A genomic dataset of single-nucleotide polymorphisms generated by ddRAD tag sequencing in Quercus petraea (Matt.) Liebl. populations from Central-Eastern Europe and Balkan Peninsula

Endre Gy. Tóth1 · Zoltán A. Köbölkuti1 · Klára Cseke1 · József D. Kámpel1 · Roland Takács1 · Vladimir T. Tomov2 · Péter Ábrán3 · Srdjan Stojnic4 · Erna Vastag5 · Milan Mataruga6 · Vanja Daničić6 · Egzon Tahirukaj7 · Petar Zhelev2 · Sasa Orlovic4 · Attila Benke1 · Attila Borovics1

Received: 28 July 2020 / Accepted: 26 February 2021 / Published online: 19 April 2021 © The Author(s) 2021

Abstract

Key message This genomic dataset provides highly variable SNP markers from georeferenced natural Quercus petraea (Matt.) Liebl. populations collected in Bulgaria, Hungary, Romania, Serbia, Bosnia and Herzegovina, Kosovo* and Albania. These SNP loci can be used to assess genetic diversity, differentiation, and population structure, and can also be used to detect signatures of selection and local adaptation. The dataset can be accessed at https://doi.org/10.5281/zenodo.3908963 (Tóth et al. 2020). Associated metadata available at https://metadata-afs.nancy.inra.fr/geonetwork/srv/fre/catalog-search#/metadata/b6fee4fa-01e9-44d0-92f5-ad19379f9693.

Keywords Single-nucleotide polymorphism · Double digest · RAD-seq · Sessile oak · Genetic variation

Handling Editor: Marianne Peiffer

Contribution of the co-authors EGT, ZAK, KC, ABE, and ABO conceived the study. EGT, ZAK, KC, ABE, JDK, RT, VT, PA, SS, EV, MM, VD, ET, PZ, SO, and ABE contributed in sampling of plant materials. KC performed laboratory analysis. EGT and ZAK processed genomic data and performed statistical analyses and simulations. EGT wrote the manuscript. All co-authors provided feedback on the manuscript drafts.

This designation is without prejudice to positions on status, and is in line with UNSCR 1244 (1999) and the ICJ Opinion’s on the Kosovo* Declaration of Independence.

Endre Gy. Tóth
toth.endre@uni-sopron.hu
Zoltán A. Köbölkuti
kobolkuti.zoltan@uni-sopron.hu
Klára Cseke
cseke.klara@uni-sopron.hu
József D. Kámpel
kampel.jozsef@uni-sopron.hu
Roland Takács
takacs.roland@uni-sopron.hu
Vladimir T. Tomov
vladimirtomov@ltu.bg
Péter Ábrán
p_abran@yahoo.com
Srdjan Stojnic
srdjan.stojnic@uns.ac.rs
Erna Vastag
erna.vastag@uns.ac.rs
Milan Mataruga
milan.mataruga@sf.unibl.org
Vanja Daničić
vanja.danicic@sf.unibl.org
1 Background

In the last decade, restriction site–associated DNA (RAD-seq) sequencing is increasingly used to identify and genotype large numbers of single-nucleotide polymorphisms (SNPs) in various organisms (Allendorf et al. 2010; Davey and Blaxter 2010). Due to the advancements of high-throughput sequencing, this approach allows to produce extremely large collections of genomic data that is fairly evenly distributed across the genome providing hundreds of thousands of polymorphic loci (Allendorf et al. 2010; Seeb et al. 2011). A variation of the approach, known as ddRAD-seq (double-digest RAD-seq), has been widely applied for population genomic analyses in non-model organisms, such as forest tree species, lacking available reference genomes (Konar et al. 2017). ddRAD-seq is suitable for genotyping a large number of loci to determine the extent of genetic diversity, to identify outlier loci being under selection (Schwartz et al. 2010; Burak et al. 2018) and to infer eco-evolutionary population dynamics (Legrand et al. 2017; Hendry 2019).

While forest tree populations are well-studied in Europe, the Central-Eastern European region, including the Balkan Peninsula is less investigated with high-resolution and genome-wide genetic markers. In fact, this region has high topographic and climatic heterogeneity, and long-term environmental stability that significantly contributed to long-term persistence of genetic diversity (Tzedakis 2004; Feliner 2011; Fassou et al. 2020). In addition, the Balkan Peninsula has been postulated to be the primary source of post-glacial expansion while functioning as a Pleistocene glacial refugia (Zhelev 2017; Gömöry et al. 2020). Therefore, this area is considered an important reservoir of genetic resources of European forest tree taxa (Gömöry et al. 2020).

Sessile oak (*Quercus petraea* (Matt.) Liebl.) is a widely distributed late-successional forest tree species of the white oak group (sect. Lepidobalanus) with great ecological and silvicultural importance in Europe (Mölder et al. 2019; Blanc-Jolivet et al. 2020). However, the species’ distribution towards the southern limit on the Balkan Peninsula tends to be fragmented (Bruschi et al. 2003; Eaton et al. 2016), and being under intensive forest management and severe ecosystem disturbances such as drought (Milad et al. 2011). Similarly, species’ intraspecific genetic differentiation increases towards the southeast due to its complex evolutionary history. Since current studies mostly focusing on western European populations, the southeastern European distribution range is poorly sampled and investigated (Lang et al. 2020; Leroy et al. 2017, 2020; Kremer and Hipp 2020). This inequality might potentially lead to underestimation of regional genetic diversity. For these reasons, we generated a novel resource of genome-wide SNPs, using ddRAD-seq, in carefully selected natural populations.

2 Methods

2.1 Sampling strategy

The sampling of plant material was designed to cover the southeastern natural distribution range of *Q. petraea*. Sites were selected based on literature data and personal communication with local forestry experts (contributing authors). For literature data, available information on neutral genetic patterns in European white oaks was reviewed (Dumolin-Lapegue et al. 1997; Gömöry et al. 2001, 2020; Bordács et al. 2002; Csaikl et al. 2002; Petit et al. 2002a, b; Slade et al. 2008; Dering et al. 2008), and the geographic regions corresponding to the major genetic clusters were included in a candidate list of locations. Existing databases of forest genetic resources also contributed in
sampling site selection. The European Information System on Forest Genetic Resources (EUFGIS, 2019; http://portal.eufgis.org/), the georeferenced database of forest conservation units, and the European Forest Reproductive Material Information System (FOREMATIS, 2019; https://ec.europa.eu/forematis) were searched and filtered, and the records of “source-identified” and “autochthonous/indigenous” locations were taken into account. Altogether, 180 individuals from 18 autochthonous populations were selected from Central-Eastern Europe and from the Balkan Peninsula (Table 1, Fig. 1). Out of these, four populations situated in Bulgaria (western Balkan, Rila, Rhodope, and Strandzha mountains), three in Hungary (Kőszeg, Bakony, and Mecsek mountains), Romania (Gurghiu, northern and southern foothills of Fagaras mountains), Serbia (Fruska Gora, Rudnik, and Stovoli mountains), Bosnia and Herzegovina (Kozara, Javorova, and Maglic mountains), and one in Kosovo* (Blinaja) and Albania (Djeravica Mt.), respectively (Tóth et al. 2020). Our sampling sites are found in natural sites either as part of nature reserves or without non-natural disturbance and human influence. Accordingly, sessile oak populations were regenerated naturally. Fresh leaves were collected from even-aged mature trees (> 70-year-old) by considering at least a 30-m isolation distance between trees to limit sampling-related individuals. Spatial coordinates (GPS) of each tree in each population were recorded. Leaf samples were both dried using silica gel and stored frozen (−20 °C) in plastic bags before DNA extraction.

2.2 Library preparation and RAD-tag sequencing

Total genomic DNA was extracted from seven leaf disks (5-mm diam.) per each individual using the method of Dumolin et al. (1995).

DNA of each Q. petraea sample was quantified by using the Qubit dsDNA BR Assay Kit and Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). A total of 50 ng of DNA per sample was double-digested with 0.1–0.1 μl of PstI and MspI enzymes at 37 °C for 2 h (FastDigest restriction enzymes; Thermo Fisher Scientific, Waltham, MA, USA). The enzyme combination was selected based on the study of Cumer et al. (2018) and a preliminary restriction site analysis carried out in CLC Genomic Workbench version 12.0 (QIAGEN Bioinformatics, Hilden, Germany). Fragments were double-sided size selected using KAPA PureBeads, with a 0.55–0.80X solution/bead ratio (Roche, Basel, Switzerland) in order to isolate fragments in the range of 300–600 bp. After quantification, inserts (3 ng) were ligated to adapters (Table 2) with using T4 DNA Ligase according to the manufacturers’ protocol (Thermo Fisher Scientific, Waltham, MA, USA).

Table 1 Summary information on sampling locations of Q. petraea populations

| Pop | Country          | Municipality   | Geographic region | Latitude | Longitude | Elevation (m a.s.l) |
|-----|------------------|----------------|-------------------|----------|-----------|--------------------|
| AL1 | Albania          | Bajram Curri   | Djeravica Mt      | 42.407   | 20.168    | 454                |
| BH1 | Bosnia and Herzegovina | Kozarska Dubica | Kozara Mts       | 45.157   | 16.979    | 311                |
| BH2 | Bosnia and Herzegovina | Teslic         | Javorova Mts     | 44.662   | 17.706    | 440                |
| BH3 | Bosnia and Herzegovina | Foca           | Maglic Mt        | 43.476   | 18.946    | 985                |
| BU1 | Bulgaria         | Botevgrad      | Balkan Mts       | 42.980   | 23.826    | 601                |
| BU2 | Bulgaria         | Samokov        | Rila Mts         | 42.353   | 23.686    | 950                |
| BU3 | Bulgaria         | Velingrad      | Rhodope Mts      | 42.073   | 23.977    | 1172               |
| BU4 | Bulgaria         | Balgari        | Strandzha Mts    | 42.117   | 27.766    | 211                |
| HU1 | Hungary          | Kőszeg         | Kőszeg Mts       | 47.371   | 16.524    | 351                |
| HU2 | Hungary          | Vállus         | Bakony Mts       | 46.835   | 17.319    | 374                |
| HU3 | Hungary          | Zengővárkony   | Mecsek Mts       | 46.200   | 18.432    | 372                |
| KO1 | Kosovo*          | Lipjan         | Blinaja          | 42.513   | 20.968    | 787                |
| RO1 | Romania          | Periş          | Gurghiu Mts      | 46.694   | 24.692    | 438                |
| RO2 | Romania          | Şura Mică      | Fagaras Mts      | 45.839   | 24.007    | 508                |
| RO3 | Romania          | Ciocănăi       | Fagaras Mts      | 44.853   | 24.793    | 455                |
| SE1 | Serbia           | Fruška Gora    | Fruška Gora Mts  | 45.172   | 19.839    | 208                |
| SE2 | Serbia           | Kragujevac     | Rudnik Mts       | 44.172   | 20.488    | 578                |
| SE3 | Serbia           | Kraljevo       | Stovoli Mts      | 43.667   | 20.603    | 414                |

*This designation is without prejudice to positions on status, and is in line with UNSCR 1244 (1999) and the ICJ Opinion’s on the Kosovo* Declaration of Independence.
Ligation products were purified using 0.8 vol KAPA PureBeads (Roche, Basel, Switzerland) and amplified by PCR with NEBNext Multiplex Oligos for Illumina (Dual Index Set 1; New England Biolabs, Ipswich, MA, USA) and KAPA HiFi Hotstart Ready Mix (Roche, Basel, Switzerland). The amount of 0.5–0.5 μl of i5 and i7 indexed primers was used per reaction. Thermal cycling conditions were as follows: a 3-min initial denaturation at 95 °C; 17 cycles of 30 s of denaturation at 95 °C, 30 s of annealing at 55 °C, and 30 s extension at 72 °C; and a final 5-min extension at 72 °C. The quality and quantity of the amplification library were determined by using High Sensitivity DNA1000 ScreenTape system with 2200 Tapestation (Agilent Technologies, Santa Clara, CA, USA) and dsDNA HS Assay Kit with Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), respectively. Equimolarly pooled libraries were diluted to 10 pM for 2 x 301-bp paired-end sequencing with 600-cycle sequencing kit v3.1 (Illumina, San Diego, CA, USA). Nucleotide sequences of the libraries were determined on a MiSeq Sequencing System (Illumina, San Diego, CA, USA) according to the manufacturer’s protocol. Raw data have been deposited in the NCBI Sequence Read Archive (SRA); BioProject ID: PRJNA699096.

2.3 De novo assembly and SNP calling

All bioinformatics processing was carried out on a Silicon Computers (SGI) HPC server, allocating 40 cores (80 threads) and 38 GB RAM, located at the University of Sopron (UOS), Sopron, Hungary.

| Restriction enzyme | Adapter sequence (5’–3’) |
|--------------------|--------------------------|
| PstI               | TCTTTCCCCTACAGACGCTTTCCCTAGCTGCA |
|                    | GATCGGAAGAGCGTCGTAAGGGAAAGAGTTG |
| MspI               | CTTGAATTCAGCAGTGTTCGCCGATC |
|                    | CGAGATCGGAAGAGCAGCACGTCCTGAACTCCAGT |

Fig. 1 Distribution of *Q. petraea* sampling locations across Europe (orange triangles). Light green areas represent the natural distribution of the species according to EUFORGEN database (http://www.euforgen.org/).
Raw short-read sequences (77,101,088) were demultiplexed and adapter-trimmed by using the MiSeq Control Software (Illumina, San Diego, CA, USA). In total, 69,993,001 sequences contained adapter sequences kept for further processing. Next, the implemented FastQ Toolkit was applied to trim bases at the 3′ and the 5′ end with a quality score lower than 30. Reads having a mean quality score less than 30 and shorter than 200 bp were filtered. Computational processing of short-read data was carried out with Stacks 2.0 (Catchen et al. 2013; Rochette et al. 2019). Whole reads were quality-filtered using a sliding-window method (15% of read length) implemented in “process_radtags.” Reads having a quality score below 90% (raw phred score of 10) were discarded (Catchen et al. 2011). In “process_radtags,” reads were truncated to 200 bp as a prerequisite of further processing and to avoid the lower-quality bases to present at the end of the reads (Catchen et al. 2011). During this filtering step, 42,273 sequences were discarded. In addition, one individual was removed from the dataset (BU2-10) since an insufficient number of high-quality reads remained after filtering. RAD loci were reconstructed following the de novo pipeline implemented in “denovo_map.pl” command, in which “ustacks” builds loci and calls SNPs in each sample, then “cstacks” creates a catalogue of all loci for all the samples and “sstacks” matches the loci of the samples against the catalogue (Catchen et al. 2011; Rochette et al. 2019). To execute “denovo_map.pl,” it is essential to optimize the parameters of the command, and define (n) the minimum number of reads required to form an allele, (M) the maximum number of mismatches allowed between two alleles, and (M) the maximum number of mismatches allowed between two individual loci to consider them as homologous (Mastretta-Yanes et al. 2015; Paris et al. 2017). We optimized these parameters throughout the “r80” method, which can effectively maximize the number of SNPs found in 80% of the individuals (Paris et al. 2017). We checked the iterative values (ranging from 2 to 10) of M and n how they affect the gained SNPs (Fig. 2a). After, based on this, we chose the threshold of $M = 3$, $n = 3$, applying the $M = n$ rule for the final run (Paris et al. 2017). For the stack-depth parameter, we selected the default $m = 3$ value (three identical reads). The program “gstacks” was used to assemble paired-end contigs and re-call SNPs using the population-wide data (Rochette et al. 2019).

The “denovo_map.pl” pipeline aligned paired-end reads for a total of 112,750 RAD loci in which 1490 loci had paired-end reads that could not be assembled (1.3%). The remaining 111,260 loci (98.7%) had an average contig size of 361.8 bp. The effective per-sample coverage resulted in 14.3× (mean) with a 2.2× standard deviation (min = 10.1×, max = 27.0×) (Fig. 2b). Population-specific statistics are presented in Table 3.

The set of SNPs were called and further processed with the “populations” program (Catchen et al. 2011), in which markers with a minor allele frequency < 0.05, missing individual rate > 0.8, and significant deviation from Hardy–Weinberg equilibrium (HWE, $P < 0.001$) were filtered out (Catchen et al. 2011). However, the “minimum number of populations” value was set to 1 (SNPs were kept when they were present in at least one population), allowing to be later filtered by the user for the specific research question. Out of the 111,260 RAD loci, 105,326 loci did not pass the missing rate of sample/population and were below the MAF threshold; also, 1099 loci were blacklisted and discarded having variants with significant HWE exception. The final dataset consisted of 5934 loci, composed of 2,443,502 bases, of which 21,951 are highly polymorphic variant sites (SNPs).

3 Access to data and metadata description

The SNP dataset is available from the ZENODO repository: https://doi.org/10.5281/zenodo.3908963 (Tóth et al. 2020). The data records are described in the metadata description files, available at https://metad ata-afs.nancy.inra.fr/geonetwork/srv/fre/catalog.search/#/metadata/b6fee4fa-01e9-44d0-92f5-ad19379f9693. The data file presented in this paper is a PLINK-formatted UTF-8 encoded, white-space (tab)-delimited, set of ped and map files. PLINK is a publicly available and widely used software for genomic data manipulation and analysis, capable of converting to other file formats and can perform several population-based tests (Purcell et al. 2007). Within the “.ped” file, the first column corresponds to the population code (Pop_ID: 1–18), the second to the individual identifier. From column three to six, values are not provided (father ID, mother ID, sex, and phenotype are not defined). From column seven, two columns code an observed allele at each SNP position. Missing data are coded as “0 0” as it is defined in the standard PLINK data format (PLINK 1.07; http://zzenodo.org/4fa92f5-ad19379f9693). The “.map” file contains four columns, the first is the chromosome number (un: unknown due to de novo assembly), the second is the SNP ID, the third is the SNP genetic position (0: not defined), and the fourth column contains the physical position (bp) within the RAD loci. Users do not need to combine or merge the files; it is ready to be loaded and
Fig. 2  a Plot of the number of new polymorphic loci (80 loci) gained for each iteration of $M$ (the distance between stacks) and $n$ (the maximum number of mismatches allowed between loci) of the preliminary test run of the “denovo_map.pl” pipeline in Stacks (Catchen et al. 2013; Rochette et al. 2019). The orange dot corresponds to the selected value. b Box plot of mean coverage of reads for each population (including all individuals).

Table 3 Detected polymorphisms and genomic coverage of *Q. petraea* populations

| Pop | Samples per locus | All sites | Variant sites | Polymorphic sites | No. of private alleles | Mean coverage (SD) | No. of reads (%) |
|-----|------------------|-----------|---------------|------------------|-----------------------|-------------------|------------------|
| AL1 | 9.1795           | 1,349,961 | 10,362        | 8681             | 61                    | 11.21 × (18.35)   | 101,988.5 (58.92%) |
| BH1 | 9.2369           | 1,458,812 | 11,546        | 9688             | 118                   | 13.43 × (14.81)   | 134,311.4 (59.60%) |
| BH2 | 9.1301           | 1,063,481 | 8473          | 7022             | 33                    | 13.84 × (18.44)   | 121,266.8 (60.83%) |
| BH3 | 9.1825           | 1,144,541 | 9160          | 7511             | 31                    | 12.61 × (13.54)   | 111,688.9 (60.78%) |
| BU1 | 9.1984           | 1,464,403 | 11,246        | 9498             | 58                    | 11.51 × (15.99)   | 109,226.2 (60.07%) |
| BU2 | 8.562            | 771,154   | 5904          | 4781             | 12                    | 12.03 × (15.32)   | 108,617.92 (59.72%) |
| BU3 | 9.1402           | 1,232,776 | 9960          | 8054             | 49                    | 12.11 × (19.25)   | 110,587.2 (55.93%) |
| BU4 | 9.1919           | 1,283,326 | 10,168        | 7139             | 96                    | 12.76 × (17.53)   | 116,093.4 (59.42%) |
| HU1 | 9.2404           | 1,184,791 | 9285          | 7142             | 54                    | 11.78 × (12.99)   | 98,395.2 (59.46%) |
| HU2 | 9.1575           | 1,138,892 | 9043          | 7589             | 53                    | 11.56 × (16.01)   | 98,240.1 (55.80%) |
| HU3 | 9.1747           | 1,454,340 | 11,267        | 9476             | 69                    | 11.85 × (12.78)   | 107,480.7 (63.90%) |
| KO1 | 9.2007           | 1,493,052 | 11,508        | 9749             | 62                    | 11.76 × (14.99)   | 113,011.8 (59.98%) |
| RO1 | 9.0935           | 1,334,593 | 10,036        | 8369             | 37                    | 10.82 × (17.47)   | 105,582.6 (54.86%) |
| RO2 | 9.1707           | 1,412,779 | 10,747        | 9029             | 42                    | 11.84 × (20.97)   | 110,162.2 (56.85%) |
| RO3 | 9.2393           | 1,479,964 | 11,691        | 9837             | 72                    | 12.26 × (16.73)   | 119,783.8 (58.06%) |
| SE1 | 9.2025           | 1,542,620 | 11,886        | 10,255           | 63                    | 12.12 × (13.00)   | 128,291.5 (59.53%) |
| SE2 | 9.2263           | 1,593,744 | 12,442        | 10,481           | 102                   | 12.51 × (17.42)   | 123,251.7 (61.60%) |
| SE3 | 9.2137           | 1,469,120 | 11,649        | 9692             | 59                    | 11.91 × (16.35)   | 111,958.5 (61.31%) |
analyzed. Population IDs (Pop_ID), individual IDs (Ind_ID), and the corresponding GPS coordinates (GPS_1: latitude and GPS_2: longitude) are located in the Appendix Table 5. Further information can be requested from the corresponding author.

4 Technical validation

Validation of the dataset was performed with manual, visual, and numerical tests. The quality of raw short-read data, of all individuals (179), was assessed using FastQC v0.11.9 (Andrews 2010) by investigating per-base quality and per-sequence quality at three steps: at raw read state directly after sequencing, after sequence quality trimming (3′ and 5′ end), and following sequence processing (whole read filtering). Results were consistent with a standard Illumina run, and therefore considered to be of sufficiently high quality for further analysis (Kircher et al. 2011).

5 Reuse potential and limits

Data papers on genome-wide polymorphisms are extremely scarce especially on geographic regions that are sources of biodiversity and in case of forest trees lacking a reference genome. Our data can significantly contribute to the growing number of RAD-seq studies by providing genotype information on species’ rarely investigated southeastern distribution range.

To evaluate the potential of the data, at the population level, we calculated genetic diversity indices including the expected heterozygosity (\(H_e\)), observed heterozygosity (\(H_o\)), and the inbreeding coefficient (\(F_{IS}\)) with R (R Core Team 2013) using the “adegenet” (Jombart and Ahmed 2011) and “hierfstat” (Goudet 2005) packages (Table 4). Trends of our resulting indices were compared to Blanc-Jolivet et al. (2020), since up until now, no SNP dataset (i.e., SNP genotype collection) from genomic resources has so far been published for Q. petraea. Since our estimators are in accordance with the Blanc-Jolivet et al. (2020) data, as well as with other publications, we can confirm that our panel of markers is suitable for population genomic studies.

It is important to note, however, that polymorphisms from reduced representation methods, such as RAD-seq, might introduce biases in population genetics estimates (Arnold et al. 2013; Andrews et al. 2016), and phylogenetic reconstructions and genomic mappings might require even larger genomic coverage as well as higher SNP density (Sims et al. 2014; Eaton et al. 2017). Despite of these, our generated high-quality SNP resource consisting of 21,951 SNPs provides a valuable tool for future population genetics and genomics studies allowing to carry out wide variety of reference-genome-free investigations (Peterson et al. 2012). The available spatial information (GPS coordinates; Table S1) allows the analyses of spatial genetic pattern, structure (including admixture), and gene flow and hybridization. Due to the relatively large-scale genome representation, fine-scale determination of relationship among individuals and populations is also possible. For example, in the context of phyleogeography, our genotype data may supplement studies on evolutionary history, migration from glacial refugia, and past hybridization events increasingly studied nowadays (Leroy et al. 2017, 2020). Since ddRAD-seq has the potential to generate strongly differentiating markers, our data could be used to investigate patterns of divergent selection processes, thus can infer population-level adaptation.

Overall, this dataset provides a unique opportunity to study the genetic background of natural populations and reveals important ecological and evolutionary insights into the present and future distribution of sessile oak.

| Table 4 Diversity statistics calculated for each Q. petraea population during the data validation process |
|----------------------------------|----------|----------|----------|
| Pop  | \(n\)  | \(H_o\)  | \(H_e\)  | \(F_{IS}\) |
|------|-------|----------|----------|------------|
| AL1  | 10    | 0.221    | 0.249    | 0.092\(^*\) |
| BH1  | 10    | 0.216    | 0.249    | 0.108\(^*\) |
| BH2  | 10    | 0.217    | 0.250    | 0.103\(^*\) |
| BH3  | 10    | 0.218    | 0.248    | 0.097\(^*\) |
| BU1  | 10    | 0.221    | 0.249    | 0.085\(^*\) |
| BU2  | 10    | 0.219    | 0.247    | 0.085\(^*\) |
| BU3  | 9     | 0.207    | 0.241    | 0.109\(^*\) |
| BU4  | 10    | 0.189    | 0.216    | 0.104\(^*\) |
| HU1  | 10    | 0.209    | 0.235    | 0.089\(^*\) |
| HU2  | 10    | 0.214    | 0.249    | 0.113\(^*\) |
| HU3  | 10    | 0.221    | 0.251    | 0.093\(^*\) |
| KO1  | 10    | 0.221    | 0.252    | 0.097\(^*\) |
| RO1  | 10    | 0.220    | 0.246    | 0.082\(^*\) |
| RO2  | 10    | 0.219    | 0.248    | 0.089\(^*\) |
| RO3  | 10    | 0.219    | 0.251    | 0.106\(^*\) |
| SE1  | 10    | 0.223    | 0.254    | 0.100\(^*\) |
| SE2  | 10    | 0.215    | 0.248    | 0.107\(^*\) |
| SE3  | 10    | 0.222    | 0.250    | 0.088\(^*\) |

\(N\) number of individuals, \(H_o\) observed heterozygosity (Nei 1978), \(H_e\) expected heterozygosity (Nei 1978), \(F_{IS}\) inbreeding coefficient (ns, non-significant; Nei 1987)
## Appendix

Table 5 Identifiers of individuals (Ind_ID) and populations (Pop_ID), and their corresponding GPS coordinates (GPS_1: latitude and GPS_2: longitude)

| Ind_ID  | GPS_1   | GPS_2   |
|---------|---------|---------|
| AL1-1   | 42.4070 | 20.1680 |
| AL1-2   | 42.4071 | 20.1685 |
| AL1-3   | 42.4067 | 20.1685 |
| AL1-4   | 42.4072 | 20.1679 |
| AL1-5   | 42.4073 | 20.1685 |
| AL1-6   | 42.4069 | 20.1685 |
| AL1-7   | 42.4070 | 20.1681 |
| AL1-8   | 42.4074 | 20.1683 |
| AL1-9   | 42.4073 | 20.1686 |
| AL1-10  | 42.4070 | 20.1688 |
| QPET-BH1-1 | 45.1584 | 16.9805 |
| QPET-BH1-2 | 45.1583 | 16.9808 |
| QPET-BH1-3 | 45.1582 | 16.9806 |
| QPET-BH1-4 | 45.1581 | 16.9805 |
| QPET-BH1-5 | 45.1571 | 16.9794 |
| QPET-BH1-6 | 45.1572 | 16.9790 |
| QPET-BH1-7 | 45.1574 | 16.9795 |
| QPET-BH1-8 | 45.1577 | 16.9795 |
| QPET-BH1-9 | 45.1603 | 16.9794 |
| QPET-BH1-10 | 45.1605 | 16.9791 |
| QPET-BH2-1 | 44.6607 | 17.7058 |
| QPET-BH2-2 | 44.6619 | 17.7057 |
| QPET-BH2-3 | 44.6620 | 17.7062 |
| QPET-BH2-4 | 44.6622 | 17.7062 |
| QPET-BH2-5 | 44.6623 | 17.7064 |
| QPET-BH2-6 | 44.6624 | 17.7067 |
| QPET-BH2-7 | 44.6623 | 17.7069 |
| QPET-BH2-8 | 44.6623 | 17.7068 |
| QPET-BH2-9 | 44.6623 | 17.7069 |
| QPET-BH2-10 | 44.6626 | 17.7064 |
| QPET-BH3-1 | 43.4763 | 18.9460 |
| QPET-BH3-2 | 43.4764 | 18.9465 |
| QPET-BH3-3 | 43.4766 | 18.9465 |
| QPET-BH3-4 | 43.4765 | 18.9464 |
| QPET-BH3-5 | 43.4765 | 18.9458 |
| QPET-BH3-6 | 43.4763 | 18.9460 |
| QPET-BH3-7 | 43.4763 | 18.9463 |
| QPET-BH3-8 | 43.4765 | 18.9466 |
| QPET-BH3-9 | 43.4764 | 18.9469 |
| QPET-BH3-10 | 43.4768 | 18.9466 |
| BU1-1   | 42.9798 | 23.8259 |
| BU1-2   | 42.9794 | 23.8259 |
| BU1-3   | 42.9797 | 23.8254 |
| BU1-4   | 42.9795 | 23.8255 |
| BU1-5   | 42.9800 | 23.8253 |
| BU1-6   | 42.9802 | 23.8256 |
| BU1-7   | 42.9802 | 23.8256 |
| Ind_ID     | GPS_1   | GPS_2   |
|-----------|---------|---------|
| BU1-8     | 42.9803 | 23.8255 |
| BU1-9     | 42.9803 | 23.8254 |
| BU1-10    | 42.9804 | 23.8254 |
| QPET-BU2-1| 42.3534 | 23.6863 |
| QPET-BU2-2| 42.3535 | 23.6863 |
| QPET-BU2-3| 42.3536 | 23.6864 |
| QPET-BU2-4| 42.3536 | 23.6865 |
| QPET-BU2-5| 42.3538 | 23.6868 |
| QPET-BU2-6| 42.3538 | 23.6866 |
| QPET-BU2-7| 42.3537 | 23.6864 |
| QPET-BU2-8| 42.3537 | 23.6862 |
| QPET-BU2-9| 42.3537 | 23.6857 |
| QPET-BU3-1| 42.0727 | 23.9771 |
| QPET-BU3-2| 42.0726 | 23.9769 |
| QPET-BU3-3| 42.0725 | 23.9768 |
| QPET-BU3-4| 42.0724 | 23.9768 |
| QPET-BU3-5| 42.0722 | 23.9767 |
| QPET-BU3-6| 42.0721 | 23.9767 |
| QPET-BU3-7| 42.0720 | 23.9768 |
| QPET-BU3-8| 42.0718 | 23.9767 |
| QPET-BU3-9| 42.0717 | 23.9768 |
| QPET-BU3-10| 42.0716 | 23.9771 |
| QPET-BU4-1| 42.1173 | 27.7659 |
| QPET-BU4-2| 42.1174 | 27.7662 |
| QPET-BU4-3| 42.1172 | 27.7664 |
| QPET-BU4-4| 42.1171 | 27.7663 |
| QPET-BU4-5| 42.1169 | 27.7661 |
| QPET-BU4-6| 42.1166 | 27.7658 |
| QPET-BU4-7| 42.1164 | 27.7656 |
| QPET-BU4-8| 42.1163 | 27.7658 |
| QPET-BU4-9| 42.1161 | 27.7659 |
| QPET-BU4-10| 42.1158| 27.7658 |
| QPET-HU1-1| 47.3718 | 16.5246 |
| QPET-HU1-2| 47.3718 | 16.5258 |
| QPET-HU1-3| 47.3716 | 16.5251 |
| QPET-HU1-4| 47.3712 | 16.5252 |
| QPET-HU1-5| 47.3710 | 16.5246 |
| QPET-HU1-6| 47.3714 | 16.5242 |
| QPET-HU1-7| 47.3719 | 16.5245 |
| QPET-HU1-8| 47.3714 | 16.5242 |
| QPET-HU1-9| 47.3719 | 16.5245 |
| QPET-HU1-10| 47.3715| 16.5242 |
| QPET-HU2-1| 46.8349 | 17.3188 |
| QPET-HU2-2| 46.8350 | 17.3170 |
| QPET-HU2-3| 46.8365 | 17.3179 |
| QPET-HU2-4| 46.8355 | 17.3197 |
| QPET-HU2-5| 46.8369 | 17.3179 |
| QPET-HU2-6| 46.8374 | 17.3192 |
| QPET-HU2-7| 46.8369 | 17.3179 |
| QPET-HU2-8| 46.8374 | 17.3192 |
| QPET-HU2-9| 46.8369 | 17.3179 |
Table 5 (continued)

| Ind_ID     | GPS_1   | GPS_2   |
|------------|---------|---------|
| QPET-HU2-10| 46.8374 | 17.3192 |
| QPET-HU3-1 | 46.2007 | 18.4317 |
| QPET-HU3-2 | 46.2001 | 18.4307 |
| QPET-HU3-3 | 46.1994 | 18.4295 |
| QPET-HU3-4 | 46.2007 | 18.4297 |
| QPET-HU3-5 | 46.1994 | 18.4317 |
| QPET-HU3-6 | 46.1995 | 18.4336 |
| QPET-HU3-7 | 46.2016 | 18.4310 |
| QPET-HU3-8 | 46.1997 | 18.4258 |
| QPET-HU3-9 | 46.1980 | 18.4286 |
| QPET-HU3-10| 46.2005 | 18.4272 |
| KO1-1      | 42.5134 | 20.9681 |
| KO1-2      | 42.5134 | 20.9681 |
| KO1-3      | 42.5133 | 20.9679 |
| KO1-4      | 42.5132 | 20.9677 |
| KO1-5      | 42.5132 | 20.9674 |
| KO1-6      | 42.5131 | 20.9671 |
| KO1-7      | 42.5133 | 20.9670 |
| KO1-8      | 42.5134 | 20.9669 |
| KO1-9      | 42.5133 | 20.9666 |
| KO1-10     | 42.5135 | 20.9664 |
| RO1-1      | 46.6943 | 24.6921 |
| RO1-2      | 46.6942 | 24.6921 |
| RO1-3      | 46.6945 | 24.6922 |
| RO1-4      | 46.6948 | 24.6911 |
| RO1-5      | 46.6946 | 24.6924 |
| RO1-6      | 46.6947 | 24.6922 |
| RO1-7      | 46.6946 | 24.6920 |
| RO1-8      | 46.6950 | 24.6920 |
| RO1-9      | 46.6949 | 24.6916 |
| RO1-10     | 46.6948 | 24.6914 |
| RO2-1      | 45.8394 | 24.0073 |
| RO2-2      | 45.8394 | 24.0076 |
| RO2-3      | 45.8394 | 24.0079 |
| RO2-4      | 45.8393 | 24.0082 |
| RO2-5      | 45.8391 | 24.0086 |
| RO2-6      | 45.8391 | 24.0089 |
| RO2-7      | 45.8392 | 24.0091 |
| RO2-8      | 45.8397 | 24.0092 |
| RO2-9      | 45.8397 | 24.0095 |
| RO2-10     | 45.8398 | 24.0089 |
| RO3-1      | 44.8533 | 24.7932 |
| RO3-2      | 44.8533 | 24.7932 |
| RO3-3      | 44.8533 | 24.7933 |
| RO3-4      | 44.8534 | 24.7939 |
| RO3-5      | 44.8533 | 24.7939 |
| RO3-6      | 44.8536 | 24.7944 |
| RO3-7      | 44.8535 | 24.7946 |
| RO3-8      | 44.8537 | 24.7946 |
| RO3-9      | 44.8538 | 24.7951 |
| RO3-10     | 44.8540 | 24.7952 |
Table 5 (continued)

| Ind_ID | GPS_1   | GPS_2   |
|--------|---------|---------|
| SE1-1  | 45.1720 | 19.8393 |
| SE1-2  | 45.1721 | 19.8395 |
| SE1-3  | 45.1721 | 19.8392 |
| SE1-4  | 45.1720 | 19.8395 |
| SE1-5  | 45.1723 | 19.8393 |
| SE1-6  | 45.1723 | 19.8392 |
| SE1-7  | 45.1723 | 19.8398 |
| SE1-8  | 45.1725 | 19.8393 |
| SE1-9  | 45.1725 | 19.8396 |
| SE1-10 | 45.1726 | 19.8392 |
| SE2-1  | 44.1722 | 20.4881 |
| SE2-2  | 44.1721 | 20.4881 |
| SE2-3  | 44.1719 | 20.4882 |
| SE2-4  | 44.1717 | 20.4880 |
| SE2-5  | 44.1716 | 20.4884 |
| SE2-6  | 44.1714 | 20.4882 |
| SE2-7  | 44.1713 | 20.4877 |
| SE2-8  | 44.1713 | 20.4875 |
| SE2-9  | 44.1717 | 20.4876 |
| SE2-10 | 44.1719 | 20.4875 |
| SE3-1  | 43.6675 | 20.6026 |
| SE3-2  | 43.6673 | 20.6029 |
| SE3-3  | 43.6672 | 20.6030 |
| SE3-4  | 43.6670 | 20.6031 |
| SE3-5  | 43.6672 | 20.6033 |
| SE3-6  | 43.6674 | 20.6034 |
| SE3-7  | 43.6673 | 20.6037 |
| SE3-8  | 43.6671 | 20.6038 |
| SE3-9  | 43.6672 | 20.6039 |
| SE3-10 | 43.6674 | 20.6041 |

Supplementary Information  * The online version contains supplementary material available at https://doi.org/10.1007/s13595-021-01051-6.

Acknowledgements The authors are grateful to Ivan Iliev (University of Forestry, Bulgaria), Lazar Kesic and Velislav Karaklić (University of Novi Sad, Serbia), Zoltán Barina (Hungarian Natural History Museum, Hungary), Ibrahim Muja (Ministry of Agriculture Forestry and Rural Development (MAFRD), Kosovo*), Naser Krasniqi (Kosovo* Forest Agency, Kosovo*), Vlado Kaurin, Alen Gačić, Aleksandar Jorgić, Stojan Božić, Mijović Goran, and Spomenko Stojanović; the PFE “ŠumeRS” a.d. Sokolac (Bosnia and Herzegovina); the Institute of Lowland Forestry and Environment (project number: 451-03-68/2020-14/200197); and the Romanian forest authorities from Sibiu and Arges counties for the assistance they offered with the collection of plant material and for the valuable discussions in the course of which they shared their insights. We also thank Zoltán Bihari (Xenovea Ltd.) for providing sequencing resources and useful comments on data analysis.

Funding This study was supported by the Ministry of Agriculture of Hungary (Kaán Károly project, no: EVGF/549/2018, EGF/178/2019), by the Ministry of Education, Science and Technological Development of the Republic of Serbia (project no.: 451-03-68/2020-14/200197), and by the National Science Fund of Bulgaria (research grant no.: KP-06-H26/4).

Data availability The datasets generated during and analyzed during the current study are available in the ZENODO repository, https://doi.org/10.5281/zenodo.3908963/.

References

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11:697–709. https://doi.org/10.1038/nrg.2010.28

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Gen 17:81. https://doi.org/10.1038/nrg.2015.28
Arnold B, Corbett-Detig RB, Hartl D, Bombíes K (2013) RAD seq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. MolEcol 22:3179–3190. https://doi.org/10.1111/mec.12276

Blanc-Jolivet C, Bakhitina S, Yanbaev R, Yanbaev Y, Mader M, Guichoux E, Degen B (2020) Development of new SNPs loci on Quercus robur and Quercus petraea for genetic studies covering the whole species’ distribution range.Conserv Genet Resour 1-4. https://doi.org/10.1007/s12686-020-01141-z

Bordács S, Popescu F, Slade D, Csáikl UM, Lesur I, Borovics A, Kézdy P, König AO, Gömöry D, Brewer S, Burg K, Petit RJ (2002) Chloroplast DNA variation of white oaks in northern Balkans and in the Carpathian Basin. For Ecol Manage 156:197–209. https://doi.org/10.1016/s0378-1139(01)00643-0

Bruschi P, Vendramini GG, Bussotti F, Grosponi P (2003) Morphological and molecular diversity among Italian populations of Quercus petrea (Fagaceae). Ann Bot 91:707–716. https://doi.org/10.1093/aob/mcg075

Burak MK, Monk JD, Schmitz OJ (2018) Eco-evolutionary dynamics: the predator-prey adaptive play and the ecological theater. Yale J Biol Med 91:481–489

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. MolEcol 22:3124–3140. https://doi.org/10.1111/mec.12354

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes Genom Genet 1:171–182. https://doi.org/10.1534/g3.111.00240

Csáikl UM, Burg K, Fineschi S, König AO, Mátys G, Petit RJ (2002) Chloroplast DNA variation of white oaks in the alpine region. For Ecol Manage 156:131–145. https://doi.org/10.1016/s0378-1127(01)00639-9

Cumer T, Pouchon C, Boyer F, Yannic G, Rioux D, Bonin A, Capblanquet T (2018) Double-digest RAD-sequencing: do wet and dry protocol parameters impact biological results?. bioRxiv 445122. https://doi.org/10.1101/445122

Davey JW, Blaxter ML (2010) RADSeq: next-generation population genomics. Brief Funct Genomics 9:416–423. https://doi.org/10.1093/bfgp/ele031

Dering M, Lewandowski A, Ufnalski K, Kedzierska A (2008) How far to the east was the migration of white oaks from the Iberian refugium? Silva Fenn 42:327. https://doi.org/10.1111/j.1155-7359.2008.tb01563.x

Gömöry D, Yakovlev I, Zhelev P, Jednáňákóvá J, Paule L (2001) Genetic differentiation of oak populations within the Quercus robur/Quercus petraea complex in Central and Eastern Europe. Heredity 86:557–563. https://doi.org/10.1046/j.1365-2640.2001.00874.x

Gömöry D, Zhelev P, Brus R (2020) The Balkans: a genetic hotspot but not a universal colonization source for trees. Plant SystEvol 306:5. https://doi.org/10.1007/s00606-020-01647-x

Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical Fsstatistics. MolEcol Notes 5:184–186. https://doi.org/10.1111/j.1471-2124.2004.00828.x

Hendry AP (2019) A critique for eco-evolutionary dynamics. FunctEcol 33:84–94. https://doi.org/10.1111/1365-2435.13244

Jombart T, Ahmed I (2011) Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Kircher M, Heyn P, Kelso J (2011) Addressing challenges in the production and analysis of illumina sequencing data. BMC Genomics 12:382. https://doi.org/10.1186/1471-2164-12-382

Konar A, Choudhury O, Bullis R, Fiedler L, Kruser JM, Stephens MT, Gailing O, Schlarbaum S, Coggeshall MV, Staton ME, Carlson JE, Emrich S, Romero-Severson J (2017) High-quality genetic mapping with ddRADseq in the non-model tree Quercus rubra. BMC Genomics 18:1–12. https://doi.org/10.1186/s12866-017-3765-8

Kremer A, Hipp AL (2020) Oaks: an evolutionary success story. New Phytol 226:987–1011. https://doi.org/10.1111/nph.16274

Lang T, Abadie P, Léger V, Decourcelle T, Frigerio JM, Barbun C, Bodenes C, Guichoux E, Le Provost G, Robin C, Tani N, Léger P, Lepoittevin C, El Mujtar VA, Hubert F, Tibbits J, Paiva J, Franc A, Raspail F, Mariette S, Reviron MP, Plomion C, Kremer A, Desprez-Loustau ML, Garnier-Gere P (2020) High-quality SNPs from genic regions highlight introgression patterns among European white oaks (Quercus petrea and Q. robur). bioRxiv 388447. https://doi.org/10.1101/388447

Legrand D, Cote J, Frohnofer EA, Holt RD, Ronce O, Schtickzelle N, Clobert J (2017) Eco-evolutionary dynamics in fragmented landscapes. Ecography 40:9–25. https://doi.org/10.1111/ecog.02537

Leroy T, Rougemont Q, Dupouey JL, Bodénes C, Lalanne C, Belser C, Labadie K, Le Provost G, Aury JM, Kremer A, Plomion C (2020) Massive postglacial gene flow between European white oaks uncovered genes underlying species barriers. New Phytol 226:1183–1197. https://doi.org/10.1111/nph.16039

Leroy T, Roux C, Villate L, Bodénes C, Romiguier J, Paiva JA, Dossat C, Aury JM, Plomion C, Kremer A (2017) Extensive recent secondary contacts between four European white oak species. New Phytol 214:865–878. https://doi.org/10.1111/nph.14413

Mastresta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC (2015) Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. MolEcolResour 15:28–41. https://doi.org/10.1111/1755-0998.12291

Miliad M, Schaich H, Bürgi M, Konold W (2011) Climate change and nature conservation in Central European forests: a review of consequences, concepts and challenges. For Ecol Manage 261:829–843. https://doi.org/10.1016/j.foreco.2010.10.038

Mölder A, Meyer P, Nagel RV (2019) Integrative management to sustain biodiversity and ecological continuity in Central European temperate oak (Quercus robur, Q. petraea) forests: an overview. For Ecol Manage 437:324–339. https://doi.org/10.1016/j.foreco.2019.01.006

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

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York
Paris JR, Stevens JR, Catchen JM (2017) Lost in parameter space: a road map for stacks. Methods Ecol Evol 8:1360–1373. https://doi.org/10.1111/2041-210x.12775

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135. https://doi.org/10.1371/journal.pone.0037135

Petit RJ, Brewster S, Bordács S, Burg K, Cheddadi R, Coart E, Cottrell J, Csaikl UM, van Dam BC, Deans JD, Espinel S, Fineschi S, Finkeldey R, Glaz I, Goicoechea PG, Jensen JS, König AO, Lowe AJ, Madsen SF, Mátéyás G, Munro RC, Popescu F, Slade D, Tabbener H, de Vries SGM, Ziegenhagen B, de Beaulieu JL, Kremer A (2002a) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. For Ecol Manage 156:49–74. https://doi.org/10.1016/S0378-1127(01)00634-x

Petit RJ, Csaikl UM, Bordács S, Burg K, Coart E, Cottrell J, van Dam BC, Deans JD, Dumolin-Lapegue S, Fineschi S, Finkeldey R, Gillies A, Glaz I, Goicoechea PG, Jensen JS, König AO, Lowe AJ, Madsen SF, Mátéyás G, Munro RC, Olalde M, Pemonge MH, Popescu F, Slade D, Tabbener H, Taurichini D, de Vries SGM, Ziegenhagen B, Kremer A (2002b) Chloroplast DNA variation in European white oaks: phylogeography and patterns of diversity based on data from over 2600 populations. For Ecol Manage 156:5–26. https://doi.org/10.1016/S0378-1127(01)00645-4

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Julian Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. https://doi.org/10.1086/519795

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rochette NC, Rivera-Colón AG, Catchen JM (2019) Stacks 2: analytical methods for paired-end sequencing improve RADseq-based population genomics. Mol Ecol 28:4737–4754. https://doi.org/10.1111/mec.15253

Schwartz MK, McKelvey KS, Cushman SA, Luikart G (2010) Landscape genomics: a brief perspective. In: Cushman SA, Huettmann F. Spatial complexity: informatics, and wildlife conservation. Springer Japan, pp 165–174. https://doi.org/10.1007/978-4-431-87771-4_9

Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Mol Ecol Resour 11:1–8. https://doi.org/10.1111/j.1755-0998.2010.02979.x

Sims D, Sudbery I, Iliot NE, Heger A, Ponting CP (2014) Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Gen 15:121–132. https://doi.org/10.1038/nrg3642

Slade D, Škvorc Ž, Ballian D, Gračan J, Papes D (2008) The chloroplast DNA polymorphisms of White Oaks of section Quercus in the Central Balkans. Silvae Genet 57:227–234. https://doi.org/10.1515/sen-2008-0035

Tóth EG, Köbölkuti ZA, Cseke K, Kámpel JD, Takács R, Tomov VT, Ábrán P, Stojnic S, Vastag E, Mataruga M, Daničić V, Tahirukaj E, Zhelev P, Orlovic S, Benke A, Borovics A (2020) A genomic data set of single-nucleotide polymorphisms (SNPs) generated by ddRAD tag sequencing in Q. petraea (Matt.) Liebl. populations from Central-Eastern Europe and Balkan Peninsula. ZENODO. https://doi.org/10.5281/zenodo.3908963/

Tzedakis PC (2004) The Balkans as prime glacial refugial territory of European temperate trees. In: Griffiths HI, Kryštufek B, Reed JM (Eds.) Balkan biodiversity - pattern and process in the European hotspot; Springer, Dordrecht, pp 49–68. https://doi.org/10.1007/978-1-4020-2854-0

Zhelev P (2017) Studies on the glacial refugia of forest trees on Balkan Peninsula. Contrib Sec Nat Math Biotech Sci 38:129–135. https://doi.org/10.20903/csnmbs.masa.2017.38.2.108

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.