Flora 260 (2019) 151477

Contents lists available at ScienceDirect

Flora

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

Highlighted Student Research

The Eurasian steppe belt in time and space: Phylogeny and historical biogeography of the false flax (*Camelina* Crantz, Camelineae, Brassicaceae)

Anže Žerdoner Čalasan[a,*], Alexey P. Seregin[b], Herbert Hurka[a], Nathaniel P. Hofford[c], Barbara Neuffer[a]

[a] Department 5: Biology/Chemistry, Botany, University of Osnabruck, Barbarastraße 11, 49076 Osnabruck, Lower Saxony, Germany
[b] Herbarium, Department of Geobotany, Faculty of Biology, M.V. Lomonosov Moscow State University, 119991 Moscow, Russia
[c] Department of Evolution, Ecology & Organismal Biology, The Ohio State University, Columbus, Ohio, 43210, United States

ARTICLE INFO

Edited by: Karsten Wesche

Keywords:
Florogenesis
Eurasian steppe
Dated phylogeny
Ancestral area reconstruction
Camelina
Brassicaceae

ABSTRACT

Stretching 8000 km from the Pannonian basin and the Danube delta in the West to the Manchuria region in the East and reaching up to 1000 km in width, the Eurasian steppe belt is the vastest steppe region worldwide. However, our knowledge about the temporal and spatial patterns of floral origin and evolution of the Eurasian Steppe is limited and inconclusive. Case studies on typical steppe flora may help us close such gaps. The study subject of this project was *Camelina* – a taxon which occupies open dry habitats in temperate zones of Eurasia. To infer the evolutionary history of this genus, maximum likelihood optimisation in RAxML and Bayesian Inference approach were carried out, based on the nuclear external transcribed spacer region. Furthermore, we performed a secondarily calibrated time estimation analysis using Bayesian optimisation in BEAST to infer potential influence of climatic shifts and paleogeographic events on the distribution patterns of *Camelina* and carried out an ancestral area reconstruction analysis using a Bayesian Binary Method. Our study resulted in a well-supported phylogeny that corresponds with the species morphology and uncovered several genetically distinct inter- and intraspecific lineages which appear to correlate geographically. Time divergence estimation argue for the diversification of *Camelina* to have taken place in the Middle East around the transition from Pliocene to Pleistocene (3-2 mya), and its historical biogeography to have been under a strong influence of several glacial periods and their palaeoclimatic and palaeoenvironmental consequences. Its young age also explains the subtle morphological character differences among species and high interspecific hybridisation potential. We further discuss the rediscovery of wild *Camelina sativa* populations and propose the external transcribed spacer as a ribotype identifying region for young and rapidly evolving core eudicot lineages.

1. Introduction

The Eurasian steppe belt stretches an impressive 8000 km from the Danube basin in Romania all the way to northeast China and can be up to 1000 km wide (Lavrenko, 1969). Despite being the largest grassland region in the world, little is known about the florogenesis (= floral origin and evolution) of this belt. It has been argued that the Eurasian steppe developed during the early Miocene around 20 million years ago (mya) in Central Asia (Jiang and Ding, 2009; Strömberg, 2011), but there is accumulating evidence for an earlier origin (Gomes Rodrigues et al., 2012; Hurka et al., 2019; Mai, 1995; Quan et al., 2014). Regardless of the exact age, the onset and development of the Eurasian steppe belt was under strong influence of the worldwide Palaeogene cooling trend interrupted by warming periods and development of the Asian monsoon system. Together with the retreat of Paratethys and the uplift of the Tibetan Plateau, early Oligocene disappearance of the Turgai Strait played a pivotal role in the initial aridification of Central Asia (Hurka et al., 2019 and references therein). However, it remains unclear to what extent each of these features contributed to the climate-landscape history of the Eurasian steppes.

Earliest records of steppe climate and vegetation for the East European Plate are from the late Miocene time, less than 9 mya (Bruch et al., 2011; Ivanov et al., 2011; Jacobs et al., 1999; Velichko, 1999). The steppe ecoregion was strongly influenced by past climatic shifts. One of the most prominent transitions was the Last Glacial Maximum followed by Late Glacial climate warming, in which the vegetation went through dramatic changes caused by extensive flora die out (Binney et al., 2017; Frenzel, 1968; Tarasov et al., 2000). Megafaunal

* Corresponding author.
E-mail address: Anze.ZerdonerCalasan@biologie.uni-osnabrueck.de (A. Žerdoner Čalasan).

https://doi.org/10.1016/j.flora.2019.151477
Received 28 June 2019; Received in revised form 13 September 2019; Accepted 25 September 2019
Available online 08 October 2019

extinctions on local and global levels at the end of the late Pleistocene and early Holocene took place (Barnosky, 2004; Bartlett et al., 2016; Fiedel, 2009; Orlova et al., 2004; Owen-Smith, 1987), resulting in a sudden lack of herbivores likely influencing the development of the Eurasians grasslands (Strömberg, 2011). Recent studies argue that the decline of large herbivores induces major alterations in landscape structure and ecosystem function (Bakker et al., 2016; Sandom et al., 2014). Furthermore, during glacial and interglacial cycles, continuous expansions, contractions, and latitudinal/longitudinal range shifts of the Eurasian steppe belt have been documented (Frenzel, 1968; reviewed in Hurka et al., 2019). These were caused by transgressive-regressive events of the Caspian Sea and ice damming of the Siberian river systems (Arkhipov et al., 1999; Mamedov, 1997; Svitoch, 2012; Tudryn et al., 2013).

The genetic footprint of typical dry-adapted steppe species mirrors these changes and allows for detailed studies of the Eurasian steppe belt florogenesis in time and space (Franzke et al., 2004; Friesen et al., 2016; Herden et al., 2016; Volkova et al., 2017). Brassicaceae is one of the biggest and highly diverse plant families of the steppes (Bone et al., 2015), whose representatives have diversified mostly in cool and dry environments around the Eocene-Oligocene boundary (Hohmann et al., 2015). The tribus Camelineae DC. has diversified in the late Miocene (Guo et al., 2017) and is one of the most comprehensively studied Brassicaceae tribes, including important model systems such as *Arabidopsis* and *Capsella* and economically important *Camelina* (Koch et al., 2018).

False flax (*Camelina*, Camelineae, Brassicaceae) is currently comprised of nine species—*Camelina alpkoyensis* Yild., *Camelina alyssum* (Mill.) Thell., *Camelina anomala* Boiss. & Hausskn. ex Boiss., *Camelina hispida* Boiss., *Camelina laxa* C.A. Mey., *Camelina microcarpa* Andrz. ex DC., *Camelina neglecta* J.Brock, Mandáková, Lysak & Al-Shehbaz, *Camelina rumelica* Velen. and *Camelina sativa* (L.) Crantz (Al-Shehbaz, 2012; Brock et al., 2019). They are annual or biennial herbs, presently growing mostly in open and disturbed habitats, including pastures, dry grasslands and along roadsides and railways (Francis and Warwick, 2009). While *C. alyssum*, *C. microcarpa* and *C. sativa* have been introduced to the New World, the rest of the species show a more restricted distribution, not expanding further than the Western Asiatic part of the Irano-Turanian floristic region (sensu Takhtajan, 1986). The latter region is also speculated to represent the area of origin not only of *Camelina*, but of the whole Brassicaceae family (Hedge, 1976). Interspecific and intergeneric hybridisation is a common phenomenon in *Camelina* (Julié-Galau et al., 2014; Séguin-Swartz et al., 2013) and together with subtle morphological character differences between species and numerous cytotypes, these make the taxonomy of *Camelina* notoriously difficult (Brock et al., 2018). Additionally, several contradictory cytotypes have been reported for this species group (Francis and Warwick, 2009; Koch et al., 2018; Martin et al., 2017), possibly due to a high hybridisation potential which further confounds accurate taxonomic inference.

Fast-evolving rRNA operon regions such as the internal transcribed spacers 1 and 2 (ITS) (Schultz, 2005; Suh et al., 1993) have failed to delimitate the species in question, further preventing the recovery of true evolutionary relationships between the species (Brock et al., 2018). Until now, no study has investigated other nuclear regions that could be useful as a barcode for species level taxonomic identification within the *Camelina* clade. Species delineation and their evolutionary relationships have been tackled successfully in a recent study using a double digest restriction site associated sequencing (ddRADSeq) approach, recovering distinct and highly supported clades that were congruent with morphologically assigned species names (Brock et al., 2018). However, not all *Camelina* species were taken into account. Had all the known species been included and sampled across their entire geographic range, the topology of the phylogenetic tree might be different.

Identifying plant species with DNA-barcodes is already a well-established methodology that cost effectively promotes biodiversity assessment, conservation, ecological and life history studies of plants (Kress et al., 2005). As such, it is a widely used tool despite possibly being confounded by plant hybridisation, asexual reproduction, genome duplications and incomplete lineage sorting (Fazekas et al., 2009; Stebbins, 1950). Fast and efficient species determination of *Camelina* is of practical nature due to its agricultural implications and hybridisation potential. For example, important oilseed crop species *C. sativa* is limited by lack of certain desirable agronomic, yield and oil quality traits (Iskandarov et al., 2014). Recovery of potential wild genetic resources would thus come in handy to enhance its breeding pool and potentially improve these inhibiting characteristics. However, because of its young age and supposed rapid speciation, finding an informative barcode for *Camelina* is likely a difficult task.

The overall aim of our studies is to place the development of the flora of the Eurasian steppe belt into time and space. To achieve this goal, case studies are needed. Here, we carried out a phylogenetic and historical biogeographic reconstruction of the *Camelina* genus and consequently characterised the spatial and temporal patterns of this group within the florogenesis of the Eurasian steppe belt. Furthermore, we demonstrated the ribotype identification applicability of a well-known nuclear coding region, the external transcribed spacer (ETS), for the *Camelina* group and investigated the congruency between morphological and genetic species delimitation methods.

2. Materials and methods

2.1. Taxon sampling and distribution range surveys

We compiled a taxon sample covering the whole distribution area and all the known species of *Camelina* (also see Discussion). Distribution maps were put together primarily using geographical range information from the literature (Vassilczenko, 1939; Nikiforova, 1994; Meusel et al., 1965) and our own field data, and only secondarily using online databases (Plants of the World Online, http://www.plantsoftheworldonline.org/), whose entries were critically assessed before being included in the survey. Plant material was obtained from specimens deposited in HBG, JE, MW, NSK and OSBU herbaria (Fig. S1). In addition, several accessions collected as seeds during field trips have been sowed and grown in the greenhouse of the University of Osnabrück. In some cases, fresh leaves were collected in the field, dried in silica gel and used directly for DNA isolation (for the taxon list, see Fig. S1).

Currently, the tribe Camelineae encompasses *Arabidopsis* (DC.) Heynh. nom. cons., *Camelina* Crantz, *Capsella* Medik. nom. cons., *Catolobus* (C.A. Mey.) Al-Shehbaz, *Chrysochamela* (Fenzl) Boiss., *Neslia* Desv. nom. cons. and *Pseudoarabidopsis* Al-Shehbaz, O'Kane & R.A. Price (Koch et al., 2018). To reduce potential topological and time estimation artefacts due to a biased taxon sampling, we included at least one representative per higher taxonomic unit (genus) in our analysis. The taxonomic placement of *Noccidium* endemic to Iran has been controversial, as morphological evidence and some molecular studies (Couvreur et al., 2010) have consistently placed it into Coluteocarpaeae (German, 2018). Therefore, it was excluded from our analysis. For the same reason, all but two currently accepted *Camelina* species were used in this study to ensure our analyses were representative and robust. The following *Camelina* species were included in the analyses: *Camelina alyssum* (Mill.) Thell., *Camelina anomala* Boiss. & Hausskn. ex Boiss., *Camelina hispida* Boiss., *Camelina laxa* C.A. Mey., *Camelina microcarpa* Andrz. ex DC., *Camelina rumelica* Velen. and *Camelina sativa* (L.) Crantz.

The identity of *Camelina alpkoyensis* Yıld. (Yıldırımlı, 2011) remained unresolved and was therefore not included in the analysis. The type specimen is currently at the author's disposal only and neither leaf nor seed material is available for molecular analysis (Yıldırımlı, pers. comm. Sept. 19th 2018). The holotype material was supposedly located in the Herbarium of Hacettepe University (Turkey); however, their database does not hold any information of it (Özüdoğru, pers. comm.).

Thus, molecular analysis, which would clarify whether this truly is a species or just an ecotype of *Camelina laxa*, could not have been carried out. During the course of this study, a new species *Camelina neglecta* J.Brock, Mandáková, Lysak & Al-Shehbaz was described (Brock et al., 2019). Nevertheless, as the origin of the type material is an unspecified seed collection without geo-referenced data deposited at the USDA and since no original voucher was ever recovered, we also excluded this taxon from the present analysis.

Geographical data was recorded at the specimen collection site, recovered directly from the voucher specimens or reconstructed by using TopGlobus (http://www.topglobus.ru/), Wikimapia (http://www.wikimapia.org/) and Loadmap (http://loadmap.net/) based on text information from the voucher specimens. Distribution maps were produced using QGIS (https://qgis.org/en/site/). The Eurasian steppe belt shape included the following geographic units: Altai Steppe and Semi-Desert, Daurian Forest Steppe (NE Mongolia, S Siberia), Eastern Anatolian Montane Steppe, Emin Valley Steppe (China-Kazakhstan border), Kazakh Forest Steppe, Kazakh Steppe, Kazakh Upland, Mongolian-Manchurian Grassland, Pontic Steppe, Sayan Intermontane Steppe, Selenge-Orkhon Forest Steppe (north-central Mongolia) and South Siberian Forest Steppe. Their descriptions were taken from The Nature Conservancy (http://maps.tnc.org/gis_data.html) and WWF For a Living Planet – Temperate Grasslands, Savannas and Shrubland Ecoregions (https://www.worldwildlife.org/biomes/temperate-grasslands-savannas-and-shrublands). The geographic terminology of the East European Plain followed the floristic division of Komarov's U.S.S.R Flora included in the Flora Europaea (as outlined in Tutin et al., 1993).

2.2. Molecular analyses and ancestral area reconstruction

Total genomic DNA was isolated from herbarium specimens using the InnuPREP Plant DNA Kit (Analytic Jena AG, Jena, Germany) according to the instructions of the manufacturer and used directly in PCR amplifications. PCR was carried out using 10 µl of 2X HS Taq Mix Red (Biozym Scientific, Hessisch Oldendorf, Germany), 7 µl of ddH_2O, 1 µl of each 10 µM primer and 1 µl of template DNA. Alternatively, PCR was carried out using 20.8 µl of ddH_2O, 3 µl of 10x Taq Buffer with $MgCl_2$, 2 µl of 10 mM dNTP mixture, 1 µl of DMSO, 0.2 µl of 5U Taq Polymerase, 1 µl of corresponding primer, and 1 µl of the template DNA. The analysed ETS region spanned from a conserved sequence segment in the 18S rRNA operon to a semi-conserved 9-bp motif near the 5′ end of ETS (Cordesse et al., 1993). Slightly shortened primer 18S-IGS (5′-GAGACAAGCATATGACTACTGGCAGGATC-3′; Baldwin and Markos, 1998) and ETS-9 primer (5′-CATGGGCGTGTGAGTGGTGA-3′; Wright et al., 2003) were used. PCR conditions included initial denaturation at 95 °C for 5 min, 35 cycles of denaturation at 95 °C for 30 s, primer annealing at 55 °C and extension at 68 °C for 1 min, and additional extension at 68 °C for 10 min. Samples that yielded single bands once run through gel electrophoresis were sent to Microsynth Seqlab (Göttingen, www.microsynth.ch) for purification and sequencing. Sequences from each accession were manually edited in Chromas Lite 2.6.4 (Chromas | Technelysium Pty Ltd) and aligned and manually corrected in AliView v1.24 (Larsson, 2014). In addition to the ETS region, internal transcribed spacer (ITS) and trnL intron, psbA-trnH, ndh-rpL32 and trnQ-rps16 chloroplast regions were also investigated on a reduced taxon sample to infer potential intraspecific variation, using corresponding primers published previously in Blattner (1999); Taberlet et al. (1991); Sang et al. (1997) and Shaw et al. (2007), respectively.

Maximum likelihood tree inference relied on RAxML-HPC v8.2.10 (Stamatakis, 2014) with 1000 ML bootstrap iterations with the default number of distinct rate categories. Tree searches were executed from a random maximum-parsimony tree employing either the GTRCAT or the GTRGAMMA bootstrapping phase model. In addition, Bayesian Inference was carried out using MrBayes v3.2.6 (Ronquist et al., 2012) under GTR + Γ substitution model using the random-addition-sequence method with 10 replicates. Two independent Markov chain Monte Carlo (MCMC) analyses of four chains were run – one cold and three heated. Runs were carried out for 20 million cycles, with parameters and trees sampled every 1000th cycle including an appropriate burn-in (10%) as inferred from the evaluation of the trace files using Tracer v1.6.0 (Rambaut et al., 2018).

Molecular dating analyses relied on BEAUti & BEAST v1.8.4 (Suchard et al., 2018) and the strict clock model. The use of the strict clock model was justified, after assessing the coefficient of variation in the trace file, which did not exceed 0.3. We used a Coalescent: Constant Size tree prior, the GTR + Γ substitution model, and three independent Monte Carlo Markov chains (MCMC) runs for 20 million generations, with parameters sampled every 10,000th generation. Effective sample sizes (ESS) for all estimated parameters were assessed using Tracer v1.6.0 (Rambaut et al., 2018). TreeAnnotator v1.8.4 (Suchard et al., 2018) was used to discard 10% of the saved trees and annotate the rest of them. Maximum clade credibility tree with median node heights was visualized using FigTree graphical viewer of phylogenetic trees v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). All sequence evolution models used in this study were assessed using the Akaike information criterion (AIC) implemented in the jModelTest2 v2.1.6 (Darriba et al., 2012). All the analyses were carried out at the CIPRES Science Gateway computing facility (Miller et al., 2010). The aligned matrices are available as *.nex files upon request.

Time estimation analysis was carried out using only one representative per ribotype, reducing the whole dataset to 26 entries. The following two secondary calibration points were used in the analysis: the crown age of tribus Camelineae dated to 8.16 ± 2.0 mya (Hohmann et al., 2015) and the crown group of *Arabidopsis* dated to 5.72 ± 1.40 mya (Hohmann et al., 2015; Koch, 2019; Novikova et al., 2016). To elucidate the historical biogeography of *Camelina*, an ancestral area reconstruction was carried out, using the Bayesian binary MCMC (BBM) method (Ali et al., 2012) implemented in RASP v4.1.0 (Yu et al., 2015). Distribution ranges were divided into four regions: (A) Europe, including Pontic steppe, (B) western and central regional sub centre of the Irano-Turanian region (Léonard, 1988), (C) Kazakh steppe, uplands and East Kazakhstan and (D) Mongolian-Chinese steppe. The BBM was run with a fixed state frequencies model (JC model) including either equal among-site rate variation or gamma distribution (0.001, 200) for two million generations, ten chains each, and two parallel runs. Maximal number of allowed areas was set to two and we sampled every 1000th tree with a final burn-in of 10%. Furthermore, region combinations A and D, as well as B and D were not allowed.

3. Results

3.1. Phylogeny and geographic distribution

In total, 120 sequences were generated and deposited to GenBank under MN120314—MN120430 over the course of the study. The ETS alignment encompassed 122 entries, was 431 bp long and comprised 280 constant characters, 52 variable characters and 99 parsimoniously informative characters. Fig. S2 shows the best-scoring Maximum Likelihood (ML, LBS values) tree (–ln = 1835.01) carried out by employing the GTRCAT bootstrapping phase model and the Bayesian tree (BI, BPP values). Both trees show a well resolved backbone and congruent topologies, but not fully resolved internal topology. This is the first well-supported phylogeny of Camelineae that includes all the representatives of the tribus. *Arabidopsis* (100LBS, 1.00BPP), *Capsella* (100LBS, 1.00BPP) and *Camelina* (97LBS, 1.00BPP) taxa were recovered as monophyletic. *Neslia* was placed as the sister group to *Camelina* (99LBS, 1.00BPP) and *Catolobus* tentatively as the sister group to *Capsella* (80LBS, .94BPP). The exact position of *Chrysochamela* remained unclear. *Camelina laxa* was ranked at the base of *Camelina* group (97LBS, 1.00BPP). Internal relationships of *C. sativa* ribotypes (R1 and R2), *C. alyssum*, *C. anomala* and *C. hispida* remained unclear, albeit moderately supported and grouped according to the morphology.

Eastern *C. microcarpa* ribotypes (E1 and E2) were placed as sister ribotypes (94LBS, .91BPP) to a combined *C. microcarpa* Western ribotypes (W1 and W2) and *C. rumelica* (98LBS, 1.00BPP). Further subgrouping within Eastern *C. microcarpa* ribotype as well as within Western *C. microcarpa* ribotype was recognised, albeit with a very low statistical support (66LBS, .57BPP for the eastern ribotypes and 61LBS, .00BPP for the western ribotypes). No delimitating names have been assigned to the *Camelina microcarpa* and *Camelina sativa* ribotypes. The investigated internal transcribed spacer (Fig. S3) as well as the chloroplast regions (data not shown) displayed little to no variation in *Camelina* and were therefore not used for further studies.

On the species level, ribotypes and morphological determinations recovered highly congruent results. Morphological determinations were carried out based on the inflorescence structure, fruit size and shape, life strategy, and stem hairiness, following Al-Shehbaz (1987); Brock et al. (2019) and Mirek (1984). *Camelina rumelica* was not only genetically, but also morphologically well separated from the rest of *Camelina* species. Its stem is namely covered in numerous long simple trichomes and its fruits are more densely spaced in the upper than in the lower part of the infructescence. Life strategy and size of the fruits separated the *C. microcarpa* (Fig. 1) from *C. sativa-alyssum*. While *C. sativa* and *C. alyssum* are the only summer annual species and have (probably through human induced directional selection) developed significantly bigger fruits and seeds, *C. microcarpa* exhibits a winter annual life strategy and produces significantly smaller fruits and seeds. *Camelina sativa* and *C. alyssum* were told apart based on the stem hairs. While *C. sativa* builds mostly stellate trichomes, stems of *C. alyssum* are covered in short simple non-branching trichomes. *Camelina anomala*, *C. hispida* and *C. laxa* were the only three morphologically distinct species. *Camelina anomala* possesses simple stiff hairs only on the lower stem parts and linear-cylindrical siliques with seeds arranged in one row only, whilst all the other *Camelina* species build pyriform fruits with two-rowed seeds. *Camelina hispida* has pubescent middle stems and inflorescences and *C. laxa* exhibits heavily flexuous infructescences. *Camelina hispida* and *C. laxa* are also self-incompatible, while all the other *Camelina* species are self-compatible. The self-compatibility of *C. anomala* is currently unknown.

Camelina microcarpa can be found across the whole Eurasian Steppe belt and is the most widely distributed species of *Camelina* in Eurasia (Fig. 3). A similar geographic area was observed for *Camelina sativa*, however its natural distribution range remained unclear. Our surveys found *C. sativa* across the whole Eurasian steppe belt, with a gap in the Central Kazakh steppe (Fig. 4). Within *C. sativa* two ribotypes have been recognised, where one ribotype was discovered only eastwards from the Pontic steppe region. In the high altitude steppes of the Emin Valley (Kazakhstan-China border) both ribotypes co-occur nowadays. The distribution area of *Camelina rumelica* is more restricted, and encompasses southern parts of the Pontic steppe, lowlands of South-Eastern Europe, Middle East steppes and Western Kazakh steppes (Fig. 3). *Camelina anomala*, *Camelina hispida* and *Camelina laxa* grow only in Central Anatolian steppes and Middle East steppes (data not shown). *Camelina alyssum* was once found across most of Europe and Western Asia but is nowadays treated as extinct in Western and Northern Europe. However, it is still present, albeit scarcely, in the East European Plain.

3.2. Time estimation analysis and historical biogeography

The overall topology of the phylogenetic tree and its rooting corresponded with other partial and whole plastome studies that have analysed phylogenetic and temporal aspects of Brassicaceae evolution (Guo et al., 2017; Hohmann et al., 2015). Time estimation analysis (Fig. 2) generated a highly congruent topology (compared to Maximum Likelihood and Bayesian Inference analyses), with ESS values exceeding 4000. It placed *Camelina* diversification on the Pliocene/Pleistocene border. The first major split in *Camelina* occurred at 1.5 ± 1 mya, resulting in a *C. microcarpa-rumelica* lineage and *C. sativa-alyssum* lineage including other species with highly restricted distribution area. The east-west split in the *C. microcarpa-rumelica* lineage was dated to 1.2 ± 1 mya and the split between *C. microcarpa* and *C. rumelica* first at 0.75 ± 1 mya. Irrefutable conclusions on *C. sativa-alyssum* lineage cannot be drawn due to a modest statistical support.

Due to a smaller size of reconstructed areas and moderate statistical support on deeper nodes within *Camelina*, a Bayesian Binary Method (BBM) was used for ancestral area reconstruction, which accepts polytomies and tends to suggest single distribution areas for ancestral nodes more often than other methods (Müller et al., 2015). Both BBM models JC and JC + G generated highly congruent results (data not shown). The inferred ancestral area of *Camelina* was shown to be the western and central regional sub centre of the Irano-Turanian region (Fig. 5: node 21), from which both *Camelina* lineages spread across the Kazakh steppes and uplands (Fig. 5: node 20). Eastern ribotypes of *C. microcarpa* evolved east of the Caspian Sea and potentially spread eastwards (Fig. 5: node 19 → 16). While western ribotypes of *C. microcarpa* and *C. rumelica* lineage never re-occupied the Irano-Turanian region when migrating from east towards west (Fig. 5: node 19 → 18), *C. hispida* and *C. anomala* possibly returned to this floristic region (proportional increase of B in Fig. 5: nodes 15 → 14 → 12). *Camelina rumelica* plausibly developed in the Pontic steppe and later on occupied the Irano-Turanian floristic region (Fig. 5: node 18). Ribotype 2 of *C. sativa* did not migrate with the rest of species and rather stayed in the eastern steppes (east of Caspian Sea; Fig. 5: node 15). Contrary, *C. alyssum* and ribotype 1 of *C. sativa* migrated to the west (Fig. 5: node 14 → 13). For details on dispersal, vicariance and/or extinction events refer to the Fig. S5.

4. Discussion

The use of molecular markers has greatly enhanced our knowledge on past vegetation patterns, including small, sub-regions of the Eurasian

Fig. 1. Habitus of *Camelina microcarpa*. A – Leaves (Tatarstan, Bavly; May 31 st 2010; by E. Izmailov), B – inflorescence (Tatarstan, Bavly; May 31 st 2010; by E. Izmailov), C – close up of an inflorescence (Krasnodar, Novorossiysk, Markotkh Ridge; May 16th 2016; by A. Malykhina), D – dried silicles (Krasnodar, Anapa, Utrish; May 10th 2016; by S. Banketov), E – seeds (Krasnodar, Anapa, Utrish; May 10th 2016; by S. Banketov).

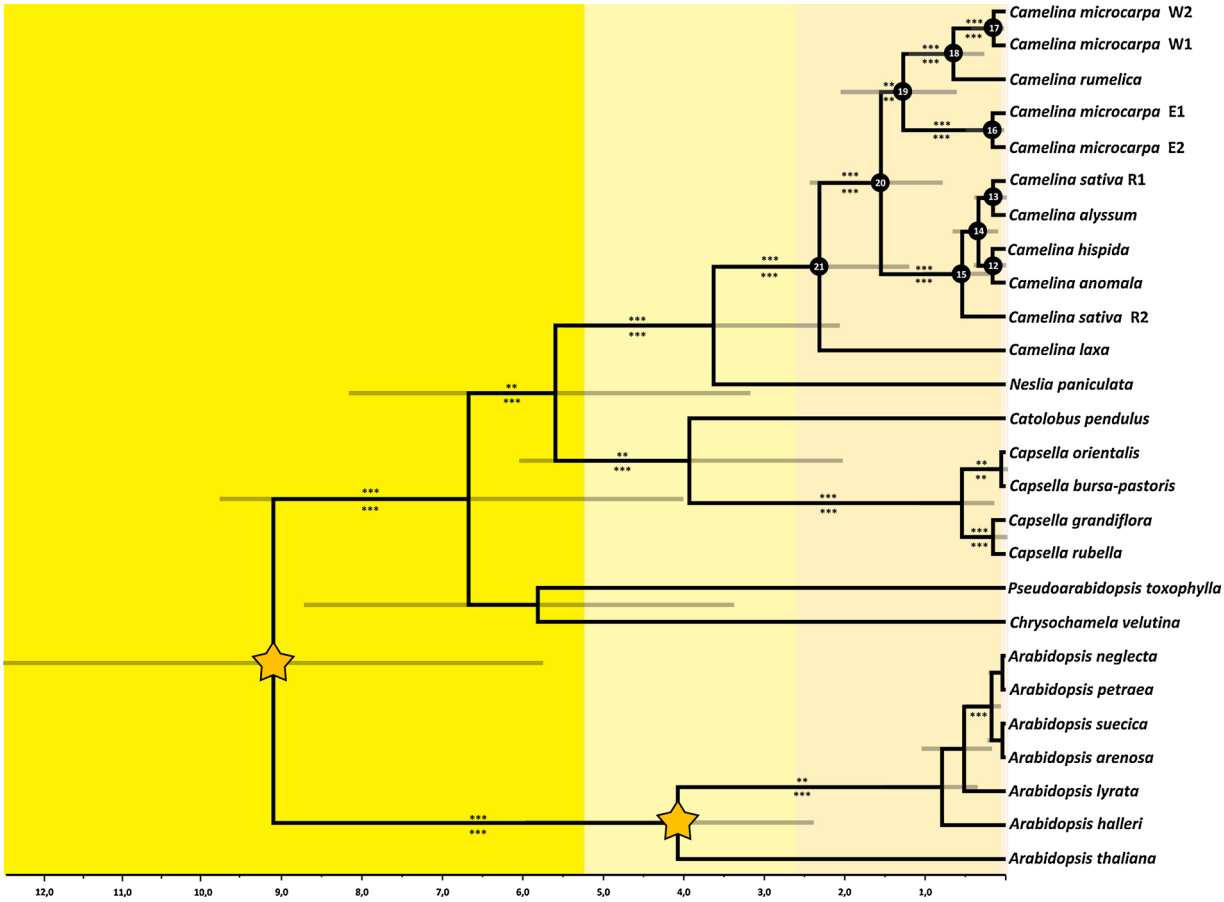

Fig. 2. Dated phylogeny of Camelineae. Median rate is given in units of substitutions per million years (including 95% confidence intervals). Absolute ages are in millions of years, and epochs are indicated in the same colours as in Gradstein et al. (2012). The numbers on the branches are statistical support values (above: ML bootstrap values, values < 75 are not shown; below: Bayesian posterior probabilities, values < .90 are not shown). Significance levels: highly supported (***): BI ≥ 0.98, BS ≥ 95%; well supported (**): BI ≥ 0.90, BS ≥ 85% to < 95%; supported (*): BI ≥ 0.90, BS ≥ 75% to < 85%. Orange stars indicate the calibration points.

steppe belt (Kajtoch et al., 2016; Meindl et al., 2016; Meng et al., 2015; Plenk et al., 2017; Qiu et al., 2011). However, only a handful of studies (Franzke et al., 2004; Friesen et al., 2016; Hurka et al., 2012; Volkova et al., 2017) have focused on the Eurasian steppe belt as a whole. In our study, we put the biogeographical history of *Camelina* into time and space and linked it with past climate-landscape features (Fig. 6). Furthermore, we present the first comprehensive phylogeographic study of *Camelina* encompassing the whole documented distribution ranges of all species. The only potential sampling artefact could be the lack of *C. sativa* accessions from the Central Kazakh steppe. *Camelina sativa* and *C. alyssum* are namely the only two *Camelina* summer annuals (Al-Shehbaz, 1987). As most of our fieldwork in the Kazakh steppe was carried out in early summer, these individuals could have escaped our notice.

4.1. Phylogenetic analysis of the genus Camelina

The tribus Camelineae was constrained secondarily with a node by Hohmann et al. (2015), where the tribus was represented using only three different genera – *Arabidopsis*, *Camelina* and *Capsella*. To justify this step, an additional time estimation analysis was carried out including only the taxa used in the Hohmann et al. (2015) study (Fig. S4). Topologies including time spans of both analyses were compared and no significant changes in time constraints or topologies were detected. Therefore, using the same node to date the whole Camelineae tribe seemed to be justified. Furthermore, crown group age time spans and time spans of speciation events in *Capsella* and *Arabidopsis* corresponded with previous independently published time span ages: crown age of *Capsella* dated to be up to approximately 1 mya (Douglas et al., 2015), the diversification of *Arabidopsis* to have taken place approximately 6 mya (Novikova et al., 2018), split between *Arabidopsis* and sister clade including *Capsella* dated to approximately 9 mya (Koch et al., 2000) and the Pleistocene origin of *Capsella rubella* and *Capsella grandiflora* (Hurka et al., 2012). Thus, despite using only secondarily calibrated nodes resulting in wider 95% confidence intervals, our results are supported by different independent studies that used different calibration methods.

Time estimation analysis and ancestral area reconstruction placed *Camelina* diversification at approximately 2.5 ± 1 mya on the Pliocene/Pleistocene border into the western and central regional subcentre of the Irano-Turanian region. This time frame coincides with a global decline in temperatures and humidity (Yang and Ding, 2010), resulting in a cooler and dryer climate that promoted grassland expansion at the cost of forest decline. As a mostly continuous Eurasian steppe belt already existed on the Miocene/Pliocene border (Mai, 1995; Velichko, 1999), *C. alyssum*, *C. microcarpa*, *C. rumelica* and *C. sativa* consecutively might have spread across an already well-established steppe belt. According to our time estimation analysis, *Camelina* is a young genus, which explains the high levels of hybridisation potential and minute morphological character differences between species. Our study uncovers several ribotypes within *Camelina* group, indicating that despite minute morphological character differences even between species, intraspecific variation (at least in the ETS region) in *Camelina* is present.

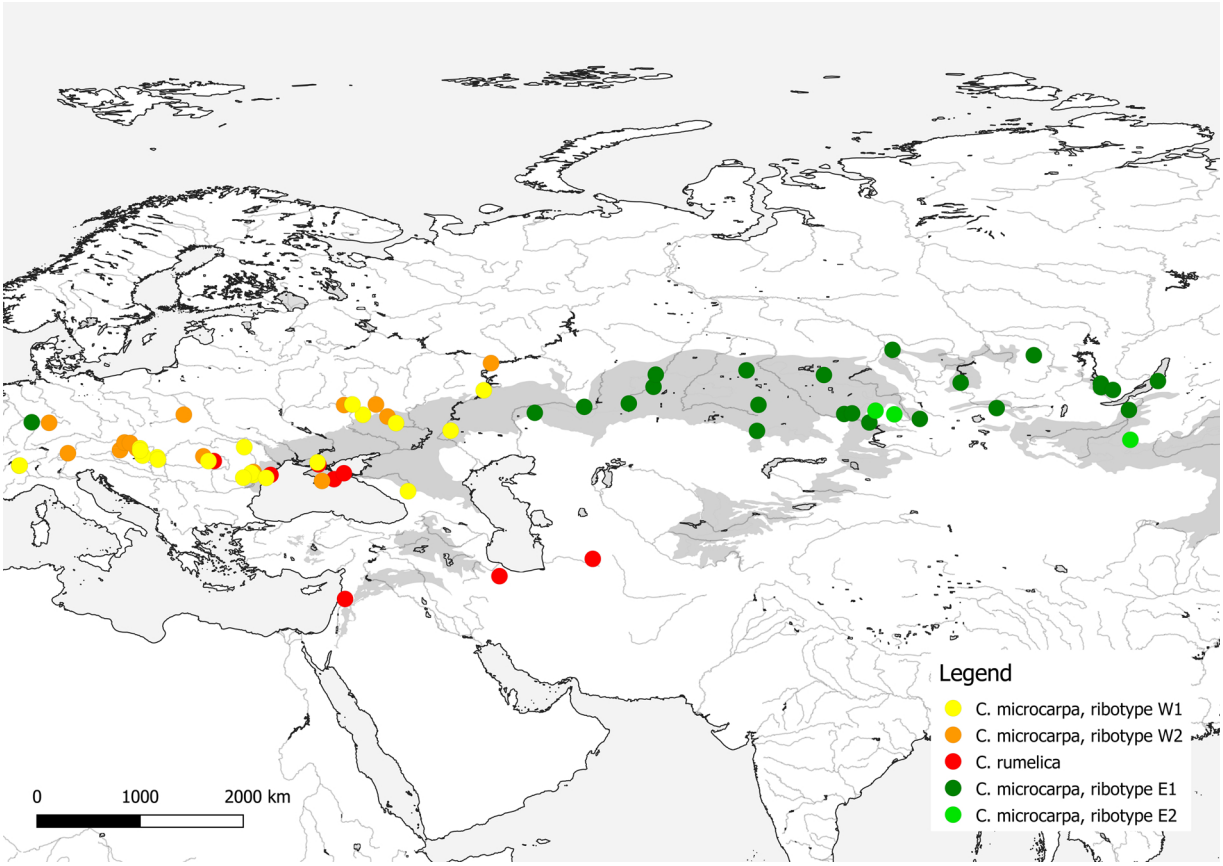

Fig. 3. Distribution map of *Camelina rumelica* (ribotype indicated in red) and *Camelina microcarpa* (western ribotype 1 (W1) indicated in yellow, western ribotype 2 (W2) indicated in orange, eastern ribotype 1 (E1) indicated in dark green and eastern ribotype 2 (E2) indicated in light green).

4.2. Historical biogeography of the genus Camelina

4.2.1. Camelina microcarpa

The split between *C. microcarpa* ribotypes E1 and E2 and joined *C. microcarpa* ribotypes W1 and W2, and *C. rumelica* is dated to 1.2 ± 1 mya, which coincides with the short-lived Apsheronian transgression of the Caspian Sea (Kroonenberg et al., 1997; Svitoch, 2012; Tudryn et al., 2013). The last common ancestor of these taxa also occupied the Kazakh steppes, which furthermore supports the potential influence of the Caspian Sea transgression. This transgression was a result of a massive freshwater influx caused by the rise of temperatures and a consequent meltdown of glaciers. At its peak, Caspian Sea mean sea level was approximately 100 m higher than nowadays, covering the modern Volga–Kama catchment (Starobogatov, 1994; Van Baak et al., 2013). The western parts of the Pontic–Caspian steppe were not affected by this transgression nor by other consequential events of the glacial periods, and thus provided a refugium for steppe vegetation (Stewart et al., 2010; Varga, 2010).

Within each *C. microcarpa* ribotype subgroups, two additional ribotypes can be detected. However, the statistical support is modest. The two eastern ribotypes (coded in yellow and orange in Fig. 3) and the two western *C. microcarpa* ribotypes (coded in dark and light green in Fig. 3) display only a weak geographical correlation. As the statistical support does not allow for irrefutable conclusions, we can only speculate that the weak geographical correlation of these two ribotypes could be related to human activity. Human influence might also explain that an outlier 12-0076-10-00 has a distinct E1 Kazakh ribotype, but is found in the German state of Rhineland-Palatinate in the Mainz Sand Dunes (dark green point in the left corner of Fig. 2). Another explanation could be the dispersal mechanism of *C. microcarpa*. Wet seeds namely produce a thick mucilage coat, which promotes adherence ability and hence long-distance dispersal by either birds or cattle. An alternative explanation, however, would be that the distribution of western ribotypes is the result of extensive human influence. If this is the case, the Mainz outlier could represent a leftover of a once widely distributed E1 (eastern) ribotype. The Mainz Sand Dunes are namely home to a number of steppe elements (i. e. *Gypsophila fastigiata*, *Onosma arenaria*, *Stipa capillata*), whose distribution, with the exception of Mainz Sand Dunes, do not reach westwards of the Pannonian Basin (Hecker, 1987).

4.2.2. Camelina rumelica

Camelina rumelica separated from western ribotypes of *C. microcarpa* approximately 0.75 ± 1 mya. This time span is known for intensified development of permafrost, loess and glacial sheets and consequential radical transformation of the zonal landscape structure. The Baku transgression of the Caspian Sea also belongs to this era (Hurka et al., 2019). The native distribution range of *C. rumelica* encompasses the southern parts of the East European Plain, Central Asia and Middle East, however it has also been introduced and naturalised in Central and western Europe, and in the United States (Plants of the World Online, http://www.plantsoftheworldonline.org/). We argue that this species developed in the Pontic-Caspian steppe in the mid-upper Pleistocene and expanded its distribution area westwards through southern East European Plain and southwards through Asia Minor. Such a scenario likely occurred during only recent glacial periods that favoured dry-adapted steppe species, which would also explain the lack of ribotype differentiation. Despite our extensive field trips and herbarium surveys, no *C. rumelica* plants were discovered east of the 70th meridian. Furthermore, there was also no distribution data on *C. rumelica* that would indicate its presence east of the abovementioned meridian. This together with our ancestral area reconstruction further supports our

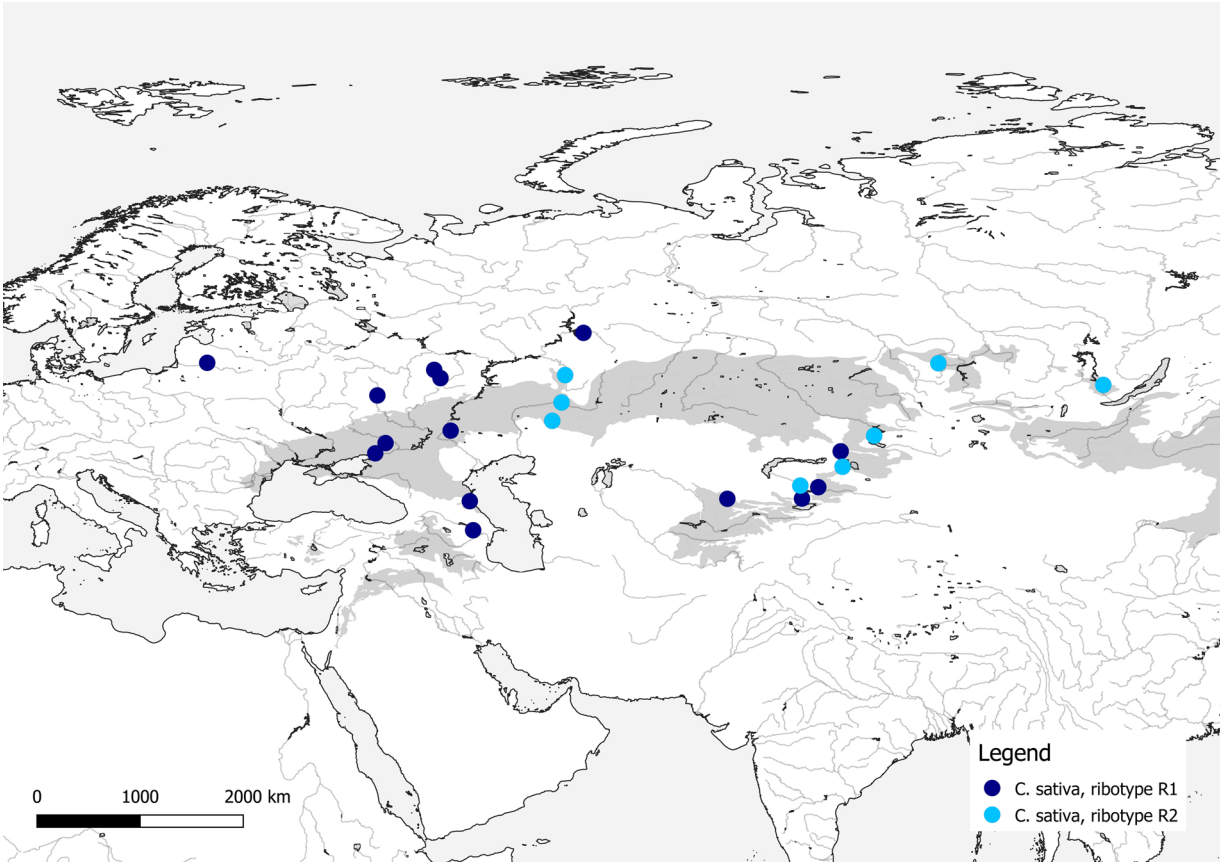

Fig. 4. Distribution map of *Camelina sativa* (ribotype R1 indicated in dark blue and ribotype R2 indicated in light blue).

argument that *C. rumelica* conceivably developed in the western parts of the Eurasian steppe belt and stepwise moved (south)-east and west from the Pontic steppe.

4.2.3. *Camelina sativa and Camelina alyssum*

Contrarily to previous assumptions (Brock et al., 2018), populations of *Camelina sativa* have been reported from the grasslands of the East European Plain, the Caucasus, the coastline of Caspian Sea, south-eastern parts of Central Asia (Emin Valley; China-Kazakhstan border) and southern Siberia. It remains unclear, however, if those populations are of anthropogenic origin or indeed reflect the species' natural distribution range. With an exception of Emin Valley, all the other regions

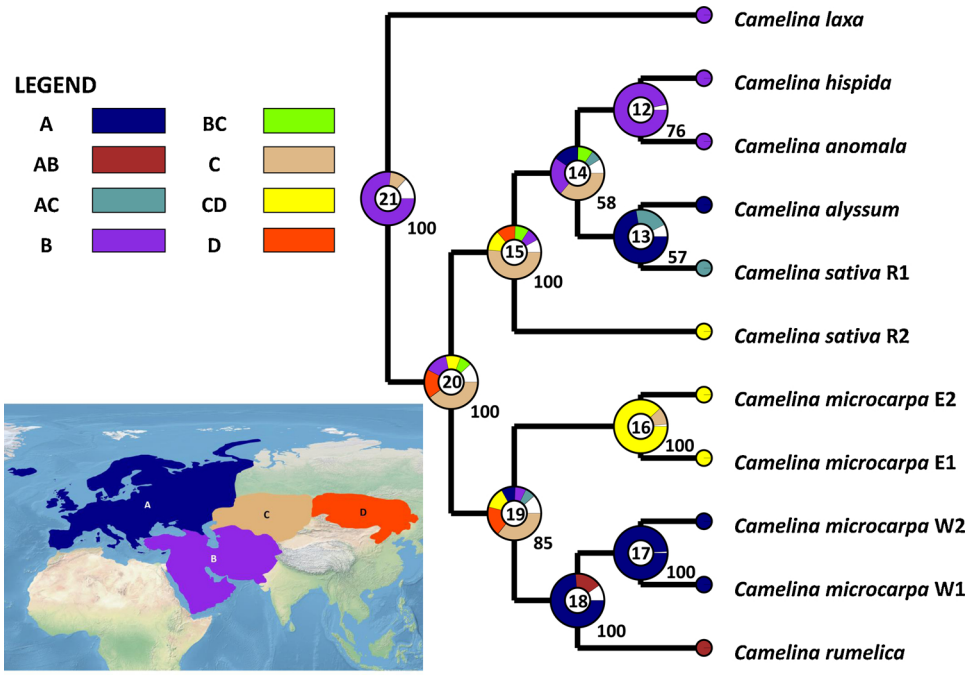

Fig. 5. Ancestral area reconstruction of *Camelina* using the Bayesian binary MCMC (BBM) method, based on the time phylogeny derived from the BEAST analysis. The areas are coded as follows: (A) Europe, including Pontic steppe, (B) western and central regional sub centre of the Irano-Turanian region (Léonard, 1988), (C) Kazakh steppe, uplands and East Kazakhstan and (D) Mongolian-Chinese steppe. Unresolved parts are indicated in **white.** Numbers next to the nodes represent node frequency (%).

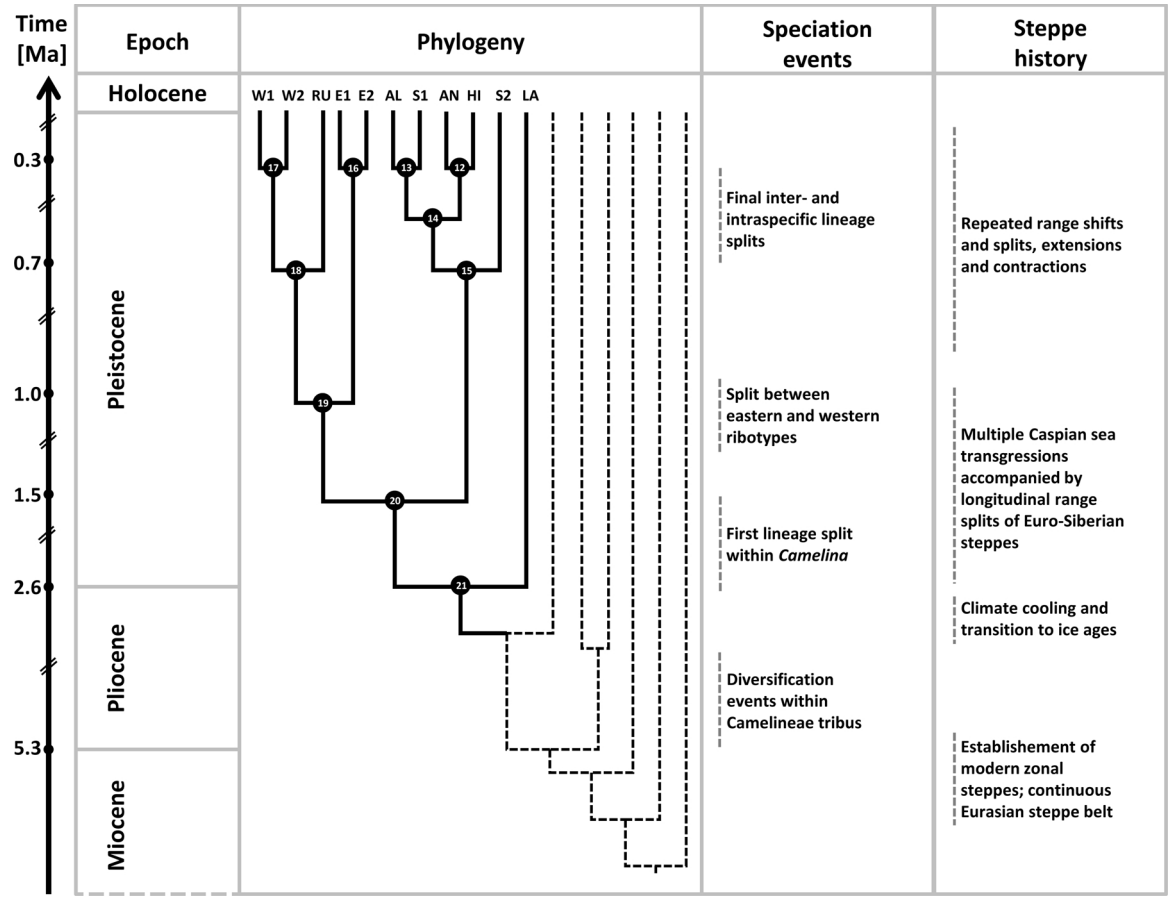

Fig. 6. Outline of the evolutionary history of Camelineae. W1 (western *C. microcarpa* ribotype 1), W2 (western *C. microcarpa* ribotype 2), RU (*C. rumelica*), E1 (eastern *C. microcarpa* ribotype 1), E2 (eastern *C. microcarpa* ribotype 2), AL (*C. alyssum*), S1 (*C. sativa* ribotype 1), AN (*C. anomala*), HI (*C. hispida*), S2 (*C. sativa* ribotype 2), LA (*C. laxa*). Taxa indicated with interrupted lines are *Neslia*, *Catolobus*, *Capsella*, *Pseudoarabidopsis*, *Chrysochamela* and *Arabidopsis* (from left to right). Geological epochs according to the International Commission on Stratigraphy (Cohen et al., 2018).

are well known for flax production in the past (Vavilov, 1992), thus it is possible that most of the investigated populations are the remnants of the former cultivated flax (*Linum usitatissimum*) fields. It has been stated that *C. sativa* is the ancestor of *C. alyssum* which has developed in flax fields as a result of directional selection, causing morphological changes in the plant's habitus, fruit morphology and ripening type to resemble flax (Barrett, 1983). Thus, it also shares the same distribution dynamics as flax (Stebbins, 1950; Vavilov, 1992). This is supported by our topology, previous crossing experiments and detailed morphological studies that all point towards a close relationship between *C. sativa* and *C. alyssum* (Sinskaia and Bezluzheva, 1931; Stebbins, 1950; Tedin, 1925; Zinger, 1909). With the abandonment of flax fields, the distribution range of *C. alyssum* also shrank and the species is nowadays considered extinct from most of Central Europe (Francis and Warwick, 2009). Both *C. sativa* ribotypes that were uncovered show a moderate east-west geographical correlation. Similar delineation into two groups has been observed recently, albeit with a different geographical correlation (Luo et al., 2019). The results are, however, incomparable as their study focused on a different geographic region and used a different set of markers.

The overall geographic signal of *Camelina* has not been overshadowed by the human influence, despite being irrefutably proven to be extensively used already in the Neolithic as an oil source (Hovsepyan and Willcox, 2008). However, it remains unclear which species exactly were used by humans in the Neolithic, if *Camelina* has just been gathered or actively cultivated, and if the latter, when and where the cultivation started (Brock et al., 2018; Hovsepyan and Willcox, 2008).

4.3. Species' determination and utility of molecular markers for the genus Camelina

Morphological determinations were highly congruent with the ETS sequences; hence the external transcribed region functions as a barcode for *Camelina* at least on a species level. Furthermore, we have uncovered at least two different ribotypes within *C. microcarpa* and at least two within *C. sativa* while amplifying this only 430 bp-long-region. This finding provides a powerful tool for further studies to undoubtedly identify the investigated *Camelina* species on a molecular basis, which is of great importance, as *Camelina* species are hard to delimitate based on morphology only. This locus may even work to delimitate species within other young, morphologically similar and highly plastic species groups, as the flanking regions for primers are highly conserved at least throughout the whole core eudicots (Baldwin and Markos, 1998; Wright et al., 2003).

5. Conclusion

Temporal and spatial aspects of florogenesis within the Eurasian steppe belt are poorly understood, despite their great importance in climate-landscape reconstructions. We have uncovered hidden genetic diversity of *Camelina*, which harbours an important species for the agronomic oil industry – *Camelina sativa*. We have shown that phylogeny, differentiation and range evolution of *Camelina* indeed mirrors the climate/landscape dynamics of the steppe (Fig. 6) (Hurka et al., 2019). Furthermore, we put the development of the group into time and

space and found a likely correlation between palaeogeographical and palaeoenvironmental events and the evolutionary history of *Camelina*. Current taxonomic treatments of *Camelina* have been confirmed using the ETS locus, however, only on the species level. Genetic dissimilarities of subspecies and their names were incongruent and their nomenclature remains to be completed. We trust that using the ETS will facilitate an easy and reliable species determination, possibly not limited to *Camelina* only. Despite sharing history with humans, *Camelina* has developed on its own prior to its use in Neolithic period, however many aspects of the domestication are still yet to be answered. Our approach proved successful at uncovering another missing point-of-view of florogenesis within the Eurasians steppe belt.

Ethical approval

This article does not contain any studies with animals carried out by any of the authors.

Sampling and field studies

The study was performed in compliance with the Convention on Biological Diversity (CBD).

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

We would like to thank Simon Pfanzelt to kindly provide us with the Eurasian Steppe belt shape file and all the curators of the HBG, JE and NSK herbaria, who kindly provided the material for sequencing. We also thank Juliana Chacón, who helped us with the ancestral area reconstruction. This work was supported by the Deutsche Forschungsgemeinschaft [grant to BN NE 314/15-1] and the German Federal Foreign Office (via DAAD Rise Germany programme) that supported NPH during his stay in Germany. Furthermore, we would like to acknowledge current digitisation activity in the Moscow University Herbarium [under Government order #AAAA-A16-116021660039-1 headed by APS]. We would also like to thank the 'Plantarium' web-community members, who allowed us to use their photographs: Sergey Banketov, Elvir Izmailov and Anna Malykhina; and Nikolai Friesen, who contacted the photographers on our behalf.

References

Ali, S.S., Yu, Y., Pfosser, M., Wetschnig, W., 2012. Inferences of biogeographical histories within subfamily Hyacinthoideae using S-DIVA and Bayesian binary MCMC analysis implemented in RASP (reconstruct Ancestral State in Phylogenies). Ann. Bot. 109, 95–107.

Al-Shehbaz, I.A., 2012. A generic and tribal synopsis of the Brassicaceae (Cruciferae). Taxon 61, 931–954.

Al-Shehbaz, I.A., 1987. Camelina. J. Arnold Arbor. 68, 234–240.

Arkhipov S. A., Volkova V. S., Zolnikov I. D., Zyklina V. S., Krukover A. A., Kulkova I. A. 1999. West Sibiria. In: Climate and Environment Changes during the last 65 Million Years; Cenozoic: from Paleocene to Holocene. Eds. A. A. Velichko, V. P. Nechaev. Chapter 4. GEOS, Moscow, 84–109 pp. [In Russian]. (Архипов С. А., Волкова В. С., Зольников И. Д., Зыкина В. С., Круковер А. А., Кулькова И. А. 1999. Западная Сибирь // Изменение климата и ландшафтов за последние 65 миллионов лет (кайнозой: от палеоцена до голоцена. Под ред. А. А. Величко, В. П. Нечаева. М.: ГЕОС. С. 84–109).

Bakker, E.S., Gill, J.L., Johnson, C.N., Vera, F.W.M., Sandom, C.J., Asner, G.P., Svenning, J.-C., 2016. Combining paleo-data and modern exclosure experiments to assess the impact of megafauna extinctions on woody vegetation. Proc. Natl. Acad. Sci. 113, 847–855.

Baldwin, B.G., Markos, S., 1998. Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). Mol. Phylogenet. Evol. 10, 449–463.

Barnosky, A.D., 2004. Assessing the causes of late Pleistocene extinctions on the continents. Science 306, 70–75.

Barrett, S.H., 1983. Crop mimicry in weeds. Econ. Bot. 37, 255–282.

Bartlett, L.J., Williams, D.R., Prescott, G.W., Balmford, A., Green, R.E., Eriksson, A., Valdes, P.J., Singarayer, J.S., Manica, A., 2016. Robustness despite uncertainty: regional climate data reveal the dominant role of humans in explaining global extinctions of Late Quaternary megafauna. Ecography 39, 152–161.

Binney, H., Edwards, M., Macias-Fauria, M., Lozhkin, A., Anderson, P., Kaplan, J.O., Andreev, A., Bezrukova, E., Blyakharchuk, T., Jankovska, V., Khazina, I., Krivonogov, S., Kremenetski, K., Nield, J., Novenko, E., Ryabogina, N., Solovieva, N., Willis, K., Zernitskaya, V., 2017. Vegetation of Eurasia from the last glacial maximum to present: key biogeographic patterns. Quat. Sci. Rev. 157, 80–97.

Blattner, F.R., 1999. Direct amplification of the entire ITS region from poorly preserved plant material using recombinant PCR. BioTechniques 27, 1180–1186.

Bone, M., Johnson, D., Panayoti, K., Kintgen, M., Vickerman, L.G., 2015. Steppes: the Plants and Ecology of the World's Semi-arid Regions. Timber Press, Portland, Oregon.

Brock, J.R., Dönmez, A.A., Beilstein, M.A., Olsen, K.M., 2018. Phylogenetics of *Camelina* crantz. (Brassicaceae) and insights on the origin of gold-of-pleasure (*Camelina sativa*). Mol. Phylogenet. Evol. 127, 834–842.

Brock, J.R., Mandáková, T., Lysak, M.A., Al-Shehbaz, I.A., 2019. *Camelina neglecta* (Brassicaceae, camelineae), a new diploid species from Europe. PhytoKeys 115, 51–57.

Bruch, A.A., Utescher, T., Mosbrugger, V., 2011. Precipitation patterns in the Miocene of Central Europe and the development of continentality. Palaeogeogr. Palaeoclimatol. Palaeoecol. 304, 202–211.

Cohen, K.M., Harper, D.A.T., Gibbard, P.L., 2018. ICS International Chronostratigraphic Chart 2019/02. International Commission on Stratigraphy, IUGS (accessed last on 2 February 2019). www.stratigraphy.org.

Cordesse, F., Cooke, R., Tremousaygue, D., Grellet, F., Delseny, M., 1993. Fine structure and evolution of the rDNA intergenic spacer in rice and other cereals. J. Mol. Evol. 36, 369–379.

Couvreur, T.L.P., Franzke, A., Al-Shehbaz, I.A., Bakker, F.T., Koch, M.A., Mummenhoff, K., 2010. Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae). Mol. Biol. Evol. 27, 55–71.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9 772–772.

Douglas, G.M., Gos, G., Steige, K.A., Salcedo, A., Holm, K., Josephs, E.B., Arunkumar, R., Ågren, J.A., Hazzouri, K.M., Wang, W., Platts, A.E., Williamson, R.J., Neuffer, B., Lascoux, M., Slotte, T., Wright, S.I., 2015. Hybrid origins and the earliest stages of diploidization in the highly successful recent polyploid *Capsella bursa-pastoris*. Proc. Natl. Acad. Sci. 112, 2806–2811.

Fazekas, A.J., Kesanakurti, P.R., Burgess, K.S., Percy, D.M., Graham, S.W., Barrett, S.C.H., Newmaster, S.G., Hajibabaei, M., Husband, B.C., 2009. Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? Mol. Ecol. Resour. 9, 130–139.

Fiedel, S., 2009. Sudden deaths: the chronology of terminal pleistocene megafaunal extinction. In: Haynes, G. (Ed.), American Megafaunal Extinctions at the End of the Pleistocene. Springer, Netherlands, Dordrecht, pp. 21–37.

Francis, A., Warwick, S.I., 2009. The biology of Canadian weeds. 142. *Camelina alyssum* (Mill.) Thell.; *C. microcarpa* Andrz. ex DC.; *C. sativa* (L.). Crantz. Can. J. Plant Sci. 89, 791–810.

Franzke, A., Hurka, H., Janssen, D., Neuffer, B., Friesen, N., Markov, M., Mummenhoff, K., 2004. Molecular signals for Late Tertiary/Early Quaternary range splits of an Eurasian steppe plant: *Clausia aprica* (Brassicaceae). Mol. Ecol. 13, 2789–2795.

Frenzel, B., 1968. The pleistocene vegetation of northern Eurasia. Science 161, 637–649.

Friesen, N., German, D.A., Hurka, H., Herden, T., Oyuntsetseg, B., Neuffer, B., 2016. Dated phylogenies and historical biogeography of *Dontostemon* and *Clausia* (Brassicaceae) mirror the palaeogeographical history of the Eurasian steppe. J. Biogeogr. 43, 738–749.

German, D.A., 2018. On the identity of some endemic SW Asian Cruciferae. I. *Thlaspi s. l.* Turczaninowia 21, 180–187.

Gomes Rodrigues, H., Marivaux, L., Vianey-Liaud, M., 2012. Expansion of open landscapes in Northern China during the Oligocene induced by dramatic climate changes: paleoecological evidence. Palaeogeogr. Palaeoclimatol. Palaeoecol. 358–360, 62–71.

Gradstein, F.M., Ogg, J.G., Schmitz, M., Ogg, G., 2012. The Geologic Time Scale 2012. Elsevier, Amsterdam.

Guo, X., Liu, J., Hao, G., Zhang, L., Mao, K., Wang, X., Zhang, D., Ma, T., Hu, Q., Al-Shehbaz, I.A., Koch, M.A., 2017. Plastome phylogeny and early diversification of Brassicaceae. BMC Genomics 18.

Hecker, U., 1987. Die Farn- und Blütenpflanzen des Mainzer Sandes. Mainzer Naturw. Archiv 25, 85–133.

Hedge, I.C., 1976. A systematic and geographical survey of the Old world Cruciferae. In: Vaughan, J.G., MacLeod, A.J., Jones, B.M.G. (Eds.), Biology and Chemistry of the Cruciferae. Academic Press, pp. 1–45.

Herden, T., Hanelt, P., Friesen, N., 2016. Phylogeny of *Allium* l. Subgenus *anguinum* (G. Don. Ex W.D.J. koch) N. Friesen (Amaryllidaceae). Mol. Phylogenet. Evol. 95, 79–93.

Hohmann, N., Wolf, E.M., Lysak, M.A., Koch, M.A., 2015. A time-calibrated road map of Brassicaceae species radiation and evolutionary history. Plant Cell tpc.15.00482.

Hovsepyan, R., Willcox, G., 2008. The earliest finds of cultivated plants in Armenia: evidence from charred remains and crop processing residues in pisé from the Neolithic settlements of Aratashen and Aknashen. Veg. Hist. Archaeobotany 17, 63–71.

Hurka, H., Friesen, N., German, D.A., Franzke, A., Neuffer, B., 2012. 'Missing link' species *Capsella orientalis* and *Capsella thracica* elucidate evolution of model plant genus *Capsella* (Brassicaceae). Mol. Ecol. 21, 1223–1238.

Hurka, H., Friesen, N., Bernhardt, K.G., Neuffer, B., Smirnov, S., Schmakov, A., Blattner, F., 2019. The Eurasian steppe belt: status quo, origin and evolutionary history. Turczaninowia 22 (3), 5–71.

Iskandarov, U., Kim, H.J., Cahoon, E.B., 2014. *Camelina*: an emerging oilseed platform for

advanced biofuels and bio-based materials. In: McCann, M.C., Buckeridge, M.S., Carpita, N.C. (Eds.), Plants and BioEnergy. Springer, New York, New York, NY, pp. 131–140.
Ivanov, D., Utescher, T., Mosbrugger, V., Syabryaj, S., Djordjević-Milutinović, D., Molchanoff, S., 2011. Miocene vegetation and climate dynamics in Eastern and Central Paratethys (Southeastern Europe). Palaeogeogr. Palaeoclimatol. Palaeoecol. 304, 262–275.
Jacobs, B.F., Kingston, J.D., Jacobs, L.L., 1999. The origin of grass-dominated ecosystems. Ann. Mo. Bot. Gard. 86, 590–643.
Jiang, H., Ding, Z., 2009. Spatial and temporal characteristics of Neogene palynoflora in China and its implication for the spread of steppe vegetation. J. Arid Environ. 73, 765–772.
Julié-Galau, S., Bellec, Y., Faure, J.-D., Tepfer, M., 2014. Evaluation of the potential for interspecific hybridization between *Camelina sativa* and related wild Brassicaceae in anticipation of field trials of GM camelina. Transgenic Res. 23, 67–74.
Kajtoch, Ł., Cieślak, E., Varga, Z., Paul, W., Mazur, M.A., Sramkó, G., Kubisz, D., 2016. Phylogeographic patterns of steppe species in Eastern Central Europe: a review and the implications for conservation. Biodivers. Conserv. 25, 2309–2339.
Koch, M.A., Haubold, B., Mitchell-Olds, T., 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). Mol. Biol. Evol. 17, 1483–1498.
Koch, M.A., 2019. The plant model system *Arabidopsis* set into an evolutionary, systematic and spatio-temporal context. J. Exp. Bot. 70, 55–67.
Koch, M.A., German, D.A., Kiefer, M., Franzke, A., 2018. Database taxonomics as key to modern plant biology. Trends Plant Sci. 23, 4–6.
Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A., Janzen, D.H., 2005. Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. 102, 8369–8374.
Kroonenberg, S.B., Rusakov, G.V., Svitoch, A.A., 1997. The wandering of the Volga delta: a response to rapid Caspian sea-level change. Sediment. Geol. 107, 189–209.
Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30, 3276–3278.
Lavrenko, E.M., 1969. Über die Lage des eurasiatischen Steppengebiets in dem System der pflanzengeographischen Gliederung des aussertropischen Eurasiens. Vegetatio 19, 11–20.
Léonard, J., 1988. Contribution à l'étude de la flore et de la végtéation des déserts d'Iran: Etude des aires de distribution les phytochories, les chorotype, Fascicules 8–9. Jardin Botanique National de Belgique, Meise (In French).
Loadmap, 2018. http://loadmap.net/ (accessed last on 2 February 2019).
Luo, Z., Brock, J., Dyer, J.M., Kutchan, T., Schachtman, D., Augustin, M., Ge, Y., Fahlgren, N., Abdel-Haleem, H., 2019. Genetic diversity and population structure of a *Camelina sativa* Spring Panel. Front. Plant Sci. 10.
Mai, D.H., 1995. Tertiäre Vegetationsgeschichte Europas: Methoden und Ergebnisse. Gustav Fischer Verlag Jena, Stuttgart.
Mamedov, A.V., 1997. The late Pleistocene-Holocene history of the Caspain sea. Quat. Int. 41 (42), 161–166.
Martin, S.L., Smith, T.W., James, T., Shalabi, F., Kron, P., Sauder, C.A., 2017. An update to the Canadian range, abundance, and ploidy of *Camelina* spp. (Brassicaceae) east of the Rocky Mountains. Botany 95, 405–417.
Meindl, C., Brune, V., Listl, D., Poschlod, P., Reisch, C., 2016. Survival and postglacial immigration of the steppe plant *Scorzonera purpurea* to Central Europe. Plant Syst. Evol. 302, 971–984.
Meng, H.-H., Gao, X.-Y., Huang, J.-F., Zhang, M.-L., 2015. Plant phylogeography in arid Northwest China: retrospectives and perspectives. J. Syst. Evol. 53, 33–46.
Meusel, H., Jäger, E., Weinert, E., 1965. Vergleichende Chorologie der Zentraleuropäischen Flora Vol. 1 German Democratic Republic, Gustav Fischer Verlag, Jena.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE). Presented at the 2010 Gateway Computing Environments Workshop (GCE), IEEE. New Orleans, LA, USA. pp. 1–8.
Mirek, Z., 1984. Monographic studies in genus *Camelina* cr. 1. *Camelina anomala* Boiss. Et Hausskn. Acta Soc. Bot. Pol. 53, 429–432.
Müller, S., Salomo, K., Salazar, J., Naumann, J., Jaramillo, M.A., Neinhuis, C., Feild, T.S., Wanke, S., 2015. Intercontinental long-distance dispersal of Canellaceae from the New to the Old World revealed by a nuclear single copy gene and chloroplast loci. Mol. Phylogenet. Evol. 84, 205–219.
Nikiforova, O.D., 1994. *Camelina* Crantz - Ryzhik. In: Malyschev, L.I., Peschkova, G.A. (Eds.), Flora Sibirica Volume VII: Berberidaceae – Grossulariaceae. Nauka, Novosibirsk, pp. 148–149.
Novikova, P.Y., Hohmann, N., Nizhynska, V., Tsuchimatsu, T., Ali, J., Muir, G., Guggisberg, A., Paape, T., Schmid, K., Fedorenko, O.M., Holm, S., Säll, T., Schlötterer, C., Marhold, K., Widmer, A., Sese, J., Shimizu, K.K., Weigel, D., Krämer, U., Koch, M.A., Nordborg, M., 2016. Sequencing of the genus *Arabidopsis* identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. Nat. Genet. 48, 1077–1082.
Novikova, P.Y., Hohmann, N., Van de Peer, Y., 2018. Polyploid *Arabidopsis* species originated around recent glaciation maxima. Curr. Opin. Plant Biol. 42, 8–15.
Orlova, L.A., Kuzmin, Y.V., Dementiev, V.N., 2004. A review of the evidence for extinction chronologies for five species of Upper Pleistocene megafauna in Siberia. Radiocarbon 46, 301–314.
Owen-Smith, N., 1987. Pleistocene extinctions: the pivotal role of megaherbivores. Paleobiology 13, 351–362.
Plenk, K., Bardy, K., Höhn, M., Thiv, M., Kropf, M., 2017. No obvious genetic erosion, but evident relict status at the westernmost range edge of the Pontic-Pannonian steppe plant *Linum flavum* L. (Linaceae) in Central Europe. Ecol. Evol. 7, 6527–6539.
POWO, 2018. Plants of the World Online portal. Facilitated by the Royal Botanic Gardens, Kew. (accessed last 2 February 2019). http://www.plantsoftheworldonline.org/.
QGIS Development Team, 2018. QGIS Geographic Information System. Open Source Geospatial Foundation Project. (accessed last on 2 February 2019). http://qgis.osgeo.org.
Qiu, Y.-X., Fu, C.-X., Comes, H.P., 2011. Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. Mol. Phylogenet. Evol. 59, 225–244.
Quan, C., Liu, Z., Utescher, T., Jin, J., Shu, J., Li, Y., Liu, Y.-S., 2014. Revisiting the Paleogene climate pattern of East Asia: a synthetic review. Earth-Sci. Rev. 139, 213–230.
Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol 67, 901–904.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Sandom, C., Faurby, S., Sandel, B., Svenning, J.-C., 2014. Global late Quaternary megafauna extinctions linked to humans, not climate change. Proc. R. Soc. B Biol. Sci. 281 20133254.
Sang, T., Crawford, D.J., Stuessy, T.F., 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84, 1120–1136.
Schultz, J., 2005. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA 11, 361–364.
Séguin-Swartz, G., Nettleton, J.A., Sauder, C., Warwick, S.I., Gugel, R.K., 2013. Hybridization between *Camelina sativa* (L.) Crantz (false flax) and North American *Camelina* species. Plant Breed. 132, 390–396.
Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L., 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am. J. Bot. 94, 275–288.
Sinskaia, E.N., Bezluzheva, A.A., 1931. The forms of *Camelina sativa* in connection with climate, flax and man. Bull. Appl. Bot. Genet. Plant Breed. 25, 89–200.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Starobogatov, Y.I., 1994. The Freshwater Zebra Mussel; Systematics, Ecology and Applications. Systematics and palaeontology, Moscow, pp. 18–46.
Stebbins, G.L.J., 1950. Variation and Evolution of Plants. Columbia University Press, New York, USA.
Stewart, J.R., Lister, A.M., Barnes, I., Dalén, L., 2010. Refugia revisited: individualistic responses of species in space and time. Proc. R. Soc. B Biol. Sci. 277, 661–671.
Strömberg, C.A.E., 2011. Evolution of grasses and grassland ecosystems. Annu. Rev. Earth Planet. Sci. 39, 517–544.
Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4.
Suh, Y., Thien, L.B., Reeve, H.E., Zimmer, E.A., 1993. Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of Ribosomal DNA in Winteraceae. Am. J. Bot. 80, 1042–1055.
Svitoch, A.A., 2012. The Caspian Sea shelf during the Pleistocene regressive epochs. Oceanology 52, 526–539.
Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.
Takhtajan, A., 1986. Floristic Regions of the World. University of California Press, Berkeley.
Tarasov, P.E., Volkova, V.S., Webb, T., Guiot, J., Andreev, A.A., Bezusko, L.G., Bezusko, T.V., Bykova, G.V., Dorofeyuk, N.I., Kvavadze, E.V., Osipova, I.M., Panova, N.K., Sevastyanov, D.V., 2000. Last glacial maximum biomes reconstructed from pollen and plant macrofossil data from northern Eurasia. J. Biogeogr. 27, 609–620.
Tedin, O., 1925. Vererbung, Variation und Systematik in der Gattung Camelina. Hereditas 6, 275–386.
The Nature Conservancy, 2018. http://maps.tnc.org/gis_data.html (accessed last on 2 February 2019).
TopGlobus, 2018. http://www.topglobus.ru/ (accessed last on 2 February 2019).
Tudryn, A., Chalié, F., Lavrushin, Y.A., Antipov, M.P., Spiridonova, E.A., Lavrushin, V., Tucholka, P., Leroy, S.A.G., 2013. Late Quaternary Caspian Sea environment: Late Khazarian and Early Khvalynian transgressions from the lower reaches of the Volga River. Quat. Int. 292, 193–204.
Tutin, T.G., Burges, N.A., Chater, A.O., Edmondson, J.R., Heywood, V.H., Moore, D.M., Valentine, D.H., Walters, S.M., Webb, D.A., 1993. Flora Europaea Volume I, 2nd Edition: Psilotaceae to Platanaceae. Cambridge University Press, New York.
Van Baak, C.G.C., Vasiliev, I., Stoica, M., Kuiper, K.F., Forte, A.M., Aliyeva, E., Krijgsman, W., 2013. A magnetostratigraphic time frame for Plio-Pleistocene transgressions in the South Caspian Basin, Azerbaijan. Glob. Planet. Change 103, 119–134.
Varga, Z., 2010. Extra-Mediterranean refugia, post-glacial vegetation history and area dynamics in Eastern Central Europe. Relict Species. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 57–87.
Vassilczenko, I.T., 1939. Rod 687. Ryzhik – *Camelina* Crantz. In: Komarov, V.L. (Ed.), Flora of the USSR, Volume VIII: Capparidaceae, Cruciferae & Resedaceae, Izdatel'stvo Akademii Nauk SSSR. Moskva – Leningrad. Union of Soviet Socialist Republics, pp. 596–602.
Vavilov, N.I., 1992. Origin and Geography of Cultivated Plants. Cambridge University Press, United Kingdom.
Velichko A. A., Akhlestina E. F., Borisova O. K., Gribchenko Yu. I., Zhidovinov N. Yu., Zelikson E. M., Iosifova V. P., Klimanov V. A., Morozova T. V., Nechaev V. P., Pisareva V. V., Svetlitskaya T. V., Spasskaya I. I., Udartsev V. P., Faustova M. A., Shik S. M. 1999. East European plains. In: Climate and Environment Changes during the

last 65 Million Years; Cenozoic: from Paleocene to Holocene. Eds. A. A. Velichko, V. P. Nechaev. Chapter 3. GEOS, Moscow, 43–83 pp. [In Russian]. (Величко А. А., Ахлестина Е. Ф., Борисова О. К., Грибченко Ю. Н., Жидовинов Н. Я., Зеликсон Э. М., Иосифова Ю. И., Климанов В. А., Морозова Т. Д., Нечаев В. П., Писарева В. В., Свет-лицкая Т. В., Спасская И. И., Ударцев В. П., Фаустова М. А., Шик С. М. Восточно-Европейская равнина //Изменение климата и ландшафтов за последние 65 миллионов лет (кайнозой: от палеоцена до голоцена. Под ред. А. А. Величко, В. П. Нечаева. М.: ГЕОС, 1999. С. 43–83).

Volkova, P.A., Herden, T., Friesen, N., 2017. Genetic variation in *Goniolimon speciosum* (Plumbaginaceae) reveals a complex history of steppe vegetation. Bot. J. Linn. Soc. 184, 113–121.

Wikimapia, 2018. http://www.wikimapia.org/ (accessed last on 2 February 2019).

World Wildlife Fund for a living planet, 2018. World Wildlife Fund for a Living Planet. (accessed last on 2 February 2019). https://web.archive.org/web/20110401124312/http://wwf.panda.org:80/about_our_earth/ecoregions/about/habitat_types/selecting_terrestrial_ecoregions/habitat08.cfm.

Wright, S.D., Yong, C.G., Wichman, S.R., Dawson, J.W., Gardner, R.C., 2003. Stepping stones to Hawaii: a trans-equatorial dispersal pathway for *Metrosideros* (Myrtaceae) inferred from nrDNA (ITS + ETS): stepping stones to Hawaii. J. Biogeogr. 28, 769–774.

Yang, S., Ding, Z., 2010. Drastic climatic shift at ~2.8Ma as recorded in eolian deposits of China and its implications for redefining the Pliocene-Pleistocene boundary. Quat. Int. 219, 37–44.

Yıldırımlı, Şinasi, 2011. Three new species from Turkey. Ot Sist. Bot. Derg. 18, 2–5.

Yu, Y., Harris, A.J., Blair, C., He, X., 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. Mol. Phylogenet. Evol. 87, 46–49.

Zinger, H.B., 1909. On the species of *Camelina* and *Spergularia* occuring as weeds in sowings of flax and their origin. Tr. Bot. Muz. Akad. Nauk USSR 6, 1–303.

Furtherreading

Brummitt, R.K., Powell, C.E. (Eds.), 1992. Authors of Plant Names: a List of Authors of Scientific Names of Plants, With Recommended Standard Forms of Their Names, Including Abbreviations. Royal Botanic Gardens, Kew, London.