005 Evolutionary processes in a continental island system: molecular phylogeography of the Aegean *Nigella arvensis* alliance (Ranunculaceae) inferred from chloroplast DNA. *Mol. Ecol.* **14**, 4065–4083. (doi:10.1111/j.1365-294X.2005.02725.x)
47. Strid A. 1996 Phytogeographia Aegaea and the Flora Hellenica database. *Ann. Nat. Mus. Wien* **98B**, S279–S289.
48. Zygouridis NE, Augustinos AA, Zalom FG, Mathiopoulos KD. 2009 Analysis of olive fly invasion in California based on microsatellite markers. *Heredity* **102**, 402–412. (doi:10.1038/hdy.2008.125)
49. Nardi F, Carapelli A, Boore JL, Roderick GK, Dallai R, Frati F. 2010 Domestication of olive fly through a multi-regional host shift to cultivated olives: comparative dating using complete mitochondrial genomes. *Mol. Phylogenet. Evol.* **57**, 678–686. (doi:10.1016/j.ympev.2010.08.008)
50. van Asch B, Pereira-Castro I, Rei F, Teixeira da Costa L. 2012 Mitochondrial haplotypes reveal olive fly (*Bactrocera oleae*) population substructure in the Mediterranean. *Genetica* **140**, 181–187. (doi:10.1007/s10709-012-9669-2)
51. Edwards PC, Meadows J, Sayej G, Westaway M. 2004 From the PPNA to the PPNB: new views from the southern Levant after excavations at Zahrat adh-Dhra' 2 in Jordan. *Paléorient* **30**, 21–60. (doi:10.3406/paleo.2004.1010)
52. Zeder MA. 2011 The origins of agriculture in the Near East. *Curr. Anthropol.* **52**, S221–S235. (doi:10.1086/659307)
53. Willcox G, Nesbitt M, Bittmann F. 2012 From collecting to cultivation: transitions to a production economy in the Near East. *Vegetation Hist. Archaeobot.* **21**, 81–83. (doi:10.1007/s00334-012-0348-0)
54. Colledge S, Conolly J, Shennan S. 2004 Archaeobotanical evidence for the spread of farming in the eastern Mediterranean. *Curr. Anthropol.* **45**, S35–S58. (doi:10.1086/422086)
55. Besnard G, Anthelme F, Baali-Cherif D. 2012 The Laperrine's olive tree (Oleaceae): a wild genetic resource of the cultivated olive and a model-species for studying the biogeography of the Saharan Mountains. *Acta Bot. Gallica* **159**, 319–328. (doi:10.1080/12538078.2012.724281)