

## The Mediterranean: the cradle of *Anthoxanthum* (Poaceae) diploid diversity

Zuzana Chumová<sup>1,2,\*</sup>, Eliška Závěská<sup>1,3,†</sup>, Terezie Mandáková<sup>4</sup>, Karol Krak<sup>2,5</sup> and Pavel Trávníček<sup>2</sup>

<sup>1</sup>Department of Botany, Faculty of Science, Charles University, Benátská 2, CZ-128 01 Prague, Czech Republic, <sup>2</sup>Institute of Botany, The Czech Academy of Sciences, Zámek 1, CZ-252 43 Průhonice, Czech Republic, <sup>3</sup>Institute of Botany, University of Innsbruck, AT-6020 Innsbruck, Austria, <sup>4</sup>Plant Cytogenomics Group, CEITEC – Central European Institute of Technology, Masaryk University, Kamenice 5, CZ-625 00 Brno, Czech Republic and <sup>5</sup>Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, CZ-165 21 Praha 6 – Suchbátka, Czech Republic

\*For correspondence. E-mail [zuza.chumova@gmail.com](mailto:zuza.chumova@gmail.com)

†Equal contribution.

Received: 23 September 2016 Returned for revision: 14 November 2016 Editorial decision: 30 January 2017 Accepted: 10 February 2017

- **Background and Aims** Knowledge of diploid phylogeny and ecogeography provide a foundation for understanding plant evolutionary history, diversification patterns and taxonomy. The genus *Anthoxanthum* (vernal grasses, Poaceae) represents a taxonomically intricate polyploid complex with large phenotypic variation and poorly resolved evolutionary relationships. The aims of the study were to reveal: (1) evolutionary lineages of the diploid taxa and their genetic differentiation; (2) the past distribution of the rediscovered ‘Mediterranean diploid’; and (3) possible migration routes of diploids in the Mediterranean.
- **Methods** A combined approach involving sequencing of two plastid regions (*trnL-trnF* and *rpl32-trnL*), nrDNA ITS, rDNA FISH analyses, climatic niche characterization and spatio-temporal modelling was used.
- **Key Results** Among the examined diploid species, only two well-differentiated evolutionary lineages were recognized: *Anthoxanthum gracile* and *A. alpinum*. The other taxa – *A. aristatum*, *A. ovatum*, *A. maderense* and the ‘Mediterranean diploid’ – form a rather intermixed group based on the examined molecular data. In situ rDNA localization enabled identification of the ancestral *Anthoxanthum* karyotype, shared by *A. gracile* and two taxa from the crown group. For the studied taxa, ancestral location probabilities for six discrete geographical regions in the Mediterranean were proposed and likely scenarios of gradual expansion from them were suggested. Modelling past and present distributions shows that the ‘Mediterranean diploid’ has already been occurring in the same localities for 120 000 years.
- **Conclusions** Highly congruent results were obtained and dated the origin and first diversification of *Anthoxanthum* to the Miocene. The later divergence probably took place in the Pleistocene and started polyploid evolution within the genus. The most recent diversification event is still occurring, and incomplete lineage sorting prevents full diversification of taxa at the molecular level, despite clear separation based on climatic niches. The ‘Mediterranean diploid’ is hypothesized to be a possible relic of the most recent common ancestor of *Anthoxanthum* due to their sharing of ancestral features.

**Key words:** *Anthoxanthum*, rDNA FISH, incomplete lineage sorting, Mediterranean, phylogeography.

### INTRODUCTION

The Mediterranean Basin played an important role in the evolution of many plant taxa and has substantially influenced flora of the whole of Europe and adjacent regions (e.g. Feliner, 2014). Several plant groups have significantly diversified there. Thus, in addition to other taxa, some grass genera could serve as salient examples, for example *Festuca* (Inda *et al.*, 2014), *Helictotrichon* (Wölk *et al.*, 2015), *Lolium* (Inda *et al.*, 2014) and *Vulpia* (Diaz-Perez *et al.*, 2014). Grasses have often undergone very complicated evolution including reticulate hybridization and polyploidization that results in intricate diploid–polyploid complexes (e.g. Hunt *et al.*, 2014; Sun, 2014; Triplett *et al.*, 2014; Zuo *et al.*, 2015). Unravelling their evolutionary history is often extremely difficult, especially if no or only fragmentary prior knowledge is available. In such cases, research primarily focuses on diploids and their evolutionary relationships, as it is the only possibility for dealing with these

complexes (e.g. Siqueiros-Delgado *et al.*, 2013; Rojas-Andres *et al.*, 2015). A key role for the Mediterranean area in the evolution of the relatively small grass genus *Anthoxanthum* (Pimentel *et al.*, 2013), along with still-unresolved relationships among diploid taxa with uncertain taxonomic concepts (e.g. Pimentel and Sahuquillo, 2003, 2008; Pimentel *et al.*, 2010; Chumová *et al.*, 2015, 2016), first led us to attempt to resolve diploid evolution in this genus, with a special emphasis on the Mediterranean.

*Anthoxanthum* is characterized by three-flowered spikelets which are shed as a unit at maturity. The lower two florets are sterile; the upper is bisexual and protogynous. In addition, all taxa are strongly scented with coumarin, suggesting that the genus is closely related to *Hierochloë*. *Anthoxanthum*, however, has the basic chromosome number  $x=5$ , whereas *Hierochloë* has  $x=7$  (Jones, 1964). The genus *Anthoxanthum* contains approx. 22 species (excluding *Hierochloë*; Pimentel *et al.*, 2013) that grow in Eurasia, Macaronesia, Central America and

parts of Africa (tropical East Africa, Ethiopia, southern Africa) (Watson and Dallwitz, 1992; Clayton *et al.*, 2016). The genus was traditionally divided into two sections – sect. *Anthoxanthum*, comprising 9–10 species of northern Eurasia, the Mediterranean, Macaronesia and East Africa, and sect. *Ataxia*, including approx. 10 species from southern Africa (South Africa and Madagascar) and East Asia (Stapf, 1899). Recent molecular research has shown a linkage between *Anthoxanthum* sect. *Anthoxanthum* and the closely related genus *Hierochloë*, with ancient hybridization between them very likely having formed sect. *Ataxia* (Pimentel *et al.*, 2013). This finding pointed to: (1) a paraphyletic origin of the genus *Anthoxanthum* (if it is treated as separate from *Hierochloë*); and (2) a fundamental role for species of sect. *Anthoxanthum* (especially its diploids) in its evolution. This means that Macaronesia, the Mediterranean basin and mountains of Europe and Asia are the cradle of the genus *Anthoxanthum*, because only these areas are occupied by these diploids (Table 1). In sum, there are six diploid species in sect. *Anthoxanthum*, and four of them are restricted to the Mediterranean area or its vicinity. The only perennial diploid taxon of the Mediterranean is represented by plants morphologically resembling the common tetraploid species *A. odoratum*. This remains without proper scientific description, referred to in the literature as the Cretan diploid (Jones, 1964) or diploid *A. odoratum* (Felber, 1987; Teppner, 1970). For the sake of simplicity, it will hereafter be termed the ‘Mediterranean diploid’, following Chumová *et al.* (2015). The remaining three taxa of the Mediterranean are annuals – *A. gracile*, a species of shady gorges and deep valleys of central and eastern Mediterranean islands; *A. aristatum*, a quite common species of open pine forests, fallows and pastures of the western and central Mediterranean, Macaronesia and secondarily as a weed in Europe; and *A. ovatum*, a species of sandy habitats and olive groves mainly in the western and central Mediterranean (Valdés, 1973; Tutin, 1980). The two remaining diploids, from outside the Mediterranean, are a perennial found in Macaronesia – *A. maderense*, an endemic species of Madeira Island (Teppner, 1998) – and another occurring in Eurasia, *A. alpinum*, a species of arctic–alpine zones of southern mountains and northern Eurasian lowlands (Hedberg, 1986, 1990). All other members of sect. *Anthoxanthum* are perennial polyploids varying in habitat preferences, in distribution and very likely in origin (Table 1).

Despite having rather few recognized species, difficulties in their delimitation and determining their origin have frustrated people who work with this genus, especially botanists. This problem is present not only at the polyploid level, but even among diploids. For instance, the taxonomic as well as phylogenetic position of the ‘Mediterranean diploid’ among its counterparts is completely uncertain, in spite of long-standing awareness of its existence (Briquet, 1910; Borrill, 1963; Jones, 1964; Hedberg, 1967; Teppner, 1970; Felber, 1987; Chumová *et al.*, 2015). Another example is the occasional lumping of the phenotypically similar annuals *A. aristatum* and *A. ovatum* into the *A. aristatum/ovatum* complex based on genetic data (Pimentel *et al.*, 2007b, 2010), salient morphological resemblance (Pimentel and Sahuquillo, 2003; Pimentel *et al.*, 2007a, 2010) and indistinguishable but extremely variable genome size (Chumová *et al.*, 2015, 2016). The fact is that the origin of this complex still remains unknown, as does the origin of the diploid *A. maderense*. The only exception is *A. gracile*, whose position has repeatedly been shown to be as a sister group to other European diploids (Pimentel *et al.*, 2007b, 2013).

To resolve unclear relationships among taxa, the phylogeographical approach might be more powerful than analyses based either purely on population genetic structure or on historical relationships (i.e. phylogeny). Classic models describing population structure are thought to be unsuitable for groups for which historical events such as range expansion, range fragmentation or population bottlenecks have strongly determined genetic structure. Observed genetic similarity between such populations owes more to recent common ancestry than to any ongoing process of genetic exchange (e.g. Mamidi *et al.*, 2013). On the other hand, methods dealing with historical relationships at the species level often assume that taxonomic units represent non-reticulate lineages (e.g. Schaal *et al.*, 1998). Therefore, the phylogenetic approach cannot be directly applied at either the individual or the population level, where at least some genetic exchange is expected. Characterization of population structure by recognizing geographical patterns of genealogical structure across the range of species (Avise, 1994), i.e. phylogeography, balances the disadvantages of each method and therefore seems to be a promising approach for understanding complex relationships within *A. sect. Anthoxanthum*. Similarly, the phylogeographical approach has been used to resolve

TABLE 1. List of recognized species within *Anthoxanthum* sect. *Anthoxanthum* according to WCSP (Clayton *et al.*, 2016) and other sources

Species	Note	Ploidy level	Native geographical range	Life strategy
<b>Diploids (or autopolyploid derivatives)</b>				
<i>A. alpinum</i> Á.Löve & D.Löve	Accepted name <i>A. nipponicum</i> Honda (WCSP)	2x, 4x	N Eurasia	Perennial
<i>A. aristatum</i> Boiss		2x	Macaronesia, W & C Mediterranean	Annual
<i>A. gracile</i> Biv.		2x	C & E Mediterranean, NW Africa	Annual
<i>A. maderense</i> H.Teppner		2x	Endemic to Madeira	Perennial
<i>A. ovatum</i> Lag.		2x	Mediterranean	Annual
‘Mediterranean diploid’	Still not properly described species	2x	C & E Mediterranean, N Africa	Perennial
<b>Polyploids</b>				
<i>A. aethiopicum</i> I. Hedberg		4x	Endemic to Ethiopia	Perennial
<i>A. amarum</i> Brot.		16–18x	Endemic to NW Iberian Peninsula	Perennial
<i>A. nivale</i> K.Shum.		4x, 12x	From NE Congo to E tropical Africa	Perennial
<i>A. odoratum</i> L.		4x	Macaronesia, Europe to Mongolia, NW Africa	Perennial

intraspecific evolution of arctic–alpine *Saxifraga oppositifolia* in Holderegger and Abbott (2003).

Here we use plastid and nuclear ribosomal DNA sequences and cytogenetic data to describe the genetic structure of diploid European *Anthoxanthum* species and to elucidate their history and clarify their taxonomic concept. We aim to show whether gene flow or ancestral polymorphism blurs relationships of currently recognized species. Therefore, we reconstruct gene-tree genealogies and, using a multi-species coalescent model, a population tree. We further test for the presence of recent gene flow among geographically separated populations. For the subset comprising the most intricate species of the Mediterranean, including the mysterious ‘Mediterranean diploid’, we model potential historical migrations of both their populations and suitable niches.

Specifically, we address the following questions: (1) How many major evolutionary lineages can be recognized among *Anthoxanthum* sect. *Anthoxanthum*? (2) How do the genetic relationships among populations affect the taxonomic concept of the genus? (3) Is there any linkage between the genetic and geographical patterns? (4) Is there any genetic or ecological differentiation among diploid species in the Mediterranean region? (5) What can be inferred from reconstruction of the Pleistocene history of the ‘Mediterranean diploid’?

## MATERIALS AND METHODS

### Plant material

The *Anthoxanthum* L. material was collected in 2006–2015 across the entire range of the genus in Europe and included all currently recognized taxa. Ploidy levels of all plants were determined by flow cytometry, and all detected cytotypes were validated by chromosome counts, as described in Chumová et al. (2015). Thirty-five populations were selected from the 63 sampled diploid populations in the study of Chumová et al. (2015) based on their geographical positions to fully cover the known distribution of the diploid taxa of sect. *Anthoxanthum*; for locality details see Supplementary Data Table S1. Each population was represented by one specimen.

For molecular analyses (dating analysis and phylogeographical model BASTA; see below for details, and Supplementary Data Table S1 for a list of GenBank accessions), we supplemented the sequences obtained through our sampling with sequences of 37 plants from Pimentel et al. (2013).

Additionally, location data for a number of occurrences of the ‘Mediterranean diploid’, *A. aristatum* and *A. ovatum* were taken from trustworthy sources (Felber, 1987; Teppner, 1970; Valdés, 1973; Chumová et al., 2015) and used for climatic niche analysis; see further Materials and Methods sections for details.

The taxonomic concept for our 35 specimens is taken from Chumová et al. (2015), in which the annual taxa *A. aristatum* and *A. ovatum* are lumped in one complex; they are hereafter referenced as ‘annuals’. However, this simplification cannot be used for climatic niche analysis, and therefore data taken from Valdés (1973) are retained in their original form and these are treated as two separate taxa. The validity of both taxonomic concepts is examined in the Discussion.

### Fluorescence in situ hybridization (FISH) analyses

Altogether, 16 populations (one specimen from each) of five recognized diploid taxa were chosen for FISH analysis (see Table S1). Chromosome spreads were prepared as described by Chumová et al. (2016). *Arabidopsis thaliana* BAC clone T15P10 (AF167571) bearing 45S rRNA gene repeats was used for *in situ* localization of 45S rDNA, and *A. thaliana* clone pCT 4.2 (M65137), corresponding to a 500-bp 5S rRNA repeat, was used for localization of 5S rDNA loci. The 45S and 5S rDNA probes were labelled by digoxigenin- and biotin-dUTP, respectively, using nick translation (Mandáková and Lysak, 2016). We pooled together 100 ng from each labelled BAC DNA, precipitated using ethanol, then dissolved this in 20 µL of hybridization mixture containing 50 % formamide and 10 % dextran sulphate in  $2 \times$  SSC and pipetted it onto microscopic slides. The slides were heated at 80 °C for 2 min and incubated at 37 °C overnight. Hybridized probes were visualized using fluorescently labelled antibodies for digoxigenin-dUTP (green) and biotin-dUTP (red). DNA labelling and fluorescence signal detection was performed according to Mandáková and Lysak (2016). Chromosomes were counterstained with DAPI ( $2 \mu\text{g mL}^{-1}$ ) in Vectashield. Fluorescence signals were analysed and photographed using a Zeiss Axioimager epifluorescence microscope and a CoolCube camera (MetaSystems, Newton, MA, USA). Individual images were merged and processed using Photoshop CS software (Adobe Systems).

### Molecular data collection

For all molecular analyses, we used 35 selected individuals (one from each population). We analysed two plastid regions (*trnL-trnF* and *rpl32-trnL*) and internal transcribed spacer (ITS) region of nuclear ribosomal DNA (nrDNA). Total genomic DNA was extracted from 0.5 g of dried leaf tissue using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). The plastid region *trnL-trnF* was amplified and sequenced using primers *c* and *f* (*trnL* intron plus *trnL-trnF* intergenic spacer, hereafter *trnL-F*; Taberlet et al., 1991). The PCR conditions for the *trnL-F* region were as follows: 1 min of denaturation at 94 °C, followed by 34 cycles at 94 °C for 50 s, 50 s at 52.5 °C, 90 s at 72 °C and a final extension of 10 min at 72 °C. The plastid region *rpl32-trnL* was amplified and sequenced using primers *trnL*<sup>(UAG)</sup> and *rpl32-F* under the PCR conditions following Shaw et al. (2007). Finally, amplification of the nuclear ribosomal ITS (ITS1–5.8S–ITS2) region followed Hsiao et al. (1995). Due to the occurrence of within-individual polymorphisms in some of the directly sequenced PCR products of the ITS region, PCR products for ITS were cloned using the pGEM-T Easy Vector System (Promega, Madison, WI, USA) following the manufacturer’s instructions but downscaled to half-volume reactions. Details of the procedure were as described by Závěská et al. (2012). Eight colonies from each individual were used as templates for PCR and sequencing. PCRs of all loci were done with MyTaq polymerase (Bioline, London, UK) or AmpliTaq Gold DNA Polymerase (Applied Biosystems, Foster City, CA, USA), following the manufacturer’s instructions, but with the annealing temperatures mentioned above. PCR products were purified with the Jetquick PCR Purification Spin Kit (Genomed, Warsaw, Poland) and

TABLE 2. Definitions of population tree OTUs and details about ITS ribotypes, cpDNA haplotypes, localities, karyotypes, life strategy, codes for 35 populations included in all molecular analyses, and regional affiliation in BASTA analysis. Initial small letter in ribotype/haplotype name indicates its colour in Fig. 1 (b, blue; y, yellow; r, red; g, green)

Population tree OTUs	ITS ribotype	cpDNA haplotype	Locality	No. of the karyotype (same as in Fig. 2A)	Life strategy	Populations included	BASTA region	
1	alp_ALP	bR2, bR4	bH2, bH1	Alps	2	Perennial	AT02alp, FR01alp, FR04alp, CH02alp	
2	alp_BAL	bR1	bH3	Balkans		Perennial	ME01alp	–
3	alp_CAR	bR1, bR5, bR6	bH1	Carpathians	3	Perennial	CZ01alp, RO01alp, SK03alp, UA06alp	–
4	alp_CAU	bR3	bH4	Caucas		Perennial	GE01alp	–
5	alp_DEF	bR1	bH1	‘defrost’ (Island, Norway)	2	Perennial	IS01alp, NO01alp	–
6	annu_SIP	rR11	rH8, gH15	South Iberian Peninsula	5	Annual	ES06annu, ES07annu	IBE
7	annu_CIP	yR20, yR21	yH16, yH17	Central Iberian Peninsula	4	Annual	PT06annu, ES09annu	IBE
8	annu_COR	gR13	gH13	Corsica	6	Annual	FR11annu	COR
9	med_CAP	rR8, rR9, rR10, rR12, gR16	rH6, rH7, gH14	Central Apennine Peninsula		Perennial	IT01med, IT02med, IT03med	APE
10	med_BAL	gR15	rH5, rH8, gH11, gH13	Balkans	1	Perennial	Al02med, BG03med, GR06med, HR01med, MK01med, MK02med, ME05med	BAL
11	med_COR	gR17, gR18, gR19	gH13	Corsica		Perennial	FR06med, FR08med, FR09med	COR
12	med_CRE	gR14	gH12	Crete		Perennial	GR03med	BAL
13	mad_MAD	rR7	rG9, rH10	Madeira	1	Perennial	PT01mad, PT03mad	MAD
14	grac_CRE	–	–	Crete	1	Annual	GR09grac, GR10grac	–

directly sequenced at Macrogen Inc. (<http://www.macrogen.com>) or at the DNA sequencing laboratory of the Biological Section of Charles University, with the original PCR primer sets in both directions.

The two datasets of sequences, i.e. the concatenated chloroplast DNA (cpDNA) dataset and the ITS dataset, were aligned independently using MAFFT (<http://mafft.cbrc.jp/alignment/server/>), and then the alignments were improved manually in BioEdit v.7.0.0. (Hall, 2004). Autapomorphies found in a particular alignment in only one sequence were considered as polymerase errors and corrected (Popp *et al.*, 2005). ITS sequences from cloned PCR products were inspected for presence of PCR and/or *in vivo* recombinants according to Závěská *et al.* (2012), and a maximum of two ITS alleles per diploid were kept for further analyses.

#### Reconstructions of gene genealogies

Because the interpretation of congruence (or lack thereof) between the geographical distributions of haplotypes and their genealogical relationships is the basis for phylogeographical analyses, we first reconstructed statistical parsimony networks for the plastid DNA data and ITS ribotypes (analysed as separate datasets) using TCS 1.21 (Clement *et al.*, 2000; Swofford, 2002). Gaps were treated as missing characters after manual re-coding of insertions/deletions (indels) as single-base-pair indels. For all other analyses, the unmodified alignments were used.

Phylogenetic trees based on cpDNA and ITS sequences (analysed as separate datasets) were reconstructed using maximum parsimony (MP) and Bayesian analyses (BA). MP analysis (including the MP analysis with 1000 bootstrap replicates) were performed using PAUP\* v.4.0b10 (Swofford, 2002) with settings as described by Závěská *et al.* (2012). Bootstrap support (BS) was categorized as strong (>85%), moderate (70–85%), weak (50–70%) or poor (<50%). Bayesian analysis was accomplished with MrBayes v.3.1.2. (Huelsenbeck and Ronquist, 2001). Both datasets were individually tested for the best substitution model using jModelTest v.0.1.1 (Posada, 2008) with default settings. We chose GTR+G for both of them according to the Bayesian information criterion (BIC; Schwarz, 1978; Posada and Buckley, 2004). Two parallel runs with four chains each were used, sampling every 1000th tree for 10 million generations for each dataset. The first 25% of samples (= 2500 trees) were discarded as burn-in, and the remaining 7500 trees per run were summarized. Nodes with posterior probability (PP) values of 0.95 and above were regarded as significant and those with PP values below 0.95 regarded as non-significant. For presentation of relationships between sampled individuals based on particular markers, Bayesian 50% majority rule consensus trees were used.

A test of the compatibility of the two datasets with theoretically different evolutionary histories was performed in Concaterpillar (Leigh *et al.*, 2008). To visualize potential conflicting signals, we analysed the concatenated dataset of cpDNA and ITS as a single supermatrix using NeighbourNet (Bryant and Moulton, 2004). The network was constructed based on pairwise (uncorrected) distance matrices in SplitsTree4 v.4.9.1. (Huson and Bryant, 2006).

### Species tree reconstructions

To reveal relationships between populations that are now isolated geographically or belong to separate species, we reconstructed a population tree using a multi-species coalescent model based on two independent alignments using the module \*BEAST (Heled and Drummond, 2010) implemented in BEAST v.1.8.0 (Drummond and Rambaut, 2007). For population tree reconstruction, we predefined the groups of populations [i.e. operational taxonomic units (OTUs) in the final population tree] that generally corresponded to the population grouping in the cpDNA and ITS gene trees. The final setting for population tree reconstruction sorted 35 populations sampled for cpDNA as well as ITS into 14 OTUs representing all plausibly independently evolving lineages of the studied group (see details in Table 2).

To test whether our assumption of absence of gene flow among geographically isolated populations is correct, we applied the posterior predictive checking approach (Joly et al., 2009), which enables testing the suitability of a coalescent model for the selected group of species. In this analysis, we inspected a population tree reconstructed in \*BEAST using the program JML (Joly, 2012). The scenario that gene trees within species tree(s) does not account for hybridization was simulated based on posterior distributions of species-tree parameters from \*BEAST analysis and original alignments of sequence data. Details on this analysis are described in Supplementary Data Text S1.

### Divergence dating analysis

Bayesian divergence date analyses were previously conducted for the entire genus *Anthoxanthum* by Pimentel et al. (2013). Here we follow the methods from that study so that our data can be interpreted over the same evolutionary time scale. For these analyses, in addition to the sequences obtained from our samples, we used 30 sequences from Pimentel et al. (2013) representing 19 populations of *Anthoxanthum* species and three populations of *Hierochloë* species to supplement our sampling and improve the fitting of our data (Table S1). Because Pimentel et al. (2013) did not use data from plastid region *rpl32-trnL*, our dating analyses were based only on ITS and *trnL-trnF* data. For estimation of the divergence time of main lineages, we analysed ITS and plastid data under a multispecies-coalescent model using \*BEAST as described above, but with additional priors as follows. For the divergence between tribe Meliceae and tribes Aveneae/Poeae + Triticeae, we used a normal prior distribution with mean 32.1 Myr and standard deviation of 3.65. For the divergence between tribes Aveneae/Poeae and Triticeae, we assumed a normal prior distribution with mean 23.4 Myr and standard deviation of 3.1. For the divergence of genera *Hierochloë* and *Anthoxanthum*, we employed a normal prior distribution with mean 10.4 Myr and standard deviation of 1.0, and secondary calibration points taken from Pimentel et al. (2013). Priors for substitution rates of plastid DNA were established following Wolfe et al. (1987), Drummond and Rambaut (2007) and Winkler et al. (2012), with wide ranges to cover most biologically realistic values. Therefore, we used a lognormal prior distribution with a mean 0.007 and standard deviation 1.0 Mya to enable use of a realistic

substitution rate of plastid DNA with mean  $1.75 \times 10^{-3}$  substitutions per site per million years.

Four independent analyses were run for a total of 100 000 000 generations. Log files were analysed using TRACER v.1.5 to assess convergence and ensure that the effective sample size (ESS) for all parameters was >200 (Drummond and Rambaut, 2007). In addition, tree convergence was evaluated using the tool RWTY (<https://github.com/danlwarren/RWTY>), the upgraded version of online application AWTY (Nylander et al., 2008). Resulting trees were combined using LogCombiner v.1.7.2 (Drummond and Rambaut, 2007) with a burn-in of 25 %. Subsequently, a maximum credibility tree was constructed using TreeAnnotator v.1.7.2 (Drummond and Rambaut, 2007).

### Spatio-temporal distribution of the ‘Mediterranean diploid’ and its closest relatives

Changing spatial distributions over time and migration trends of the subgroup of special interest, i.e. the ‘Mediterranean diploid’ and its closest relatives, were inferred using the recently described phylogeographical model BASTA (Bayesian Structured coalescent Approximation; De Maio et al., 2015). Applying this approach, migration rates and root locations are estimated based on the structured coalescent, which is supposed to be more accurate and less sensitive to biased sampling than methods based on discrete trait models (e.g. Lemey et al., 2009). Because the different gene genealogies might indicate different migration routes and different ancestral distribution areas (reflecting also different time scales of migration), we first analysed plastid and ITS datasets separately and then analysed both datasets together. To improve our geographical sampling, we included 16 sequences representing 11 populations of diploid species from Pimentel et al. (2013) in our dataset and therefore again analysed only ITS and *trnL-trnF* data. The analyses were done in BEAST v.2.3.1 based on a user-prepared xml file specifying the settings of the analyses. Individuals were assigned to five discrete geographical regions: Apennine Peninsula (APE); Balkan Peninsula and Crete (BAL); Corsica and Sardinia (COR); Iberian Peninsula and Morocco (IBE); and Madeira (MAD). We used an asymmetric geospatial model, i.e. we enabled migration rate in one direction to be different from that in the reverse direction. Bayesian Stochastic Search Variable Selection (BSSVS) was applied to achieve statistical efficiency. We forced the analyses to equalize the effective population sizes for different locations in order to decrease the effect of potentially biased sampling across particular regions. Log files of the final analyses were analysed using TRACER v.1.5 to assess convergence and ensure that the ESS for all parameters was >200.

### Niche differentiation of the ‘Mediterranean diploid’ and its closest relatives

Environmental parameters based on climatic data from WordClim database (Hijmans et al., 2005; <http://www.worldclim.org>) were assessed to try to elucidate the degree of ecological divergence among closely related taxa of the Mediterranean basin (the ‘Mediterranean diploid’, *A. ovatum*

and *A. aristatum*). Altogether, for 298 georeferenced locations, we extracted data from a database of 19 bioclimatic variables representing annual trends, seasonality, and extreme and limiting environmental factors. For the taxon ‘Mediterranean diploid’, only those localities (66 locations) with plants determined by flow cytometry (Chumová *et al.*, 2015) or chromosome counts (Felber, 1987; Teppner, 1970) were used as georeferenced points for range delimitation. Occurrence data for *A. aristatum* (168 locations) and *A. ovatum* (64 locations) were adopted from comprehensive work about their distribution by Valdés (1973) and georeferenced manually (see Supplementary Data Fig. S1). Data from raster layers for each of the 19 bioclimatic variables was extracted using the *extract* function in the *raster* package in R (Hijmans *et al.*, 2016). First, principal trends in variation of bioclimatic variables were detected by principal component analysis. Next, mutually uncorrelated variables were identified by stepwise forward selection and used for linear discriminant analysis. All analyses were conducted via the *MorphoTools* R package for multivariate data handling (Koutecký, 2015).

#### *Distributional modelling of the ‘Mediterranean diploid’*

The maximum entropy machine-learning approach implemented in Maxent v.3.3.3 k (Elith *et al.*, 2006; Phillips *et al.*, 2006) was used for modelling the potential distribution pattern of the ‘Mediterranean diploid’ based on climatic data from the WorldClim database (Hijmans *et al.*, 2005). We inferred the potential ranges at present and during the Last Glacial Maximum (LGM, approx. 21 000 years ago) and the Last Interglacial (LIG, approx. 120 000–140 000 years ago) periods. This approach allowed us to explore shifts in potential distribution patterns within the last 120 000 years and to reconstruct population histories (withdrawals or expansions). Sixty-six thoroughly proved populations of the ‘Mediterranean diploid’ (see Table S1 for detail) were used for modelling. Layers for the 19 bioclimatic variables were used for inferring distribution. A 30-s resolution (approx. 1 km) for current data and the LIG climatic model of Otto-Bliesner *et al.* (2006) were used. Two climatic models (both of resolution 2.5 min, approx. 4 km) were used for the LGM period: MIROC (Model for Interdisciplinary Research on Climate; Hasumi and Emori, 2004) and CCSM4 (Community Climate System Model; Collins *et al.*, 2006). First, however, the bioclimatic variables were clipped such that they reached from 25°N to 72°N and from 25°W to 45°E, including all potentially suitable habitats for *Anthoxanthum* species in Europe, the Middle East and North Africa (accordingly to Felber, 1987; Chumová *et al.*, 2015). Only eight relatively uncorrelated bioclimatic layers ( $r < 0.7$ ) were used for the analysis: mean diurnal range (BIO2); temperature seasonality (BIO4); mean temperature of wettest quarter (BIO8); mean temperature of driest quarter (BIO9); mean temperature of warmest quarter (BIO10); precipitation of wettest month (BIO13); precipitation of driest month (BIO14); and precipitation seasonality (BIO15).

The default setting of Maxent with the ‘auto features’ options was used for independent fivefold cross-validation runs. The analysis parameters were set as follows:  $10^{-5}$  convergence threshold, 1000 maximum iterations, regularization parameter

$\beta = 1$ , 100 000 background points and 20 % (13 localities) as a testing data set. The AUC (area under the curve) statistic (Peterson *et al.*, 2008) was used to evaluate the quality of the final model. For model projection, a median of five replicates and the ‘equal training sensitivity and specificity’ logistic threshold (Liu *et al.*, 2005) were used. The potential area of long-term stable occurrence of ‘Mediterranean diploid’ populations was shown by the overlap of suitable areas predicted for the present, LGM and LIG (data processed in GRASS GIS; Neteler *et al.*, 2012).

## RESULTS

### *Plastid DNA*

The lengths of the *trnL-trnF* and the *rpl32-trnL* intergenic spacers in the *Anthoxanthum* full dataset (including all 35 specimens) were 965 and 854 bp, respectively. The combined alignment was 1819 bp long and comprised 95 variable characters including 17 indels and 34 parsimony-informative sites. A total of 17 haplotypes in 35 individuals were identified (Table 2). The original alignment is available at datadryad.org; GenBank accession numbers are provided in Table S1.

The statistical parsimony network (TCS, Fig. 1A) revealed five main genetic groups, of which the most divergent group corresponds to *A. gracile*, which fell out at the 95 % connection limit. All remaining haplotypes were connected within a single network and did not differ by more than 23 mutations. The most distant was a group of haplotypes specific to populations of *A. alpinum* (‘blue haplotype group’, bH1–bH4). The second-most different group, further referred to as the ‘yellow haplotype group’, corresponds to populations of annuals from the central Iberian Peninsula (yH16, yH17). Two central groups of haplotypes correspond to two sets of populations. The first set comprises populations of *A. maderense* from Madeira, annuals from the southern Iberian Peninsula and populations of the ‘Mediterranean diploid’ spanning the Apennines and Balkan Peninsula (‘red haplotype group’, rH5–rH10). The other set comprises populations of annuals from Corsica in addition to the rest of the ‘Mediterranean diploid’ populations (‘green haplotype group’, gH11–gH15).

Hierarchical genetic relationships resolved by MP and Bayesian inference based on the cpDNA dataset (Supplementary Data Fig. S2) generally gave the same pattern as the TCS analyses, showing four main groups of haplotypes within the ingroup (*A. gracile* was used as an outgroup). A strongly supported basal group (BS = 97 %, PP = 1.00, in blue in Fig. S2) corresponds to all haplotypes of the diploid *A. alpinum*. In sister position to the *A. alpinum* group, haplotypes of all populations of the ‘Mediterranean diploid’, annual taxa and *A. maderense* were resolved (BS = 94 %, PP = 0.93). Within this group, only the yellow haplotype group resolved in TCS analysis was strongly supported as monophyletic (BS = 99 %, PP = 1.00).

### *nrDNA data – ITS*

ITS alignment in the *Anthoxanthum* full dataset was 687 bp long and comprised 86 variable characters including 18 indels and 29 parsimony-informative sites. In three individuals

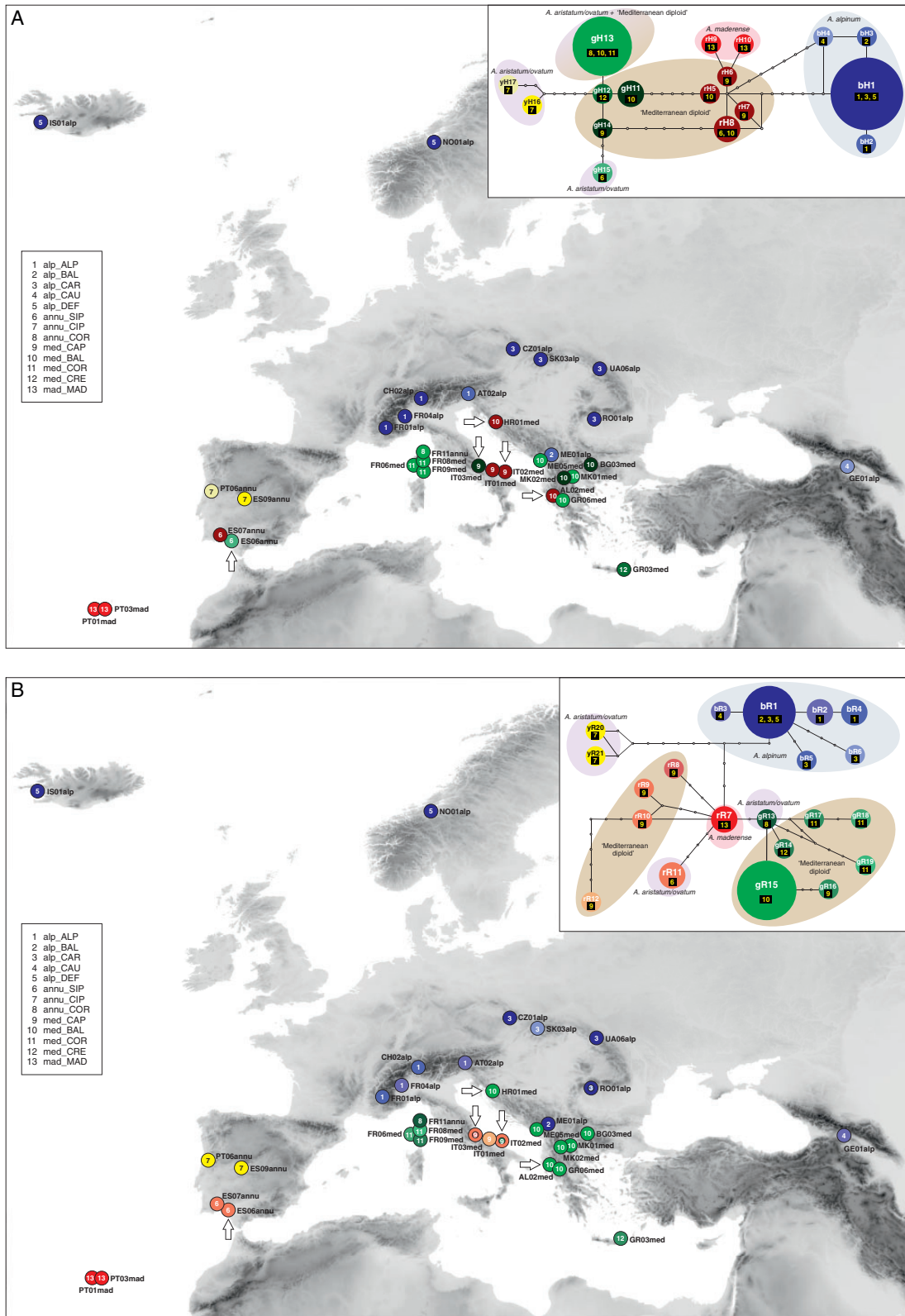


FIG. 1. Geographical patterns and statistical parsimony networks (TCS) of plastid DNA and ITS variation in European *Anthoxanthum* (outlying *A. gracile* excluded). (A) Variation in plastid DNA haplotypes based on *trnL-trnF* and *rpl32-trnL*. (B) Variation in ITS ribotypes. TCS networks: small black dots represent haplotypes/ribotypes not sampled. Haplotypes/ribotypes are divided into four groups (blue, yellow, red and green) and are shown in the insets arranged by their taxonomic membership. Names of haplotypes/ribotypes within networks correspond to names of haplotypes/ribotypes in Table 2. Numbers correspond to their regions of origin (listed in box and in Table 2). Maps: dots representing the sampled populations are coloured according to their haplotypes/ribotypes (in the same colours as in network). Arrows indicate the populations bearing mutually incongruent cpDNA haplotypes and ITS ribotypes. Population names correspond to names in Table S1 and include taxa abbreviations: alp, *A. alpinum*; annu, *A. aristatum/ovatum*; mader, *A. maderense*; med, 'Mediterranean diploid'. Numbers within population dots correspond to their regions of origin (listed in box and in Table 2).

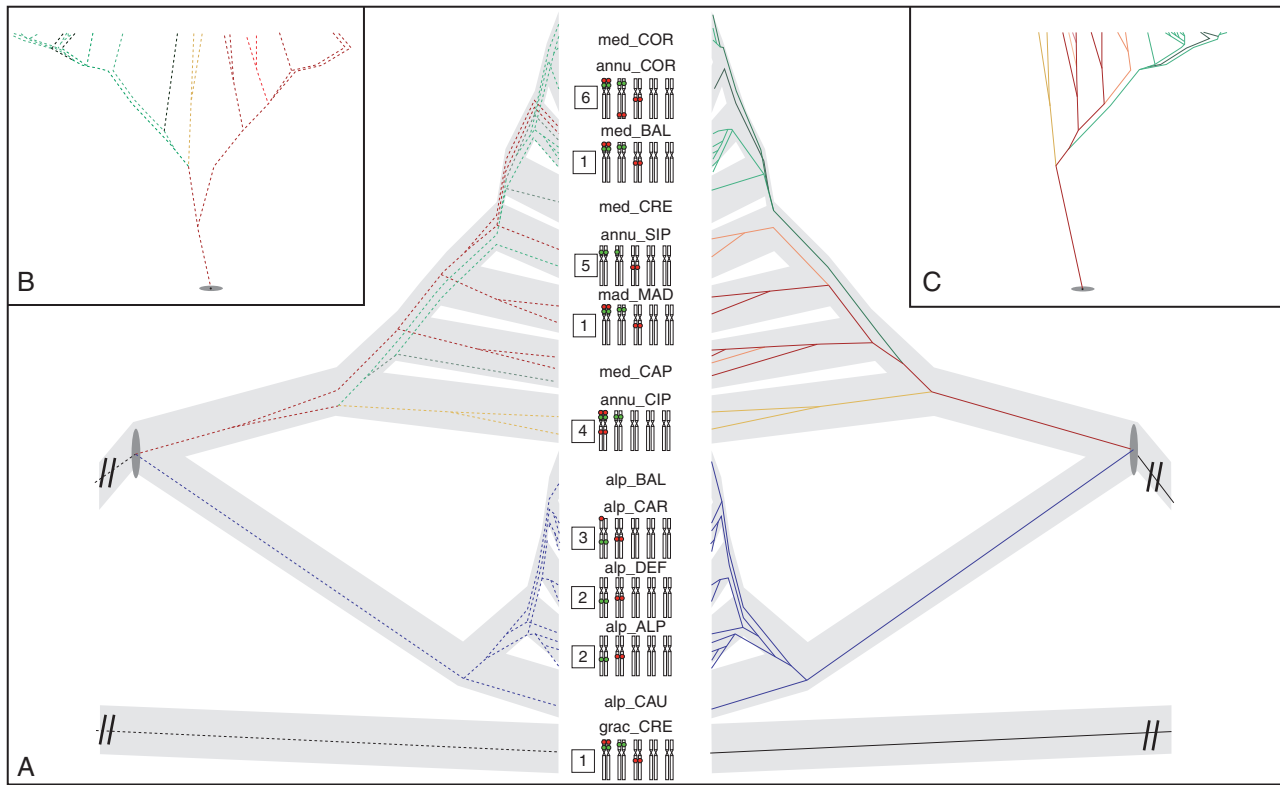


FIG. 2. (A) Population tree (two mirrored copies depicted in grey) reconstructed using multi-species coalescent model with reconciled gene trees (depicted in colour on the background of the grey population tree; cpDNA gene tree on the left, ITS tree on the right). Topology of population tree depicts relationships among geographically isolated populations; reconciled gene tree highlights gene tree–species tree incongruence explained by ILS. Between mirrored population trees, karyotypes specific for each group/OTU are shown, numbered 1–6 according to Table 2; red dot – 5S rDNA, green dot – 45S rDNA. (B) Unfolded gene tree based on the plastid data; close-up of the relationships of Mediterranean populations. Colouring of the branches corresponds to colours of haplotypes in Fig. 1. (C) As for B, but the results are based on ITS data.

(RO01alp, IT02med, IT03med; Table 2), two ITS ribotypes were detected after cloning analyses. A total of 20 ribotypes in 35 individuals were detected and analysed (Table 2). The original alignment is available at datadryad.org; GenBank accession numbers are provided in Table S1.

The statistical parsimony network (Fig. 1B) revealed four main ingroup haplotype groups with similar geographical distributions to those based on plastid data. Therefore, we kept similar naming and colouring patterns, and we highlight the differences in ITS patterns from those of cpDNA. The ‘blue ribotype group’ corresponds completely to the *A. alpinum*-specific blue cpDNA haplotype group. Similarly, the ‘yellow ribotype group’ corresponds to the yellow haplotype group. The ‘red ribotype group’ (ribotypes rR7–rR12) and the ‘green ribotype group’ (ribotypes gR13–gR19) differ from their respective cpDNA haplotype groups by having more specific geographical distributions in the Mediterranean. The red ribotype group is the only ribotype group occurring in Madeira and southern regions of the Iberian Peninsula, and its eastern-most occurrence is in the Apennines, where it co-occurs (in a single case) with the green ribotype group. Green ribotypes are also distributed in the Balkans, Corsica and Crete, where no other ribotype groups occur.

The topology of the Bayesian gene tree based on ITS data (Fig. S3) differed substantially from the gene tree based on cpDNA. While the grouping of individuals in the ITS gene tree

was generally the same as that in the cpDNA gene tree, hierarchical relationships among the blue, yellow, red and green ribotype groups based on ITS data differed from those among the haplotype groups based on cpDNA data (see detailed comparison in Fig. 2A and 2B). Additionally, the grouping based on ITS data was more reflective of the geographical origins of the individuals (as also detected in TCS analyses). The group of blue ribotypes was resolved as monophyletic but with only poor statistical support. Similarly as for cpDNA, *A. alpinum* individuals (blue ribotypes) were in sister position to the remaining samples. Within those, strongly supported were the group of yellow ribotypes (BS = 100 %, PP = 1.00) and its sister group including a group of green ribotypes (BS = 60 %, PP = 1.00) and non-monophyletic red ribotypes.

The recognized cpDNA haplotypes and ITS ribotypes were mostly congruent, with the exception of five individuals from the Iberian Peninsula (ES06annu), Apennines (IT02med, IT03med) and Balkans (HR01med, AL02med; Fig. 1).

#### Gene tree (in)congruence, population-tree reconstruction and posterior predictive check

Congruence testing of the plastid and ITS datasets indicates significant incongruence between the datasets. Results of



NeighbourNet analyses presented in Supplementary Data Fig. S4 based on the cpDNA dataset (Fig. S4A), ITS dataset (Fig. S4B) and concatenated dataset (Fig. S4C) show clearly the samples causing this incongruence. Such incongruent datasets are not recommended to be analysed as a concatenated dataset to obtain a species or population tree because of their different evolutionary histories (Lowe *et al.*, 2004; Knowles and Kubatko, 2010). On the other hand, reconstruction of a population tree in a multi-coalescence framework could explain these incongruences due to incomplete lineage sorting (ILS).

The resulting population tree visualized with reconciled cpDNA and ITS gene trees is presented in Fig. 2A (for original population tree see Supplementary Data Fig. S5). The population tree topology recovered *A. gracile* as an outgroup and two statistically supported groups in sister positions as an ingroup. The first of these groups comprised all *A. alpinum* individuals (blue branches in Fig. 2A and Fig. S5, PP = 1.00), and the second included all individuals of the ‘Mediterranean diploid’, annuals and *A. maderense* (yellow, red and green branches of gene trees in Fig. 2A and Fig. S5, PP = 0.98). In the other group, the populations of annuals from central Spain and Portugal (yellow branches in Fig. 2A) were resolved as most basal. The remaining part of the population tree is neither well resolved nor statistically supported. The distributions and relationships of cpDNA haplotypes and ITS ribotypes within the population tree are visualized using reconciled cpDNA (left) and ITS (right) gene trees, respectively, within the population tree in Fig. 2A. For better visibility, the reconciled gene trees are presented also in unfolded form (cpDNA gene tree in Fig. 2B, ITS gene tree in Fig. 2C).

The posterior predictive check performed in JML shows that the incongruence between the cpDNA gene tree and the population tree can be explained purely by the deep coalescence process, as no minimum pairwise distances within cpDNA alignment occurred with low probability ( $P < 0.05$ ) in the simulated population tree assuming no gene flow. We visualized this result by reconciling the cpDNA gene tree with the population tree (Fig. 2A on the left) and also by presenting the unfolded cpDNA gene tree (Fig. 2B). In contrast, the posterior predictive check applied to the ITS dataset yielded several pairs of individuals between which genetic distances were shorter than would be expected with pure coalescence; therefore, these distances were observed only with low probability in a simulated population tree assuming no gene flow and indicated violation of the multi-species coalescent model.

#### Divergence dating of population tree

The scaled \*BEAST analyses performed on the dataset augmented by samples from Pimentel *et al.* (2013) resolved relationships among outgroup and ingroup taxa with strong statistical support (see details in Fig. 3). The divergence of genus *Anthoxanthum* was dated to the late Miocene [95 % highest posterior density (HPD) 4.76–7.72 Mya, median 6.17 Mya], and divergence of *A. alpinum* and the remaining species from *A. sect. Anthoxanthum* (i.e. Mediterranean taxa) was estimated to have occurred in the Late Pliocene–Early Pleistocene (95 % HPD 1.33–2.8 Mya, median 2.02 Mya). The most recent common ancestor (MRCA) of the Mediterranean taxa diversified in

the Pleistocene around 1.51 Mya (95 % HPD 0.96–2.19 Mya, median 1.51 Mya; see details in Table 3).

#### Localization of rDNA loci

All of the analysed *Anthoxanthum* populations have a diploid chromosome number ( $2n = 2x = 10$ ). Eight of the 16 analysed *Anthoxanthum* accessions exhibit a common set of rDNA loci with two pairs of 5S and 45S (Table 2, Fig. 2A and Supplementary Data Fig. S6). This pattern was observed in *A. gracile*, all ‘Mediterranean diploids’ and *A. maderense*. Because *A. gracile* is the ancestral species to all other *Anthoxanthum* taxa (see above), we refer to this karyotype as the ‘ancestral *Anthoxanthum* karyotype’ (see Fig. S3). Both annual taxa (*A. aristatum* and *A. ovatum*) possess several independent lineage-specific rDNA alterations including amplification and reduction of 5S and 45S rDNA loci (Table 2, Fig. 2A, Figs S3 and S6). All *A. alpinum* exhibit one pair of both the 5S and the 45S rDNA loci, but *A. alpinum* from the Carpathians possesses an additional pair of 5S.

#### Spatio-temporal distribution of the ‘Mediterranean diploid’ and its closest relatives

Results of the phylogeographical modelling with BASTA are presented in the form of an asymmetric migration rate matrix (Table S2) and in range connectivity diagrams in Fig. 4.

An asymmetric model based on plastid data identified the most important migration pathway as from the Apennines to Madeira. ITS data revealed the strongest migration to be from the Apennines to the Balkans (and Crete). Both datasets indicated the next-most important migrations to have been from Corsica and Sardinia to the Balkans (and Crete) and also backwards from the Balkans to Corsica. Finally, ITS data showed lower support for migrations from the Iberian Peninsula (and Morocco) to Madeira.

Combined analysis based on cpDNA and ITS data basically summarized the results of the separately analysed datasets. Cumulatively, it revealed the most important migrations to be from Corsica (and Sardinia) to the Balkans (and Crete) and from the Apennines to the Balkans (and Crete). Details and particular values of BSSVS indicators are presented in Table S2.

Ancestral locations for internal tree nodes were reconstructed and interpreted based on ITS data only, as the current version of BASTA does not enable cumulative interpretation of trees, and ancestral locations from combined datasets or from analyses based on plastid data was not informative enough. The ITS gene genealogy with inferred ancestral areas is presented in Supplementary Data Fig. S7. The Balkans (and Crete) was estimated to be the most probable ancestral root location with PP of 0.42. The Iberian Peninsula is suggested to be the ancestral area for all populations of annuals with yellow haplotypes and ribotypes currently distributed in central Spain and Portugal. The populations bearing red ITS ribotypes first occurred in the Apennines and later also in both the Iberian Peninsula (along with Morocco), and especially in Madeira. The ancestral area of the green ITS ribotypes is suggested to be the Balkans together with Corsica and Sardinia

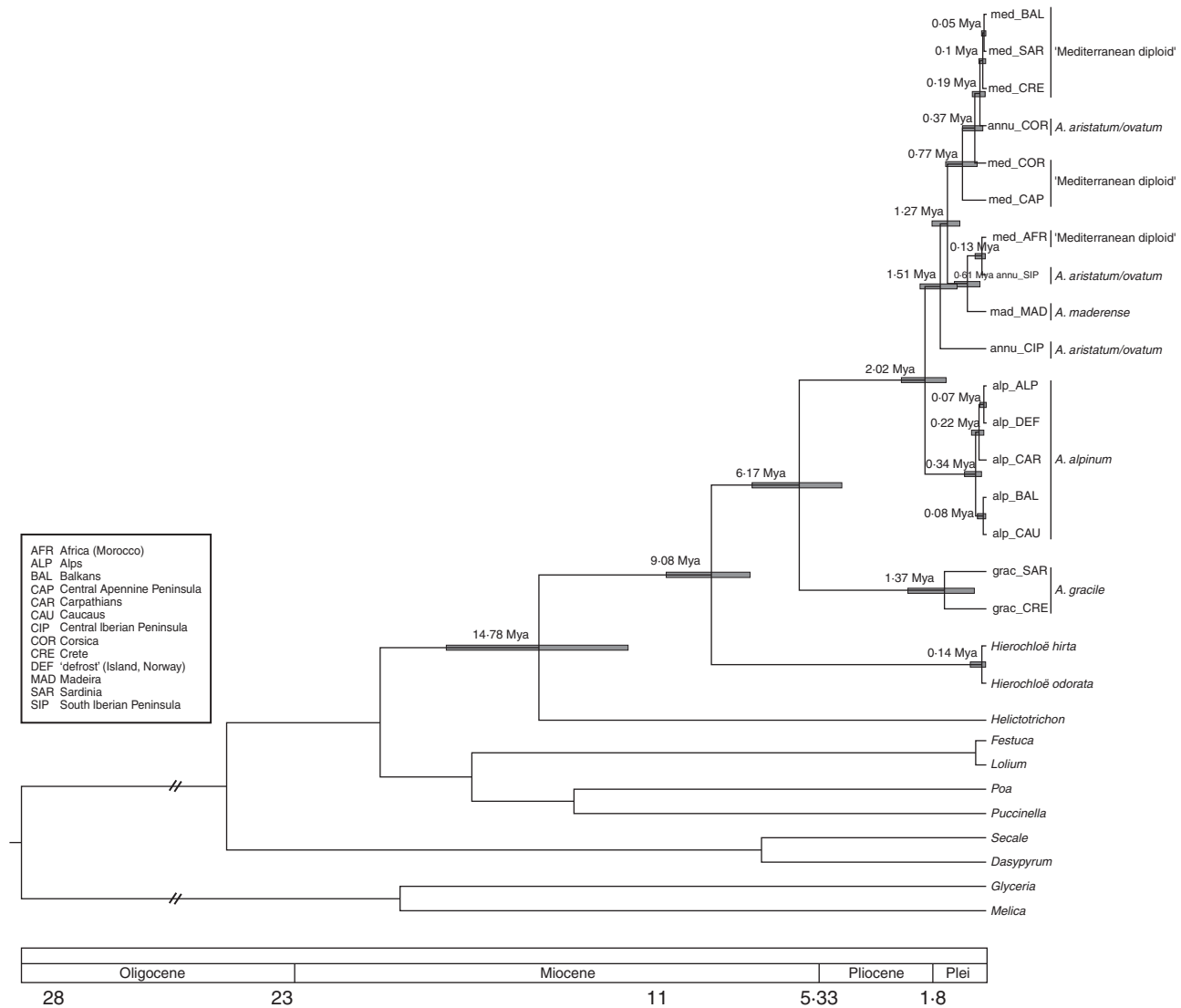


FIG. 3. Maximum clade credibility tree resulting from dated \*BEAST species tree analysis based on plastid and ITS data. Node labels correspond to divergence time estimates in millions of years ago (Mya). Grey node bars indicate 95 % highest posterior density (HPD) interval. Plei, Pleistocene.

TABLE 3. Divergence time estimates for the main groups of *Anthoxanthum* sect. *Anthoxanthum* based on plastid and ITS data analysed using a multi-species coalescent model

Divergence between/of	Median (Mya)	95 % highest posterior density (HPD)
<i>Anthoxanthum</i> vs. <i>Hierochloë</i>	9.08	7.79–10.57
<i>A. gracile</i> vs. remaining species of <i>A.</i> sect. <i>Anthoxanthum</i>	6.17	4.76–7.72
<i>A. alpinum</i> vs. Mediterranean taxa	2.02	1.33–2.8
Mediterranean taxa	1.51	0.96–2.19

*Niche differentiation of Mediterranean taxa*

Analysis of environmental/climatic factors was performed to identify whether and to what extent three closely related taxa of the Mediterranean area ('Mediterranean diploid', *A. aristatum* and *A. ovatum*) differ. Linear discriminant analysis based on 298 localities (Fig. 5) revealed a significant difference in climatic niches among taxa ( $F_{9,288} = 45.6, P < 0.001$ ) and

identified the nine most important climatic factors promoting such divergence (Fig. 5 and Supplementary Data Fig. S8). The main trend in climatic data support differentiation of *A. aristatum* from the other two taxa and is highly correlated with increasing isothermality (i.e. diurnal temperature oscillations relative to annual temperature oscillations) and decreasing mean temperature of the coldest quarter. This divergence is further supported by increasing minimum temperature of the

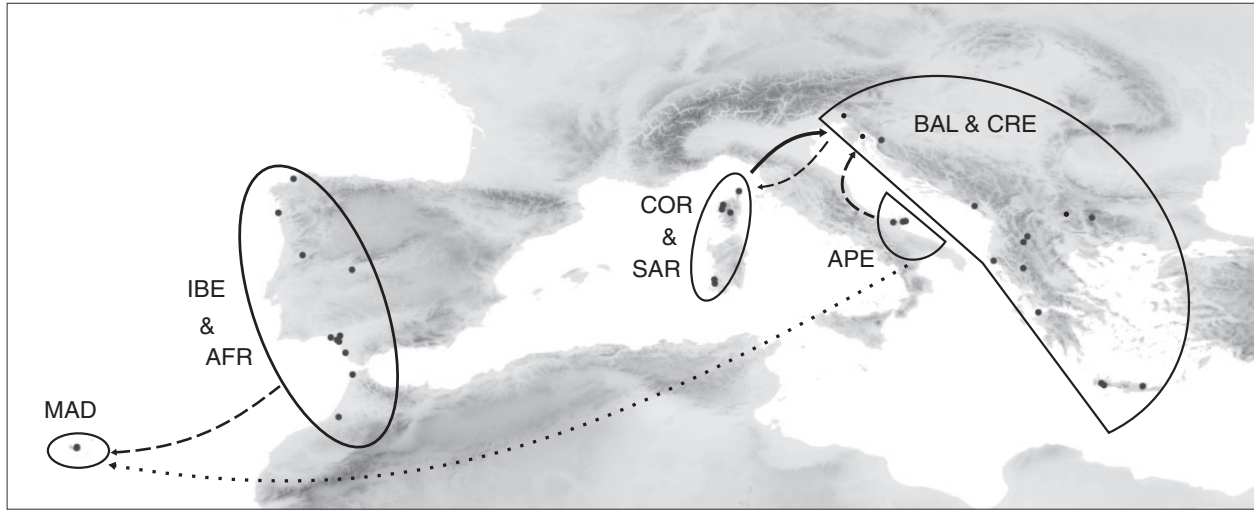


FIG. 4. Range connectivity and migration directions of Mediterranean populations of *Anthoxanthum* inferred using asymmetrical BASTA phylogeographical model. Migration supported in analysis of plastid data and in combined plastid and ITS analysis is depicted using a dotted line. Migrations supported in analysis of ITS data (and in combined analysis) are depicted using dashed lines. Migration supported in plastid, ITS as well as in combined analysis is represented by solid line. Direction of migration is indicated by arrows. The thickness of the lines is proportional to their support by BSSVS indicators (only connections with BSSVS > 0.5 are shown). Predefined discrete areas are indicated by oval-like shapes and indicated by abbreviations: APE, Apennine peninsula; BAL, Balkan Peninsula and Crete; COR, Corsica and Sardinia; IBE, Iberian peninsula and Morocco; MAD, Madeira. The most probable ancestral area is 'Balkans and Crete'.

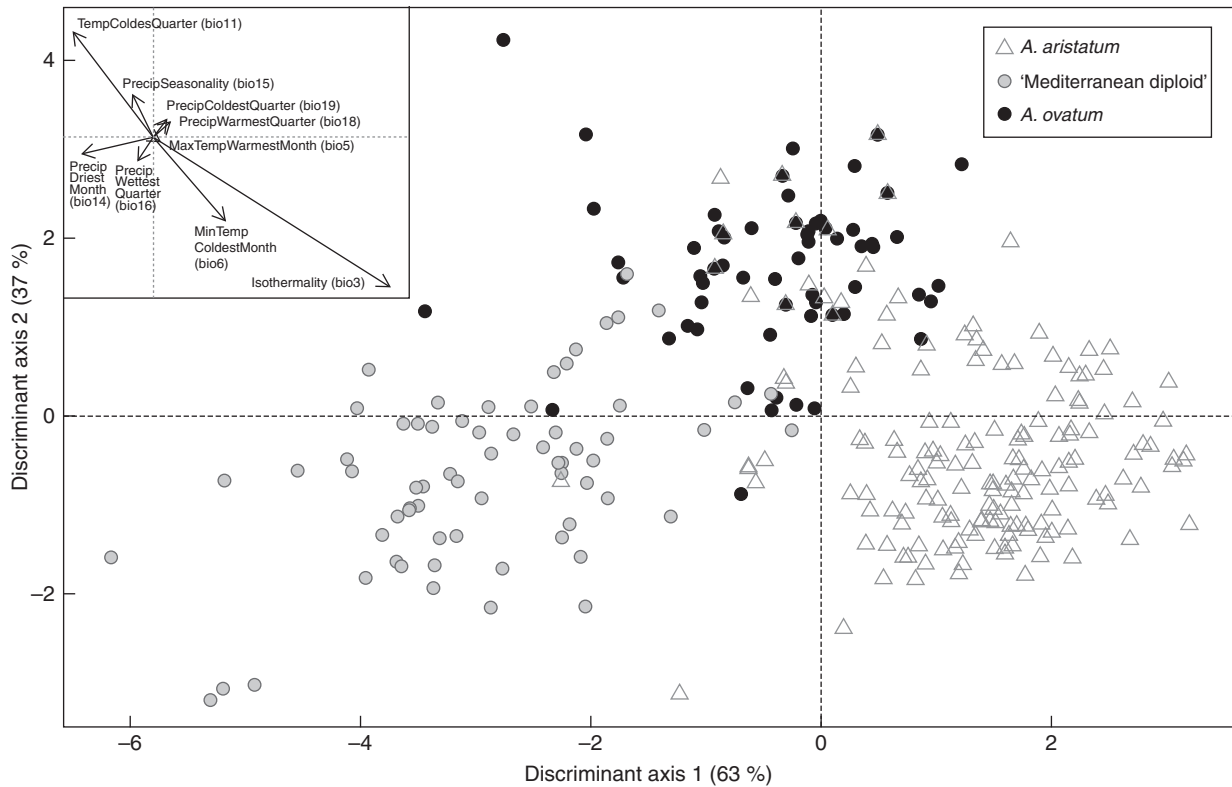


FIG. 5. Linear discriminant analysis of range-wide climatic niche variation of 298 populations of three Mediterranean taxa. Model is based on nine significantly contributing variables selected via stepwise forward selection.

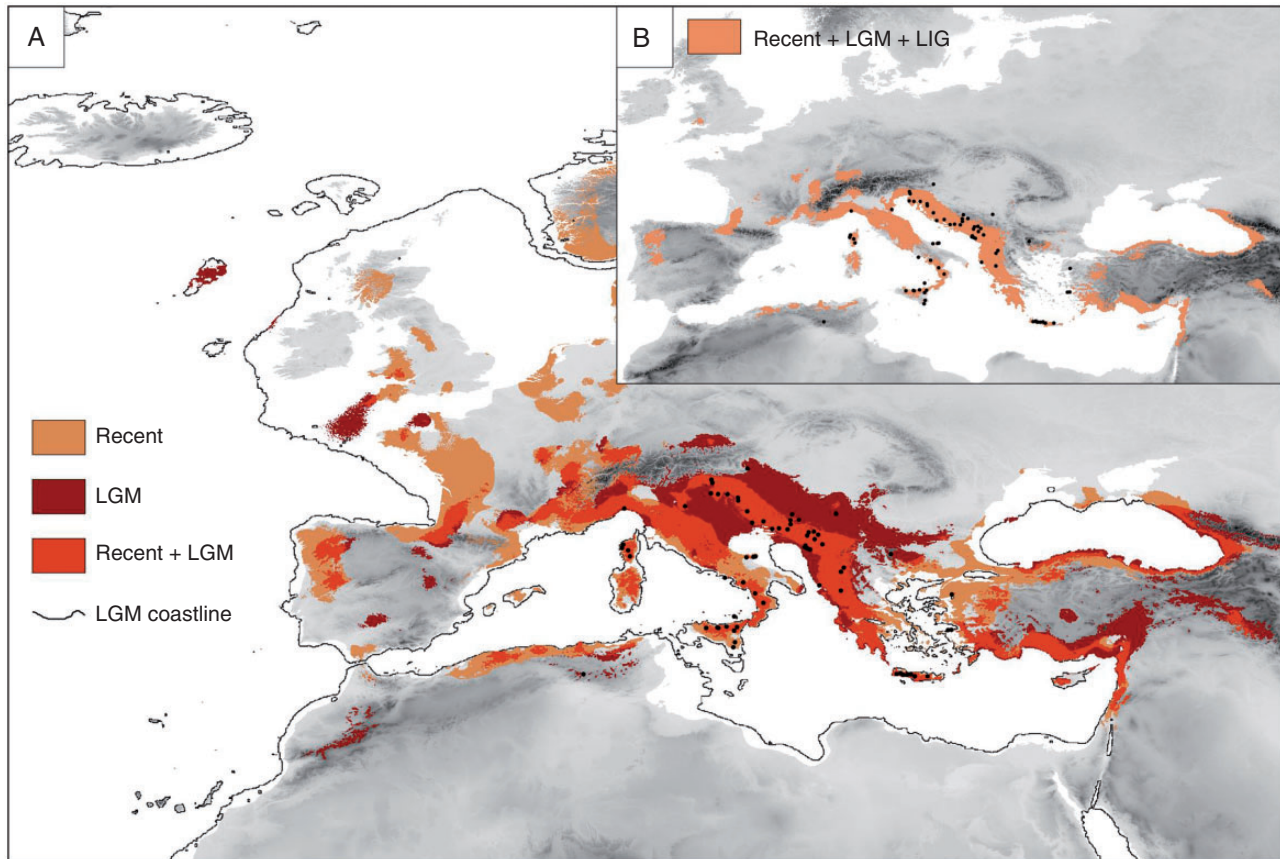


FIG. 6. Distributions of the ‘Mediterranean diploid’ since (A) the Last Glacial Maximum (LGM; approx. 21 000 years ago); and (B) the Last Interglacial (LIG; approx. 120 000 years ago) predicted based on the MIROC climatic model. ‘Recent + LGM’ and ‘Recent + LGM + LIG’ mark areas predicted as suitable in more than one time period. The black line in coastal zones indicates land areas above sea level during the LGM.

coldest month and decreasing precipitation seasonality. Separation of the ‘Mediterranean diploid’ is mainly promoted by precipitation factors (increasing tendency in precipitation of driest month and precipitation of wettest quarter, Fig. S8).

#### *Distributional modelling of the ‘Mediterranean diploid’*

The high average value for the replicates of the operating characteristic curve ( $AUC = 0.970 \pm 0.007$ ) points to a non-random distribution pattern of the ‘Mediterranean diploid’ based on eight Bioclim variables. Modelling of current distribution of the ‘Mediterranean diploid’ points to a potentially wider area than sampled, because of the wider occurrence of suitable conditions (Fig. 6). The Miroc LGM model places the core distribution area in Italy and the Balkan Peninsula (Fig. 6), and the CCSM LGM model indicates suitable conditions not only in this area, but more broadly in a region also including southern Great Britain, France and Turkey (data not shown). Reconstruction of long-standing persistence of suitable conditions over the LIG and the LGM period up to the recent state (Fig. 6) is highly concordant with the current distribution of the ‘Mediterranean diploid’. Both LGM models used identify the identical core area in the central and Eastern Mediterranean (Italy + Balkan Peninsula).

## DISCUSSION

The present study summarizes our findings on the evolution of diploid members of *Anthoxanthum* sect. *Anthoxanthum* in Europe. From the technical point of view, we are operating on the border of two evolutionary levels – below and above the species level. Both levels have their biological specificities, and different approaches are used to analyse them. On this border, special care has to be taken to recognize the amounts of gene flow among populations and species, the presence of ancestral polymorphism, and the historical relationships of populations (Schaal *et al.*, 1998). We combined phylogeographical and phylogenetic approaches to answer questions regarding phylogeny and taxonomy of *Anthoxanthum* sect. *Anthoxanthum*. It enabled us to describe in more detail processes that influenced evolution of the so far undescribed, yet key species, the ‘Mediterranean diploid’, and its closest relatives from the Mediterranean Basin. Results of karyological methods and climatic niche modelling added further insights not obtainable from the genetic data only.

#### *Phylogenetic relationships of diploid taxa from A. sect. Anthoxanthum*

We reconstructed gene trees and a population tree to illustrate the relationships among 14 groups (Table 2) of

populations currently isolated by reproductive, ecological or geographical barriers. This grouping was moreover a finer grouping of the currently recognized taxa of *A. sect. Anthoxanthum*, which enabled us to evaluate current taxonomic treatment of this section. Because of biological reasons (reproductive/ecological/geographical barriers) as well as requirement of the multi-species coalescence model used in our analysis, we assumed that there is no current gene flow among these groups. By testing this assumption *a posteriori* by simulations of plastid data, we showed that differences between plastid gene tree topology and population tree topology can be explained by persistence of ancestral polymorphism (i.e. ILS). However, the same test applied to ITS data detected a signal of recent hybridization/gene flow. In general, the ITS marker is well known for its usefulness in plant systematics as well as for its potential complexity (recombination, presence of multiple ribotypes within single individuals, concerted evolution, etc.) caused by its multi-copy character (e.g. Alvarez and Wendel, 2003). However, in some cases, the complexity of ITS can be a source of valuable information about gene flow and/or migration pathways in the studied group and it is still widely used for untangling relationships even at the infra-specific level (e.g. Comes and Abbott, 2001; Wissemann, 2002). In the case of our data, relationships of ITS ribotypes generally reflect the topology of the inferred population tree, with only a few exceptional cases violating this pattern. As these incongruences can be explained by (recent) migration between neighbouring areas (see below) and thus should not influence the multi-species coalescent model significantly, we consider these results for further interpretation and conclusions.

At the base of the phylogeny stands *Anthoxanthum gracile*, which is relatively dissimilar from the other studied taxa in terms of both its cpDNA haplotype and ITS ribotype. This is fully in accord with previous results based on morphometric analysis (Pimentel *et al.*, 2007a), amplified fragment length polymorphism (AFLP) (Pimentel *et al.*, 2007b) as well as sequence data (Pimentel *et al.*, 2013). Interestingly, *A. gracile* exhibits an rDNA pattern (FISH analyses) shared with some derived members of the group, thus making it the most common type within diploids and supporting the hypothesis of the ‘ancestral *Anthoxanthum* karyotype’. A recently revealed completely different genome size (both 2C- and 1C *x*-value; Chumová *et al.*, 2015) further promotes the idea of *A. gracile* as differing considerably from the other taxa in the genus. All evidence supports the hypothesis of a basal position for this species not only within *sect. Anthoxanthum*, but for the genus as a whole (Pimentel *et al.*, 2013).

The second well-differentiated and supported group is composed of five subgroups of *A. alpinum* (groups of populations from the Alps, Balkans, Caucasus, Carpathians and Scandinavia). *Anthoxanthum alpinum* is a taxon of arctic lowlands and mountain ranges of southern and central Eurasia. Its different position in cladograms is clearly evident even in analysis including polyploids, in which *A. alpinum* is placed in a separate lineage with north-east African *A. nivale* and some specimens of Eurasian *A. odoratum* (Pimentel *et al.*, 2013). Clear separation from other diploid taxa and connection to some polyploids supplies the first evidence in the long-term debate about participation of *A. alpinum* in polyploid formation (e.g. Borrill, 1963; Jones, 1964; Teppner, 1970; Hedberg, 1986;

Chumová *et al.*, 2015). Based on its linkage to at least some polyploid lineages we can conclude that *A. alpinum* played a substantial role in polyploid evolution.

The third group comprises the remaining taxa, two perennial taxa – *A. maderense* and the ‘Mediterranean diploid’, and two annual taxa – *A. aristatum* and *A. ovatum*. In our concept, this group is further divided into eight groups – *A. maderense* forming one group, four groups within the ‘Mediterranean diploid’ (groups of populations from the Balkans, Crete, Corsica and Italy) and three groups of annuals representing the *A. aristatum/ovatum* complex (one group of populations from the central Iberian Peninsula, one group from the southern Iberian Peninsula and one from Corsica). The relationships among these eight groups remain rather unresolved. However, by taking into account particular gene genealogies and the pattern of ancestral polymorphism observed in plastid data (Fig. 2A), we can propose a hypothesis about the relationships of these groups. Our data suggest that annuals from the central Iberian Peninsula are highly differentiated and placed basally to the remaining groups. Their karyotype differs from the ‘ancestral karyotype’ as well as from karyotypes of the other annuals. We can speculate that these populations shared a common ancestor with the remaining Mediterranean populations/taxa but have been separated from them for a long time, perhaps without any further gene flow between them. Furthermore, these populations, forming a quite well-separated lineage, should be assigned to pure *A. aristatum* based on previous phylogenetic studies (Pimentel *et al.*, 2013) and their localization in the Mediterranean area (Valdés, 1973; Fig. S8). Moreover, these plants share the same karyotype as *A. aristatum* from Poland (Drapikowska *et al.*, 2013).

Genetic differentiation of the remaining seven groups (i.e. the third major grouping discussed above, without the group from the central Iberian Peninsula) is further blurred by ancestral polymorphism in the plastid data (Fig. 2). However, the geographically structured ITS genealogy gives us the chance to see some patterns. We can hypothesize that the currently most frequently observed haplotypes (from the red and green haplotype groups) diverged approximately in the same time period, co-occurred in the past and in some areas still co-occur. Red and green ITS ribotypes, in contrast, most probably have been sorted faster and reflect geographical barriers, rather than taxonomy, as annuals and perennials still share ITS ribotypes in Western Europe as well as in Corsica, where both life forms co-occur (Lambinon and Deschatres, 1991; Felber, 1993; Fig. 1). We can only conclude that two groups of annuals (from the southern Iberian Peninsula and Corsica) most probably do not share the MRCA with annuals from the central Iberian Peninsula, but with the ‘Mediterranean diploid’. These findings would support the concept of two distinct species, *A. aristatum* and *A. ovatum*, that differ, among other ways, in their geographical distribution (Valdés, 1973) and quite well-separated climatic niches (Fig. 5).

#### *Ecological differentiation of taxa from A. sect. Anthoxanthum over time*

Based on available evidence, the MRCA of *A. gracile*, as well as *sect. Anthoxanthum*, which we henceforth refer to as

‘Ant<sub>MRC</sub>’, originated in the Miocene approx. 6–8 Mya (95 % HPD 4–8–9–4 Mya) (Pimentel *et al.*, 2013), i.e. at a time of great climatic change from humid or sub-humid summers to summers representative of the summer-dry, Mediterranean-type climate (Quan *et al.*, 2014). According to the contemporary distribution of taxa that speciated first (*A. gracile* – occupying humid gorges and deep valleys, and *A. alpinum* – inhabiting only high mountain ranges in the Mediterranean) adaptation of Ant<sub>MRC</sub> originally to humid or sub-humid climate is very likely. Adaptation to dry climate took place much later and is associated in fact only with the annuals *A. aristatum* and *A. ovatum* (Fig. 5). However, even their MRCA was probably adapted or pre-adapted to humid climate, because the ‘Mediterranean diploid’ and *A. maderense* occupied more humid habitats of the central and Eastern Mediterranean and the Macaronesian island Madeira, respectively.

If the above-described scenario is taken into account, it seems that the first diversification of *Anthoxanthum* diploids was associated with aridification of the Mediterranean Basin in the Miocene (Flower and Kennett, 1994; Zachos *et al.*, 2001). Ant<sub>MRC</sub> at first inhabited a broader area and, probably due to subsequent aridification, was thrust out to relict habitats (leading to the origin of *A. gracile*) or higher altitudes, (yielding the common ancestor of the rest of the taxa, namely *A. alpinum*). Similar scenarios including genetic diversification and speciation related to Miocene climate changes took place in several Mediterranean taxa, for example genus *Odontites* (Gaudeul *et al.*, 2016), genus *Ricotia* (Özüdoğru *et al.*, 2015) and *Smilax aspera* (Chen *et al.*, 2014). In the case of *Anthoxanthum*, the later step of diversification (Fig. 2) is dated to the end of the Pliocene or to the beginning of Pleistocene climate oscillations (median 2 Mya; Suc, 1984) and included separation of the MRCA of mountain type *A. alpinum* from the rest of the taxa. This divergence was crucial for subsequent evolution of polyploids, as *A. alpinum* played a notable role in their origin, both in Eurasia and in Africa (Pimentel *et al.*, 2013). Thus, *A. alpinum* forms a connection between Eurasian and African members of section *Anthoxanthum*. The remaining question is whether the migration route from Eurasia to north-east Africa was used by one or two different diploid taxa (certainly including *A. alpinum*) or by a newly emerged polyploid (or polyploids). However, in both cases such colonization took place quite recently, which is in congruence with some findings of Eurasian–African connections (e.g. Koch *et al.*, 2006; Popp *et al.*, 2008; Inda *et al.*, 2014). The diversification time of the rest of the Mediterranean taxa is still blurry because their separation according to the investigated molecular markers is incomplete. However, it is clear that their common ancestor was able to form the perennial ‘Mediterranean diploid’ in the central and Eastern Mediterranean (see below), to colonize Madeira (*A. maderense*; Teppner, 1998), and finally to adapt to the arid climate and change life strategy (shift to annuals *A. aristatum* and *A. ovatum*; Valdés, 1973).

relatives are greatly blurred by the high portion of shared plastid haplotypes as well as ITS ribotypes. To shed some light on the evolutionary history of the complex of taxa occurring in the Mediterranean, we reconstructed past migration routes of the populations from this complex by employing phylogeographical modelling. Based on two genetic regions with different mutation rates (e.g. Wolfe *et al.*, 1987) and also different rates of sorting of the ancestral alleles (e.g. Comes and Abbott, 2001), we were able to observe two levels of past connections among regions of the current distribution of the complex. However, due to the lack of reliable results of dating analysis for this group, we have to interpret these results without a proper time scale.

The data suggested that the area of ancestral distribution of the ‘Mediterranean diploid’ and its closest relatives was in the Balkans (including Crete), which is highly concordant with the present occurrence of the oldest lineages of *Anthoxanthum* in that area – i.e. *A. gracile* in gorges in Crete and *A. alpinum* in the highest mountain ranges of the Balkan Peninsula. From the ancestral area, most probably two independent migrations to the Iberian Peninsula happened, the first probably giving rise to *A. aristatum* (in our analyses represented by highly differentiated populations of annuals from the central Iberian Peninsula). The second migration to the Iberian Peninsula involved establishment of populations in Africa (Morocco) and Madeira and, as suggested by higher migration rates between Madeira and the Apennines, the source of the migrating population could actually have been the Apennines (perhaps together with the Balkans). The higher migration rate between the Iberian Peninsula (plus Morocco) and Madeira indicates that there was gene flow between the populations in these areas or that some of the populations from one area originated from the other. The strongest bidirectional migration was between the Balkans and Corsica (and Sardinia) and strongest unidirectional flow was from the Apennines to the Balkans. These three migration routes (i.e. Balkans–Corsica in both directions and Apennines–Balkans) are in concordance with the distribution model applied to the ‘Mediterranean diploid’, which suggests the occurrence of contiguous areas of suitable niches spanning the Balkans and Apennines and most probably easily reachable niches in Corsica and Sardinia in the past 120 000 years (Fig. 6). Because these migrations were estimated mainly based on close genetic relationships of green haplotypes from Corsica and green haplotypes from the Balkans that are probably not older than 0.5 Mya (Fig. 3), we believe that this connectivity is supported by past co-occurrence in the continuous area spanning the Balkans, Corsica, Sardinia and the Apennines. The fact that these analyses are based only on two independent markers, however, prevents us from presenting this hypothesis with any certainty. A broader sampling of annual taxa and use of genome-wide data may uncover additional past genetic connections and perhaps even recent or ongoing gene flow among some populations.

#### *Past migration routes of the ‘Mediterranean diploid’ and its closest relatives*

As mentioned above, genetic relationships among groups of populations of the ‘Mediterranean diploid’ and its closest

#### *The ‘Mediterranean diploid’*

Distribution-modelling data on the ‘Mediterranean diploid’ points to very likely long-standing persistence of the taxon in the Mediterranean area, particularly in the central and eastern

parts of this region. Only those locations where suitable conditions have persisted for at least 120 000 years (Fig. 6) host the taxon now. On the other hand, several locations with the same suitability are not occupied (mainly in France and the Iberian Peninsula; Fig. 6). We are convinced that this difference in occupancy was probably caused by differential evolution of the common ancestral taxon of the ‘Mediterranean diploid’ and annuals (*A. aristatum*, *A. ovatum*) in different parts of the Mediterranean Basin. Pursuant to this scenario, those suitable uncolonized places in the Western Mediterranean are mainly occupied by *A. aristatum* now (Valdés, 1973). The idea of long-standing stability of the ‘Mediterranean diploid’ populations is further supported by its sharing of the ‘ancestral *Anthoxanthum* karyotype’ (Fig. 2A, Fig. S3). Moreover, as we hypothesized above, the Balkan and adjacent regions are considered as a source area for further evolution of Mediterranean taxa, and the scenario of long persistence of the most recent ancestor of these taxa in this area appears likely. This finding is in accordance with various evidence of a crucial role of the Balkan Peninsula in the evolution of taxa that originated in the Miocene (Frajman and Schneeweiss, 2009; Frajman et al., 2009) or later (Surina et al., 2014; Petrova et al., 2015), making this area a European biodiversity hotspot (Médail and Diadema, 2009; Hewitt, 2011).

#### Attributes of the MRCA of section *Anthoxanthum*

Various lines of evidence enable us to speculate about the attributes of ‘Ant<sub>MRCA</sub>’, the MRCA of section *Anthoxanthum*. As we hypothesized above, its adaptation to a humid climate is very likely, because most diploid *Anthoxanthum* species (including both basal and derived species) prefer such habitats. It seems that adaptation to a Mediterranean climate with dry summer was secondary, restricted mostly to annuals and occurring more recently (in accordance with the establishment of a truly Mediterranean climate; see e.g. Thompson, 2005; Tomasello et al., 2015). The determination of the basal life strategy (annual vs. perennial) is quite tricky, but several clues suggest perenniality as the more likely basal state. Namely, the annual life strategy is a common adaptation to arid climate (e.g. Grime, 1977; Evans et al., 2005), but Ant<sub>MRCA</sub> is presumed to have been adapted for more humid habitats. Moreover, perenniality is thought to be a contributing factor to: (1) better adaptation of many grass genera to various climatic and edaphic conditions; and (2) formation of species-rich polyploid complexes (Röser, 1997). Finally, sharing of the same type of karyotype by basal *A. gracile* and two other taxa from the most derived group underlies our inference regarding the existence of an ‘ancestral *Anthoxanthum* karyotype’, which should be shared also by Ant<sub>MRCA</sub>.

## CONCLUSIONS

Our examination provided comprehensive insight into the evolution of diploid members of genus *Anthoxanthum* sect. *Anthoxanthum* and brought new understanding of their origin, time of divergence, colonization routes and niche differentiation. We corroborated the hypothesis of a Miocene origin and establishment of the basal lineage leading to evolution of *A.*

*gracile*. Other divergences are much younger, having occurred during the Pleistocene period. Based on the current occurrence and habitat preferences of diploids we further hypothesized prior adaptation of the MRCA of diploids to humid or sub-humid habitats. This provided the basis for theorizing that only a relatively recent adaptation to true Mediterranean climate underlay the shift in life strategy from perennial to annual. In sum, we assembled evidence that the Mediterranean area is the cradle of *Anthoxanthum* sect. *Anthoxanthum* and that climatic changes starting in the Miocene, leading to increased aridity and formation of the true Mediterranean climate, were the main factors promoting taxa differentiation. It seems, however, that speciation of a substantial part of the diploid taxa is still an ongoing process. This is well documented by unclear differentiation in molecular markers but clear separation of climatic niches. Phylogeographical modelling suggested the source area of diploid diversity to be the Balkans (including Crete) and Apennines, corresponding to long-standing suitable conditions for persistence of the rediscovered ‘Mediterranean diploid’ in this region. Moreover, its sharing of the ‘ancestral *Anthoxanthum* karyotype’ further supports the linkage of the ‘Mediterranean diploid’ to the MRCA of other diploid taxa and corroborates the importance of its distribution area (i.e. Central and Eastern Mediterranean) for evolution of the entire genus.

## SUPPLEMENTARY DATA

Supplementary data are available online at <https://academic.oup.com/aob> and consist of the following. Table S1: list of analysed *Anthoxanthum* populations. Table S2: asymmetrical migration rate matrix estimated by phylogeographical modelling with BASTA based on the plastid dataset, ITS dataset, and combined plastid and ITS data. Text S1: detailed description of the approach used for testing for the presence of gene flow among species/lineages. Figure S1: Map of localities of three Mediterranean *Anthoxanthum* taxa (*A. aristatum*, *A. ovatum*, ‘Mediterranean diploid’) that serve as reference points for appropriate climatic data from the WorldClim database. Figure S2: maximum clade credibility tree from the Bayesian analysis of *Anthoxanthum* species (cpDNA dataset). Figure S3: maximum clade credibility tree from the Bayesian analysis of *Anthoxanthum* species (ITS dataset) with highlighted branches indicating possible origin of the ‘ancestral *Anthoxanthum* karyotype’ (rDNA FISH analyses). Figure S4: NeighbourNet analysis of *Anthoxanthum* species based on (A) the cpDNA dataset, (B) ITS dataset and (C) concatenated cpDNA and ITS dataset. Figure S5: multi-species coalescent population tree based on cpDNA and ITS data computed in \*BEAST. Figure S6: DAPI-stained mitotic chromosome spreads of *Anthoxanthum* species. Figure S7: ancestral area reconstruction using asymmetrical BASTA phylogeographical model based on ITS dataset projected to maximum clade credibility tree from the Bayesian analysis. Figure S8: box-and-whiskers plots of nine bioclimatic variables that significantly contribute to modelling of climate niches of three Mediterranean *Anthoxanthum* taxa – *A. aristatum*, *A. ovatum* and ‘Mediterranean diploid’.

## ACKNOWLEDGEMENTS

We are grateful to R. Elven, H. Chudáčková, J. Chrtek, F. Krahulec, A. Krahulcová, F. Rooks, P. Růžičková, P. Silveira, R. Sudová, M. Štech, T. Urfus, P. Vít and J. Zahradníček for their help during fieldwork and/or providing some *Anthoxanthum* samples and to J. Rosenthal for kindly improving the English text. The research was financially supported by Charles University projects [GA UK 66109, 541413]. Additional support was supplied by the Academy of Sciences of the Czech Republic [RVO 67985939] and the Czech Science Foundation [P501/12/G090].

## LITERATURE CITED

- Alvarez I, Wendel JF. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29**: 417–434.
- Avise JC. 1994. *Molecular markers, natural history and evolution*. New York: Springer US.
- Borrill M. 1963. Experimental studies of evolution in *Anthoxanthum* (Gramineae). *Genetica* **34**: 183–210.
- Briquet J. 1910. *Anthoxanthum* L. In: Briquet J, ed. *Prodrome de la Corse*. Genève, Bale, Lyon: Georg & Co, Libraires – éditeurs, 72–74.
- Bryant D, Moulton V. 2004. Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* **21**: 255–265.
- Chen C, Qi Z-C, Xu X-H, et al. 2014. Understanding the formation of Mediterranean–African–Asian disjunctions: evidence for Miocene climate-driven vicariance and recent long-distance dispersal in the Tertiary relict *Smilax aspera* (Smilacaceae). *New Phytologist* **204**: 243–255.
- Chumová Z, Krejčíková J, Mandáková T, Suda J, Trávníček P. 2015. Evolutionary and taxonomic implications of variation in nuclear genome size: lesson from the grass genus *Anthoxanthum* (Poaceae). *PLoS One* **10**: e0133748.
- Chumová Z, Mandáková T, Trávníček P. 2016. Are B-chromosomes responsible for the extraordinary genome size variation in selected *Anthoxanthum* annuals? *Plant Systematics and Evolution* **302**: 731–738.
- Clayton WD, Govaerts R, Harman KT, Williamson H, Vorontsova M. 2016. World checklist of Poaceae. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; <http://apps.kew.org/wcsp/> (last accessed 6 December 2016).
- Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657–1659.
- Collins WD, Bitz CM, Blackmon ML, et al. 2006. The community climate system model version 3 (CCSM3). *Journal of Climate* **19**: 2122–2143.
- Comes HP, Abbott RJ. 2001. Molecular phylogeography, reticulation, and lineage sorting in Mediterranean *Senecio* sect. *Senecio* (Asteraceae). *Evolution* **55**: 1943–1962.
- De Maio N, Wu C-H, O'Reilly KM, Wilson D. 2015. New routes to phylogeography: a Bayesian structured coalescent approximation. *PLoS Genetics* **11**: e1005421.
- Díaz-Pérez AJ, Sharifi-Tahrani M, Inda LA, Catalán P. 2014. Polyphyly, gene-duplication and extensive allopolyploidy framed the evolution of the ephemeral *Vulpia* grasses and other fine-leaved Loliinae (Poaceae). *Molecular Phylogenetics and Evolution* **79**: 92–105.
- Drapikowska M, Susek K, Hasterok R, Szkudlarz P, Celka Z, Jackowiak B. 2013. Variability of stomata and 45S and 5S rDNAs loci characteristics in two species of *Anthoxanthum* genus: *A. aristatum* and *A. odoratum* (Poaceae). *Acta Biologica Hungarica* **64**: 352–363.
- Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**: 214.
- Elith J, Graham CH, Anderson RP, et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**: 129–151.
- Evans MEK, Hearn DJ, Hahn WJ, Spangle JM, Venable DL. 2005. Climate and life-history evolution in evening primroses (*Oenothera*, Onagraceae): a phylogenetic comparative analysis. *Evolution* **59**: 1914–1927.
- Felber F. 1987. *Contribution à l'étude phytogéographique, biosystématique et expérimentale du complexe polyploïde Anthoxanthum odoratum L. s. lat.* PhD Thesis, University of Neuchâtel, Switzerland.
- Felber F. 1993. Présence de trois taxon pérennes d'*Anthoxanthum* L. en Corse. *Candollea* **48**: 582–591.
- Feliner GN. 2014. Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics* **16**: 265–278.
- Flower BP, Kennett JP. 1994. The middle Miocene climatic transition: East Antarctic ice sheet development, deep ocean circulation and global carbon cycling. *Palaeogeography, Palaeoclimatology, Palaeoecology* **108**: 537–555.
- Frajman B, Schneeweiss GM. 2009. A campanulaceous fate: the Albanian sten-endemic *Asyneuma comosiforme* in fact belongs to isophyllous *Campanula*. *Systematic Botany* **34**: 595–601.
- Frajman B, Eggens F, Oxelman B. 2009. Hybrid origins and homoploid reticulate evolution within *Heliosperma* (Sileneae, Caryophyllaceae) – A multigene phylogenetic approach with relative dating. *Systematic Biology* **58**: 328–345.
- Gaudeul M, Vêla E, Rouhan G. 2016. Eastward colonization of the Mediterranean Basin by two geographically structured clades: the case of *Odontites* Ludw. (Orobanchaceae). *Molecular Phylogenetics and Evolution* **96**: 140–149.
- Grime JP. 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *The American Naturalist* **111**: 1169–1194.
- Hall T. 2004. BioEdit, Biological sequence alignment editor for Win95/98/NT/2K/XP, version 7.0.4.1. <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>.
- Hasumi H, Emori S. 2004. *K1-coupled GCM (MIROC) description. K1 Technical report No. 1*. Tokyo: Center for Climate System Research.
- Hedberg I. 1967. Cytotaxonomic studies on *Anthoxanthum odoratum* L. s. lat. 2. Investigations of some Swedish and of a few Swiss population samples. *Symbolae Botanicae Upsalienses* **18**: 5–97.
- Hedberg I. 1986. The genesis of tetraploid *Anthoxanthum odoratum*. *Symbolae Botanicae Upsalienses* **27**: 147–154.
- Hedberg I. 1990. Morphological, cytotaxonomic and evolutionary studies in *Anthoxanthum odoratum* L. s. lat. – a critical review. *Sommerfeltia* **11**: 97–107.
- Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multi-locus data. *Molecular Biology and Evolution* **27**: 570–580.
- Hewitt G. 2011. Mediterranean peninsulas: the evolution of hotspots. In: Zachos FE, Habel JC, eds. *Biodiversity hotspots*. Berlin: Springer, 123–147.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.
- Hijmans RJ, Eten J van, Cheng J, et al. 2016. Raster: geographic data analysis and modelling. <https://cran.r-project.org/web/packages/raster/index.html> (last accessed 2 June 2016).
- Holderregger R, Abbott RJ. 2003. Phylogeography of the Arctic-Alpine *Saxifraga oppositifolia* (Saxifragaceae) and some related taxa based on cpDNA and ITS sequence variation. *American Journal of Botany* **90**: 931–936.
- Hsiao C, Chatterton N, Asay K, Jensen K. 1995. Molecular phylogeny of the Pooideae (Poaceae) based on nuclear rDNA (ITS) sequences. *Theoretical and Applied Genetics* **90**: 389–398.
- Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.
- Hunt HV, Badakshi F, Romanova O, Howe CJ, Jones MK, Heslop-Harrison JSP. 2014. Reticulate evolution in *Panicum* (Poaceae): the origin of tetraploid broomcorn millet, *P. miliaceum*. *Journal of Experimental Botany* **65**: 3165–3175.
- Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**: 254–267.
- Inda LA, Sanmartín I, Buerki S, Catalán P. 2014. Mediterranean origin and Miocene-Holocene Old World diversification of meadow fescues and ryegrasses (*Festuca* subgenus *Schedonorus* and *Lolium*). *Journal of Biogeography* **41**: 600–614.
- Joly S. 2012. JML: testing hybridization from species trees. *Molecular Ecology Resources* **12**: 179–184.
- Joly S, McLenachan PA, Lockhart PJ. 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting. *American Naturalist* **174**: E54–E70.
- Jones K. 1964. Chromosomes and the origin of *Anthoxanthum odoratum* L. *Chromosoma* **15**: 248–274.
- Knowles LL, Kubatko LS (eds). 2010. *Estimating species trees: practical and theoretical aspects*. Oxford: Wiley-Blackwell.



- Koch MA, Kiefer C, Ehrich D, Vogel J, Brochmann C, Mummenhoff K. 2006. Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology* 15: 825–839.
- Koutecký P. 2015. MorphoTools: a set of R functions for morphometric analysis. *Plant Systematics and Evolution* 301: 1115–1121.
- Lambinon J, Deschatres R. 1991. Les *Anthoxanthum* annuels de Corse. *Candollea* 46: 217–221.
- Leigh JW, Susko E, Baumgartner M, Roger AJ. 2008. Testing congruence in phylogenomic analysis. *Systematic Biology* 57: 104–115.
- Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian phylogeography finds its roots. *PLoS Computational Biology* 5: e1000520.
- Liu CR, Berry PM, Dawson TP, Pearson RG. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28: 385–393.
- Lowe A, Munro R, Samuel S, Cottrell J. 2004. The utility and limitations of chloroplast DNA analysis for identifying native British oak stands and for guiding replanting strategy. *Forestry* 77: 335–347.
- Mamidi D, Rossi M, Moghaddam SM, et al. 2013. Demographic factors shaped diversity in the two gene pools of wild common bean *Phaseolus vulgaris* L. *Heredity* 110: 267–276.
- Mandáková T, Lysak MA. 2016. Painting of *Arabidopsis* chromosomes with chromosome-specific BAC clones. In: *Current Protocols in Plant Biology*, published online. Chichester: John Wiley & Sons, Inc.
- Médail F, Diadema K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography* 36: 1333–1345.
- Neteler M, Bowman MH, Landa M, Metz M. 2012. Grass GIS: a multi-purpose open source GIS. *Environmental Modelling & Software* 31: 124–130.
- Nylander JA, Wilgenbusch JC, Warren DL, Swofford DL. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24: 581–583.
- Otto-Bliesner BL, Marsha SJ, Overpeck JT, Miller GH, Hu AX. 2006. Simulating arctic climate warmth and icefield retreat in the last interglaciation. *Science* 311: 1751–1753.
- Özudogru B, Akaydin G, Erik S, Al-Shehbaz IA, Mummenhoff K. 2015. Phylogeny, diversification and biogeographic implications of the eastern Mediterranean endemic genus *Ricotia* (Brassicaceae). *Taxon* 64: 727–740.
- Peterson AT, Papes M, Soberon J. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling* 213: 63–72.
- Petrova G, Moyankova D, Nishii K, et al. 2015. The European paleoendemic *Haberlea rhodopensis* (Gesneriaceae) has an Oligocene origin and a Pleistocene diversification and occurs in a long-persisting refugial area in Southeastern Europe. *International Journal of Plant Sciences* 176: 499–514.
- Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259.
- Pimentel M, Sahuquillo E. 2003. An approach to the study of morphological relationships among the sweet vernal grass (*Anthoxanthum* L. Poaceae, Pooideae) in the Iberian Peninsula. *Bocconea* 16: 731–737.
- Pimentel M, Sahuquillo E. 2008. Relationships between the close congeners *Anthoxanthum odoratum* and *A. alpinum* (Poaceae, Pooideae) assessed by morphological and molecular methods. *Botanical Journal of the Linnean Society* 156: 237–252.
- Pimentel MP, Perez GE, Balbuena E. 2007a. European sweet vernal grasses (*Anthoxanthum*: Poaceae, Pooideae, Aveneae): a morphometric taxonomical approach. *Systematic Botany* 32: 43–59.
- Pimentel M, Sahuquillo E, Catalan P. 2007b. Genetic diversity and spatial correlation patterns unravel the biogeographical history of the European sweet vernal grasses (*Anthoxanthum* L., Poaceae). *Molecular Phylogenetics and Evolution* 44: 667–684.
- Pimentel M, Catalan P, Sahuquillo E. 2010. Morphological and molecular taxonomy of the annual diploids *Anthoxanthum aristatum* and *A. ovatum* (Poaceae) in the Iberian Peninsula. Evidence of introgression in natural populations. *Botanical Journal of the Linnean Society* 164: 53–71.
- Pimentel M, Sahuquillo E, Torrecilla Z, Popp M, Catalan P, Brochmann C. 2013. Hybridization and long-distance colonization at different time scales: towards resolution of long-term controversies in the sweet vernal grasses (*Anthoxanthum*). *Annals of Botany* 112: 1015–1030.
- Popp M, Erixon P, Eggens F, Oxelman B. 2005. Origin and evolution of a circumpolar polyploid species complex in *Silene* (Caryophyllaceae) inferred from low copy nuclear RNA polymerase introns, rDNA, and chloroplast DNA. *Systematic Botany* 30: 302–313.
- Popp M, Gizaw A, Nemomissa S, Suda J, Brochmann C. 2008. Colonization and diversification in the African ‘sky islands’ by Eurasian *Lychnis* L. (Caryophyllaceae). *Journal of Biogeography* 35: 1016–1029.
- Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* 25: 1253–1256.
- Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53: 793–808.
- Quan C, Liu Y-S, Tang H, Utescher T. 2014. Miocene shift of European atmospheric circulation from trade wind to westerlies. *Scientific Reports* 4: 5660.
- Rojas-Andres BM, Albach DC, Montserrat Martinez-Ortega M. 2015. Exploring the intricate evolutionary history of the diploid-polyploid complex *Veronica* subsection *Pentasepalae* (Plantaginaceae). *Botanical Journal of the Linnean Society* 179: 670–692.
- Röser M. 1997. Patterns of diversification in Mediterranean oat grasses (Poaceae: Aveneae). *Lagascalia* 19: 101–120.
- Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA. 1998. Phylogeographic studies in plants: problems and prospects. *Molecular Ecology* 7: 465–474.
- Schwarz G. 1978. Estimating the dimension of a model. *The Annals of Statistics* 6: 461–464.
- Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* 94: 275–288.
- Siqueiros-Delgado EM, Ainouche M, Columbus JT, Ainouche A. 2013. Phylogeny of the *Bouteloua curtipendula* complex (Poaceae: Chloridoideae) based on nuclear ribosomal and plastid DNA sequences from diploid taxa. *Systematic Botany* 38: 379–389.
- Stapf O. 1899. *Anthoxanthum*. In: Harvey WH, Sonder OW, eds. *Flora Capensis*. Dublin: Hodges, Smith and Co., 465–468.
- Suc J-P. 1984. Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature* 307: 429–432.
- Sun G. 2014. Molecular phylogeny revealed complex evolutionary process in *Elymus* species. *Journal of Systematics and Evolution* 52: 706–711.
- Surina B, Pfanzelt S, Einzmann HJR, Albach DC. 2014. Bridging the Alps and the Middle East: evolution, phylogeny and systematics of the genus *Wulfenia* (Plantaginaceae). *Taxon* 63: 843–858.
- Swofford DL. 2002. *Phylogenetic analysis using parsimony (\*and other methods)*. Sunderland, MA: Sinauer Associates.
- Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of 3 noncoding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.
- Teppner H. 1970. Caryotypes of European, perennial species of Gramineae genus *Anthoxanthum*. *Osterreichische Botanische Zeitschrift* 118: 280–292.
- Teppner H. 1998. *Anthoxanthum maderense* spec. nova and *A. odoratum* (Poaceae-Aveneae) from Madeira and their chromosome morphology. *Phyton-Annales Rei Botanicae* 38: 307–321.
- Thompson JD. 2005. *Plant evolution in the Mediterranean*. Oxford: Oxford University Press.
- Tomasello S, Álvarez I, Vargas P, Oberprieler C. 2015. Is the extremely rare Iberian endemic plant species *Castrilanthemum debeauxii* (Compositae, Anthemideae) a ‘living fossil’? Evidence from a multi-locus species tree reconstruction. *Molecular Phylogenetics and Evolution* 82: 118–130.
- Triplett JK, Clark LG, Fisher AE, Wen J. 2014. Independent allopolyploidization events preceded speciation in the temperate and tropical woody bamboos. *New Phytologist* 204: 66–73.
- Tutin T. 1980. *Anthoxanthum* L. In: Tutin T, Heywood V, Burges N, et al., eds. *Flora Europaea*, Vol. 5. Cambridge: Cambridge University Press.
- Valdés B. 1973. Revisión de las especies anuales del género *Anthoxanthum* (Graminae). *Lagascalia* 3: 99–141.
- Watson L, Dallwitz MJ. 1992. *The grass genera of the world*. Wallingford: CAB International.

- Winkler M, Tribsch A, Schneeweiss GM, et al. 2012.** Tales of the unexpected: phylogeography of the arctic-alpine model plant *Saxifraga oppositifolia* (Saxifragaceae) revisited. *Molecular Ecology* **21**: 4618–4630.
- Wissemann V. 2002.** Molecular evidence for allopolyploid origin of the *Rosa canina*-complex (Rosaceae, Rosoideae). *Journal of Applied Botany* **76**: 176–178.
- Wolfe KH, Li WH, Sharp PM. 1987.** Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 9054–9058.
- Wölk A, Winterfeld G, Röser M. 2015.** Genome evolution in a Mediterranean species complex: phylogeny and cytogenetics of *Helictotrichon* (Poaceae) allopolyploids based on nuclear DNA sequences (rDNA, topoisomerase gene) and FISH. *Systematics and Biodiversity* **13**: 326–345.
- Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001.** Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**: 686–693.
- Záveská E, Fér T, Šída O, Krak K, Marhold K, Leong-Skornickova J. 2012.** Phylogeny of *Curcuma* (Zingiberaceae) based on plastid and nuclear sequences: proposal of the new subgenus *Ecomata*. *Taxon* **61**: 747–763.
- Zuo H, Wu P, Wu D, Sun G. 2015.** Origin and reticulate evolutionary process of wheatgrass *Elymus trachycaulus* (Triticeae: Poaceae). *PLoS One* **10**: e0125417.