Clonal Differentiation and Identification of Polyploids in *Juglans* hybrids

Winston Beck[1#*], Oliver Caré[2,3#], Matthias Zander[1], Bernhard Mettendorf[4], Ludger Leinemann[2,3], and Christian Ulrichs[1]

[#]W. Beck and O. Caré contributed equally to the work.

[1] Humboldt Universität Berlin, Division Urban Ecophysiology of Plants, Lentzeallee 55-57, 14195 Berlin, Germany
[2] ISOGEN GmbH & Co. KG. Büsgenweg 2, 37077 Göttingen, Germany
[3] Georg-August-University Göttingen, Department of Forest Genetics and Forest Tree Breeding, Büsgenweg 2, 37077 Göttingen, Germany
[4] Forstbetrieb Mettendorf, Renchen, Germany

*Corresponding author: Winston Beck, e-mail: winston.beck@hu-berlin.de

Abstract

Juglans is a genus of ca. 20 species divided into four sections (*Rhysocaryon, Cardiocaryon, Trachycaryon* and *Juglans*) with a broad geographic distribution which is known for producing edible nuts and high-quality wood. Hybrid *Juglans* trees are becoming increasingly popular for forestry purposes in Europe because of their fast growth and adaptability. In this work, DNA was isolated from leaves of eighty-five individuals of *Juglans* species and hybrids and evaluated at eight nSSR and seven EST-SSR markers. The markers were able to differentiate all trees on an individual basis, and further, gave indications of polyploidy in some samples. Consequently, fifty-nine samples were then checked for ploidy by flow cytometry. Genetic structure of the population was evaluated with the most probable number of groups being K=3, which agrees with the four sections of the genus *Juglans*, where *Cardiocaryon* and *Trachycaryon* are genetically very similar and thus comprise one group. These results agree with other research using SSRs in *Juglans*. Flow cytometry revealed thirteen triploid individuals, where twelve are in agreement with the marker study, one was only detected by flow cytometry.

Keywords: *Walnut, PCR, Hybrid, × intermedia, triploid, marker, breeding*

Introduction

Forest production is undergoing changes across the globe as sustainability goals evolve to relieve changing stresses (EU Forest strat. 2030). Hybrid trees can offer faster growth through heterosis effects (Li, Howe, and Wu 1998, Woeste and Michler 2011), improved tolerance to changing environmental conditions (Fabio, Leary, & Smart, 2019; Brennan et al., 2020), and improved response to pest and disease pressure (Leslie & McGranahan, 2014; Utley et al., 2012; Brennan et al., 2020). The genus Juglans has a wide geographic distribution including approximately twenty species, and has been prized for centuries as a producer of valuable wood and edible nuts. The genus Juglans can be divided into four sections; *Juglans, Rhysocaryon, Trachycaryon* and *Cardiocaryon* (Aradhya et al. 2007), which have also been confirmed by more recent sequence analyses (Dong et al. 2017; Zhao et al. 2018). The sequence analysis of the ITS und Ubiquitin Ligase gene differentiate the sections *Juglans, Rhysocaryon* and *Trachycaryon* with *Cardiocaryon*. Sections Trachycaryon and Cardiocaryon are, on the other hand, characterised by fewer differences (Dong et al. 2017). Many interspecific and intersectional hybrids have been documented in Juglans, some of which have demonstrated commercial worth (Leslie and McGranahan, 2014).

Juglans × intermedia, the hybrid of *J. nigra* and *J. regia* first described by Carrière (1870) has shown potential as a timber tree in Europe (Ehring, et al. 2011; Fady 2003; Mettendorf 2008). *Juglans nigra* was introduced to mainland Europe in the 1600s (Booth, 1882; Nicolescu et al., 2020). However, only recently have *Juglans × intermedia* been produced in appropriate and

DOI:10.2478/sg-2023-0007
edited by the Thünen Institute of Forest Genetics

reproducible quantities for use in forestry in Europe, where plantations especially in France and more recently Germany have been successful and are becoming more widely popular. The genus *Juglans* is monoecious, dichogamous, and wind-pollinated, and spontaneous hybrids are possible (Leslie and McGranahan 2009; Pollegioni et al. 2013; Shu et al. 2019). Hand pollination of *Juglans* can be performed, though several factors influence the success of crosses, such as pollen compatibility, pollen load and the associated pistillate flower abscission, and availability of pollen (McGranahan et al. 2019; Pollegioni 2010). Additionally, manually pollinating enough flowers to generate a marketable crop of F1 trees is a challenge.

An open problem in walnut breeding is the verification of breeding stock. The genus *Juglans* offers substantial diversity and has a long relationship with humans, which has resulted in many locally-associated names, the naming of novelties over time, and synonymy. Verification of starting material can be a problem, as homonyms (different genotypes with the same name) and synonyms (same genotypes with different names) cannot be precluded without testing (Bernard et al., 2018; Hoban and Romero-Severson, 2012; Zhao et al. 2017). Further, the correct identification of *Juglans* species and respective hybrids can be difficult, especially in juvenile plants, as a high amount of plasticity in morphological characteristics can be observed. In mature trees, nuts are highly informative to a correct identification, but are not always available. When additional generations of hybrid trees are observed, especially where nuts are not necessarily produced (Grimshaw, 1985), the use of genetic markers to identify hybrids is especially useful and can additionally be used to eliminate off-target genotypes early in breeding programs. Several classes of markers have been investigated in Juglans, including isozymes (Arulsekar, Parfitt, and McGranahan 1985; Fornari et al. 2001), Restriction Fragment Length Polymorphisms (RFLPs) (Fjellstrom, Parfitt, and McGranahan, 1994), randomly amplified polymorphic DNA (RAPD) markers (Emilia et al. 1997; Nicese, Hormaza, and McGranahan 1998), and simple sequence repeats (SSRs). SSRs are short sequences of DNA (1-6 bp) that are repeated variably among genotypes, and can be applicable across a range of species, and offer a high level of resolution between genotypes. SSRs are additionally capable of determining paternity (Ebrahimi et al. 2018; Pollegioni et al. 2009) and have been increasingly used in *Juglans* in recent literature, as seen in Bernard (2018). The goal of this research was to investigate a collection of *Juglans* species and hybrid trees with varying degrees of provenance knowledge to better understand their genetic constitution.

Materials and Methods

Plant material and DNA extraction

Juglans species and hybrids were sought out from parks, botanical gardens, and landholders, and gathered into a registry. Taxonomic designations were used as provided. Leaf samples were taken from 85 individuals, and approx. 50 mg of leaf tissue was prepared for DNA extraction. The sample tissue was pulverized by freezing in liquid nitrogen and grinding for 2 min at 30 Hz in an MM300 ball mill (Retsch, Haan, Germany). Total genomic DNA was extracted using the DNeasyTM 96 Plant Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The 85 samples also include five duplicates, where two separate samples with identical numbers arrived in the laboratory. All samples were analysed to check identity, amplification and genotyping results. These duplications include sample 6, hereafter referred to as 6a and 6b, sample 94, hereafter referred to as 94.1 and 94.2 and sample 8, hereafter referred to as 8a and 8b. The duplication of sample BG22 is referred to as BG22 and BG22-2 and of sample 13 as 13a and 13b.

The species assessed here can be assigned to the following sections: Section *Juglans* with *J. regia*. Section *Trachycaryon* with *J. cinerea*. Section *Cardiocaryon* with *J. ailantifolia*, *J.mandshurica* and *J. cathayensis*. The species *J. major, J. nigra, J. hindsii, J. californica* and *J. microcarpa* belong to the section *Rhysocaryon* (Aradhya et al. 2007). An overview of the samples is presented in table S1.

Genotyping

For the genetic characterisation, 21 nuclear microsatellites (nSSR) and expressed-sequence tag microsatellites (EST-SSR) were selected from the literature (Dangl et al., 2005; Woeste et al., 2002 Bernard et al., 2018; Pollegioni et al., 2009, 2011; Torokeldiev et al., 2019; Wang et al., 2008; Zhang et al., 2013, Feng et al., 2011; Qi et al., 2011) and tested for amplification, correct genotyping and species-specific genetic variation. For the genotyping of the samples, 15 microsatellites were applied after evaluation of the test results. The corresponding primer sequences, references and observed sizes, as well as the number of different alleles of these eight nSSRs (Dangl et al., 2005; Woeste et al., 2002) and seven EST-SSRs (Feng et al., 2011; Qi et al., 2011) can be found in Supplementary table S2. All PCR reactions were performed in a 14 µL total volume, containing 1 µL of 1:10 diluted DNA. The PCR mix contained 1× reaction buffer B (Solis BioDyne, Tartu, Estonia), 2.68 mM MgCl2, 178.57 µM for each dNTP, and one unit of HOT FIREPol® (Solis BioDyne, Tartu, Estonia) Taq polymerase. For the marker combinations WGA 118, ZMZ 11, ZMZ 5, WGA 321 and WGA 376, ZMZ 45, ZY 1 it was possible to amplify them by multiplex PCR. Each forward and reverse primer was added in a concentration given in Table S3. For the other markers, a separate PCR with the specific primer sequences had to be carried out with a concentration of 3.57 µM/µl for each forward and reverse primer in the PCR-mix. The reaction started with an initial incubation for 15 min at 95 °C, 10 touch-down cycles of denaturation at 94 °C, followed by annealing at 60 °C (Δ−1°C) and extension at 72 °C each for 1 min, 25 cycles at 94 °C, 50 °C, and 72 °C with each temperature level kept for 1 min, and final extension at 72 °C for 20 min. All reactions were run in a Biometra TProfessional Basic thermocycler (Analytic Jena AG, Jena, Germany). SSR fragments were separated on an ABITM 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) with the size standard GeneScanTM 500 ROXTM as a reference (Applied Biosystems, Foster City, CA, USA). The marker combinations WGA276,

WGA4 and ZMZ16 as well as WGA9, WGA79, WGA72 and CV198434 could be measured simultaneously. The PCR products were mixed and diluted according to table S4. Marker ZMZ7 had to be amplified and measured individually. Peak calling was done in the GeneMapperTM v 4.1 software (Applied Biosystems, Foster City, CA, USA). All called peaks at all markers were checked and, if needed, rescored manually to ensure consistent scoring.

Combined analysis of mixed di- and triploid individuals

During the scoring of peak data, it became evident that polyploid allele patterns occurred beside the expected diploid pattern. The combination of different ploidy levels within a genetic data set poses challenges for the analysis. In most cases, statistical methods for genetic data sets are designed for purely diploid (2n), haploid (n) or, more rarely, polyploid (xn) individuals. To be able to apply a wide range of methods, an alternative to the usual coding of SSR genotypes was chosen. Standard convention for SSR genotypes is to define alleles by base pair size (fragment length) with the SSR-name as trait variable. The alternative is to code each allele as a single trait similar to a SNP (single nucleotide polymorphisms). In the present dataset the sum of all different alleles over all SSR-loci is 231. Thus, the data for analysis contains 231 traits coded as absent, heterozygous or homozygous. This approach is normally used to apply methods designed for SNP analysis to SSR data. For example, this procedure is suggested for SSR association analysis with the program TASSEL (Bradbury et al. 2007).

Flow Cytometry

Samples were prepared according to modified kit instructions from CyStain® PI Absolute P (Sysmex Europe GmbH, Norderstedt, Germany). Approximately 1.5 cm^2 of sample leaf tissue was used for analysis, along with ca. 1 cm^2 leaf tissue of *Pisum sativum* 'Ctirad' as an internal standard. Leaf material of sample and internal standard was placed in a petri dish with 3-4 drops of extraction buffer and chopped for 30 s with a razor blade. The remainder of the 0.5 ml of extraction buffer was added and incubated for 30 s before filtering through a 50 µm nylon mesh sieve into a 1.5 ml Eppendorf tube. 1 ml of staining buffer was added, and the samples were allowed to incubate for one hour at room temperature. All samples were prepared in duplicate. Flow Cytometry was carried out using a BD FACSAria II (BD Biosciences, San Jose, CA) set up to detect propidium iodide. Detection was gated to exclude cellular debris and run until 10,000 events were registered. Flow Cytometry output was analyzed using FlowJo Software ver. 10.7.2 (Becton, Dickinson and Company, Ashland, OR).

Statistical analysis

Using the genetic information of the 15 SSRs studied, a Kruskal's Non-metric Multidimensional Scaling was performed based on the genetic differences between individuals. The distances are based on the number of differences between two individuals, as the sum of all different alleles across the individual traits. The calculations were performed in R version 4.1.0 (R core team 2021) with the package "MASS" 7.3-54 (Venables and Ripley 2002).

The analysis of population structure was carried out using STRUCTURE 2.3.4 (Pritchard et al. 2000). STRUCTURE uses only the information of the genetic markers under the assumption that these are present in Hardy-Weinberg proportions and are not linked. Admixture proportions were estimated for K=2 to K=10 clusters with 20 iterations each. Each run consists of 25,000 burn-in steps and 100,000 steps to determine the solution. No initial population membership was used. The results of the 20 iterations per assumed cluster number were summarised using CLUMPAK (Kopelman et al. 2015) and the parameters for selecting K were calculated in STRUCTURE-HARVESTER (Earl and von Holdt 2012). K=3 and K=6 were chosen to assess the population structure present in the data. Where K=3 is the most statistically probable value for K according to the method of Evanno et al. (2005) and K=6 has been chosen as the next larger classification with some statistical support.

Results

Polyploid individuals

In the Genotyping, we found allelic patterns at all markers except ZMZ5 and ZMZ45 that had more than the expected maximum of two alleles in at least one sample. This non-diploid (>2n) allelic scheme was detected especially in hybrids, particularly in crosses designated as *Juglans* (*nigra* × *regia*) × *J. mandshurica*. Samples that show a non-diploid pattern for at least two markers, excluding WGA376 in this assessment, were considered as polyploid. The marker WGA376 was excluded from the assessment of the ploidy, as its amplification products covers a wide size range with up to four alleles in single species and five alleles in hybrid individuals. For this marker it was assumed that two gene regions are amplified, possibly by a repetition of primer binding sites, leading to this pattern. Due to the formatting of the genetic data with each allele as a trait (described above) for the analysis of genetic differentiation/structure, it was nevertheless kept as a highly polymorphic locus.

Population structure and genetic distance

The analysis of the genetic structure (Fig. 1) largely agrees with the section classification. In the statistically most probable division into three groups, the individuals can usually be clearly assigned to one or two clusters that correspond to the section or sections of the corresponding species/species proportions. The section *Rhysocaryon* is represented by the cluster (light red) containing all *J. nigra* samples and the samples of *J. californica*, *J. microcarpa* and *J. major*. Both sections *Trachycaryon* and *Cardiocaryon* are represented in the cluster (light blue) containing *J. cinerea, J. ailantifolia*, *J. mandshurica* and *J. cathayensis*. The section *Juglans* is mostly represented by the cluster (dark blue) containing *J. regia* samples and backcrosses of *J.* (*nigra* × *regia*) × *regia.* Nevertheless three out of four *J. regia* samples

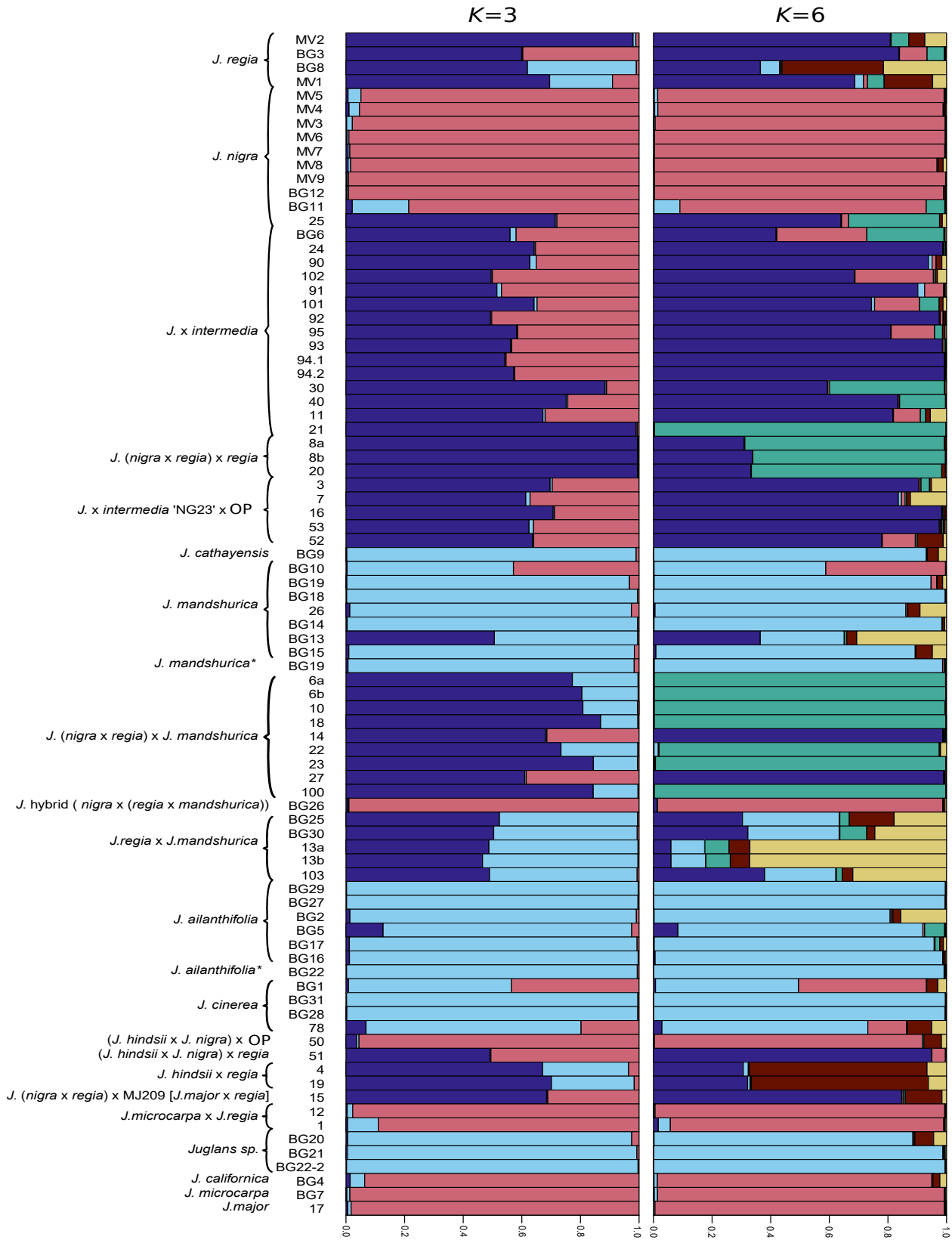


Figure 1

Mean cluster assignment result based on the STRUCTURE analysis and displayed as the average over the 20 runs obtained from CLUMPAK for K=3 and K=6 genetic groups. The admixture proportion to the color-coded clusters can be taken from the respective bar height.

(BG3, BG8, MV1) also had admixture proportions of 30 % to 40 % of the other two clusters.

We chose K=6 groups to assess higher cluster numbers, as this grouping had a slight increase in Delta K compared to the previous and next cluster numbers. Assigning the individuals to six assumed clusters did not reflect the section differentiation as clearly as *K*=3. The additional light green cluster substitutes the other clusters, at least partially, in the samples 21, 30 (*J.* × *intermedia),* 8a, 8b and 20 (*J.* (*nigra* × *regia*) × *regia)* as well as in 6a, 6b, 10, 18, 22, 23 and 100 (*J.* (*nigra* × *regia*) × *mandshurica).* Only two samples (4, 19) are assigned to the dark red cluster to a greater proportion, representing the samples of *J. hindsii* × *regia*, which in K=3 are comprised of approximately 30 % Section *Cardiocaryon*/*Trachycaryon*. Admixture proportions to the light-yellow cluster are mainly present in samples of *J. regia* × *mandshurica* and one sample of *J. mandshurica* (BG13).

The genetic distances (Fig. 2) show a similar grouping. The individuals of *J. regia* and *J.* × *intermedia* are genetically closer as compared to other samples. *J. nigra* individuals with the hybrids between *J. microcarpa* × *J. regia* and the individual specimens of the species *J. californica*, *J. microcarpa* and *J. major* group together. Whereas the species *J. mandshurica*, *J. cinerea*, *J. ailanthifolia* and *J. cathayensis*, without their hybrids, also form a group. A separate group is formed by the polyploid individuals characterised by crosses with *J. mandshurica*.

Flow Cytometry

Genome size estimates for 59 samples were found to range 1.01-1.93 pg (Table 1). Genome sizes of 1.16 to 1.4 pg have been previously published for *Juglans regia* (Khorami et al., 2018) and several species of *Juglans* have been estimated to have genome sizes of around 1.2 pg (Stevens et al., 2018). Forty-five of the selected clones have relative genome sizes from 1.01 to 1.36 pg, corresponding reasonably well to a diploid state. Thirteen clones were found to have larger relative genome sizes from 1.64-1.90 pg, which correlates with a triploid state. Clone 106 did not yield a result in the flow cytometry in both repetitions, as the sample was visibly oxidized and degraded at the time of testing. Ten of these polyploids correlate with the individuals identified as polyploid in the marker study, where clone 14 is the only outlier. Clones 96 and 107 were not analyzed for DNA markers.

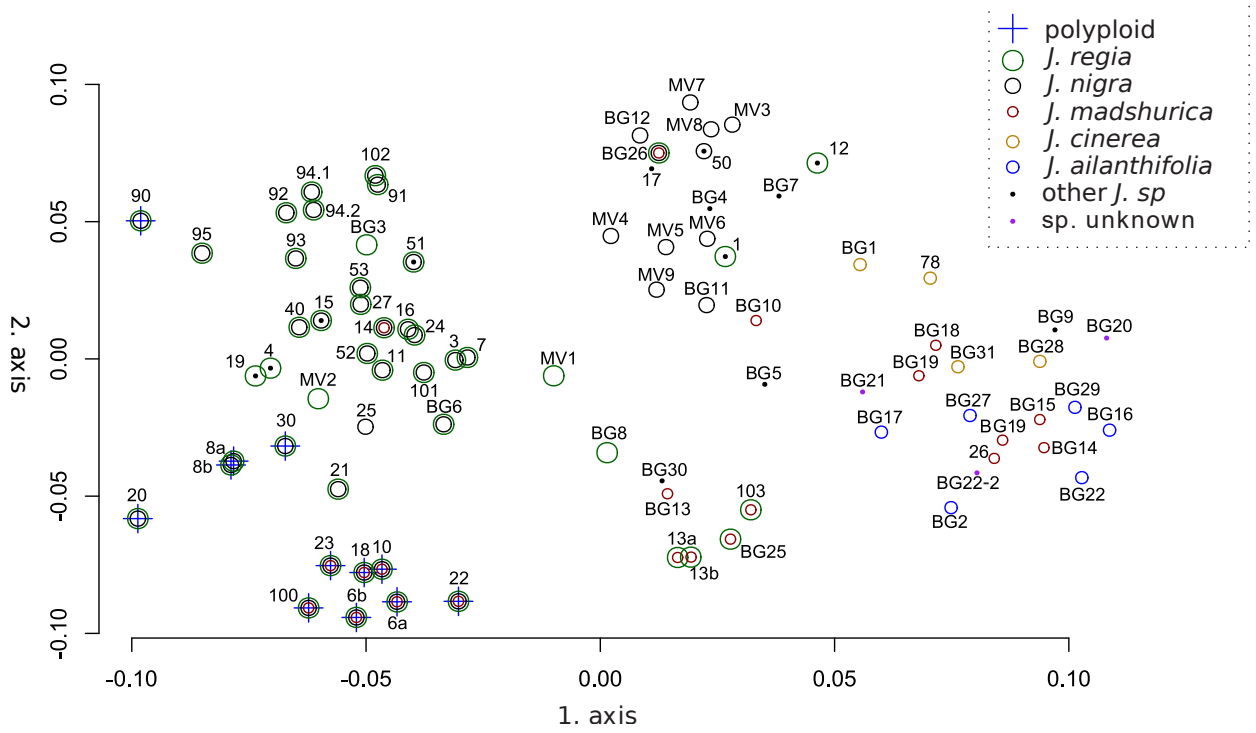


Figure 2

Principal Coordinate Analysis (PCoA) based on genetic distance at eight nSSRs and seven EST-SSRs as pseudoSNPs. Individuals are shown in two dimensions along the two principal coordinates. Species membership/description. This can be taken from the marking of the points. The individual designation is added to the points. As can be seen from the legend, different colours of the circles/dots and their size correspond to the species *J. regia*, *J. nigra*, *J. mandshurica*, *J. cinerea*, *J. ailanthifolia*. A black/pink dot stands for all other species or species that are not specified. The known crossing partners of the individuals are shown in each case. The cross additionally marks samples with more than two alleles on at least two markers.

Table 1: Genome sizes of various Juglans species and hybrids as determined by flow cytometry.

Genotype Description	Clone	2C Genome Size (pg)[z]	Ploidy Level
Diploid			
Juglans (hindsii × nigra) × open-pollinated	50	1.36 ± 0.00	2
Juglans (hindsii × nigra) × regia	51	1.26 ± 0.02	2
Juglans (nigra × major) MJ209 × regia	5	1.26 ± 0.02	2
Juglans (nigra × regia) × J. mandshurica	27	1.24 ±0.07	2
Juglans australis	109	1.29 ± 0.02	2
Juglans australis	105	1.31 ± 0.01	2
Juglans cathayensis × J. regia	106		
Juglans cinerea	108	1.26 ± 0.04	2
Juglans cinerea × mandshurica	61	1.12 ± 0.10	2
Juglans hindsii × J. regia	19	1.27 ± 0.10	2
Juglans hindsii × J. regia	4	1.21 ± 0.05	2
Juglans × intermedia	40	1.26 ± 0.01	2
Juglans major	17	1.27 ±0.02	2
Juglans mandshurica	26	1.14 ± 0.13	2
Juglans microcarpa × J. nigra	12	1.26 ± 0.00	2
Juglans microcarpa × J. nigra	1	1.31 ± 0.03	2
Juglans nigra	MV6	1.24 ± 0.02	2
Juglans nigra	25	1.26 ± 0.00	2
Juglans nigra	MV8	1.27 ± 0.01	2
Juglans nigra	MV9	1.28 ± 0.05	2
Juglans nigra	MV3	1.29 ±0.02	2
Juglans nigra	MV7	1.30 ± 0.02	2
Juglans nigra	MV5	1.30 ± 0.04	2
Juglans nigra	MV4	1.34 ± 0.04	2
Juglans regia	MV2	1.21 ± 0.04	2
Juglans regia × J. mandshurica	103	1.13 ± 0.03	2
Juglans regia × J. mandshurica	13	1.17 ± 0.03	2
Juglans regia × J. mandshurica	6-1	1.19 ± 0.04	2
Juglans regia × J. mandshurica	60	1.11 ± 0.35	2
Juglans sigillata	104	1.19 ± 0.02	2
Juglans × intermedia	92	1.19 ± 0.12	2
Juglans × intermedia	102	1.22 ± 0.01	2
Juglans × intermedia	11	1.22 ± 0.05	2
Juglans × intermedia NG23	24	1.24 ± 0.01	2
Juglans × intermedia	91	1.26 ± 0.02	2
Juglans × intermedia	21	1.26 ± 0.01	2
Juglans × intermedia	93	1.28 ± 0.04	2
Juglans × intermedia	95	1.29 ± 0.02	2
Juglans × intermedia	94	1.30 ± 0.03	2
Juglans × intermedia	101	1.30 ± 0.02	2
Juglans × intermedia NG23 × OP	3	1.21 ± 0.10	2
Juglans × intermedia NG23 × OP	2	1.33 ± 0.01	2
Juglans × intermedia NG23 × OP	7	1.19 ± 0.04	2
Juglans × intermedia NG23 × OP	16	1.28 ± 0.01	2
Juglans × intermedia NG38 × OP	53	1.22 ± 0.02	2
Juglans × intermedia NG38 × OP	52	1.31 ± 0.05	2
Triploid			
Juglans (nigra × regia) × J. mandshurica	23	1.64 ± 0.04	3
Juglans (nigra × regia) × J. mandshurica	10	1.73 ± 0.03	3
Juglans (nigra × regia) × J. mandshurica	22	1.74 ± 0.02	3
Juglans (nigra × regia) × J. mandshurica	100	1.80 ± 0.01	3
Juglans (nigra × regia) × J. mandshurica	14	1.84 ± 0.01	3
Juglans (nigra × regia) × J. mandshurica	18	1.85 ± 0.01	3
Juglans (nigra × regia) × J. mandshurica	6	1.86 ± 0.01	3
Juglans (nigra × regia) × J. regia	8	1.83 ± 0.05	3
Juglans (nigra × regia) × J. regia	20	1.87 ± 0.02	3
Juglans australis × regia ‚X-RODEO'	107	1.84 ± 0.00	3
Juglans regia '120' × nigra	30	1.83 ± 0.01	3
Juglans × intermedia × OP	90	1.75 ± 0.04	3
Juglans × intermedia × OP	96	1.85 ± 0.08	3
Internal Standard			
Pisum sativum 'Ctirad'	IS	9.09	2

zvalues are means ± SEM

Discussion

The selected combination of nSSRs and EST-SSRs is characterised by cross-species usability and a high level of detail in mapping individual genetic differences. This makes it possible to characterise and distinguish individuals belonging to different species and species hybrids. Furthermore, the set is considered suitable for clone identification and differentiation of closely related individuals. The duplicate samples of individuals 6, 8, 13 and 94 show no or only minor genetic differences. Small deviations between samples of identical individuals can be explained experimentally. Single samples partly show unequal amplification strength and quality. This results in different fragment determination due to different peak sizes and peak recognition. Consequently, the absence or occurrence of an allele at only one gene locus is not to be regarded as an exclusion criterion for the determination of identity, especially if amplification strength deviates. The sample BG22 also had a duplication, but genetic differences occurred at 10 out of 15 markers. Therefore, it is highly unlikely that these are samples of the identical individual.

Some samples, especially hybrids, showed a non-diploid (>2n) allelic scheme. A possible polyploid genotype was therefore assumed when more than two alleles occur at more than two gene loci. Due to the complex amplification pattern of WGA376, this marker was not used for this assessment. The deviations from a diploid allelic scheme for hybrids are also known from the literature (Mu and Xi 1988; Pollegioni et al. 2009). In principle, a diploid chromosome set (2n) can be assumed for the pure species of the genus *Juglans* (Mu et al. 1990; Stevens et al. 2018; Luo and Chen 2020). It is reported that hybrids are also more likely to be 2n (Zhu et al. 2019). Nevertheless, no information on the ploidy level of backcrosses is available. Deviations from a diploid allelic scheme are possible by duplication of gene regions during hybridisation and the retention of chromosome number (2n), polyploidisation or amplification of more than one gene region in different species. Since this cannot be distinguished on the basis of the SSR data alone, we performed flow cytometry to assess the average genome size of certain samples. These results show that 13 of 59 samples tested were polyploid, namely, samples 6, 8, 10, 14, 18, 20, 22, 23, 30, 90, 96, 100, and 107, where the averaged genome size for diploid (2n) and triploid (3n) plants were found to range 1.13-1.36 pg and 1.64-1.87 pg, respectively. (n) was calculated as 0.6225 pg, which is in accordance with Stevens et al. (2018) and Khorami (2018) after conversion with the factor supplied by Doležal and Bartoš (2005). Yajima et al. (1997) used colchicine to generate a tetraploid *J. regia* 'Mituru', which exhibited larger pollen grains, and additionally generated tetraploid, triploid, and diploid offspring when crossed with a known diploid parent. All triploid clones except 30, 90, 96, and 107 arose from one mother tree, Clone 21, formerly located at the Forest Arboretum Liliental near Freiburg im Breisgau, Germany. Clone 21 bears characteristics of *J.* × *intermedia* such as leaves with toothed rather than entire margins, pubescent buds, diamond-patterned smooth gray bark, and unique nut shape. According to the STRUCTURE analysis where K=3, the genome can be identified as characteristic of *J. regia*, which does not match the physical attributes, and where K=6 as polyploid, which is not supported by flow cytometry. Additionally, in the CLUSTER analysis, clone 21 falls between the cluster of *J.* × *intermedia* trees and its progeny. The progeny of clone 21, numbers 6a, 6b, 10, 14, 18, 22, 23, 27, and 100, to have hybridized with a member of section *Cardiocaryon*/*Trachycaryon*, as was suspected by Forstbetrieb Mettendorf, where the samples originated. The suspicion of hybridity with a member of section *Cardiocaryon* arose from two aspects: likelihood of paternity; the suspected pollinator(s) (a cluster of five *J. mandshurica* individuals) are geographically close to where the original mother tree stood, and secondly, morphology of the progeny; notably fast-growing among their cohort and display unusual characteristics for *J.* x *intermedia* – the leaves are exceptionally large, rather pubescent (also on stems), with an increased number of leaflets. It is also interesting that seedlings from the same seed lot were apparently also hybridized with *J. nigra* – clone 14, with a triploid genome size, despite not being detected in the STRUCTURE analysis, and clone 27, with a diploid genome size. Clones 90 and 96 (absent from the marker study) also derive from a single diploid mother tree (Clone 91, *J.* x *intermedia* in Meinier, Switzerland), though for both clone 21 and 91, it is unknown what percentage of offspring are polyploid. Initial observations suggest that neither of these polyploid-producing mother trees produce pollen, and any progeny would be obligate hybrids, assuming they come from sexual means. Apomixis has been documented in *Juglans*, occurring at varying rates in varying years (Loiko, 1990; Schanderl, 1964). These triploids could be the result of the B_{III} hybridization pathway, where an unreduced egg sack is formed, and pollination occurs, resulting in triploid embryos and pentaploid endosperm (Hojsgaard and Hörandl 2019). This method of reproduction would explain the variation in the following generation along with the ploidy.

The phylogenetic distances between sections (Dong et al. 2017) is also reflected in the genetic distances at the assessed SSR markers, as well as in the population structure for three clusters. Both groupings suggest the presence of three genetic groups representing the sections *Juglans*, *Rhysocaryon* and *Trachycaryon* combined with *Cardiocaryon*. A separation of *Trachycaryon* and *Cardiocaryon* is not possible, probably due to the high genetic similarity (Aradhya *et al.* 2007, Bernard *et al.* 2018). Even a higher number of groups does not separate the two sections. Thus, an estimation of the individual affiliation to a genus section, or sections in case of hybrids, is possible. Since the description of variation within and between species is based on a relatively small sample per species, general statements on variation differences between species and sections are only possible to a limited extent. This extent includes correction of samples submitted under incorrect names, such as BG26 – a strong seedling which was overzealously identified as hybrid, BG13 – a clear hybrid which was later confirmed morphologically, and samples 14 and 27, both of which appear to be rather *J.* x *intermedia* types not including markers from section *Cardiocaryon*. Additionally, the detection or confirmation of intersectional hybrids is possible; clone 71 was found to be a

J. × *intermedia* genotype, which was not obvious at the time of harvest, and BG8, a seedling of *J. regia* was documented as only *J. regia*, though it contains a portion of genome from section *Cadiocaryon*. The further subdivision at K=6 groups rather reflects individual relationships and differences in allele distributions. This comes essentially evident for the group of putative polypoid individuals, which outgroup from their parental species group based on the genetic distance assessment and form a separate (light green) cluster if six genetic groups are considered in the Bayesian clustering.

These results reiterate the increasing importance of genetic characterisation for various applications in the genus *Juglans*- misidentification of species and their hybrids, synonymy and homonymy are common problems that are easily propagated. Without an organized effort to benchmark genotypes and certify stock materials, these problems will likely continue. The detection of multiple polyploid F2 individuals is novel, and highlights a need for more research to understand the underlying mechanisms behind these trees' existence, and may lead to novel genotypes for use in forestry.

Acknowledgements

The authors would also like to thank Dr. Gabi Schreiber and Min Kyung Kim for their assistance with the flow cytometry.

Funding

This project was funded by the Agency of Renewable Resources (FNR), an initiative of the German Federal Ministry of Food and Agriculture (BMEL) under the project number 22025416.

References

Aradhya MK, Potter D, Gao F, Simon CJ (2007) Molecular phylogeny of Juglans (Juglandaceae): A biogeographic perspective. Tree Genetics and Genomes, 3(4): 363–378. https://doi.org/10.1007/s11295-006-0078-5.

Arulsekar S, Parfitt DE, Mcgranahan GH (1985) Isozyme gene markers in Juglans species. The Journal of Heredity 76:103-106. https://academic.oup.com/jhered/article/76/2/103/875376

Bernard A, Barreneche T, Lheureux F, Dirlewanger E (2018) Analysis of genetic diversity and structure in a worldwide walnut (*Juglans regia* L.) germplasm using SSR markers. PLoS ONE, 13(11), 1–19. https://doi.org/10.1371/journal.pone.0208021.

Brennan AN, McKenna JR, Hoban SM, Jacobs DF (2020) Hybrid Breeding for Restoration of Threatened Forest Trees: Evidence for Incorporating Disease Tolerance in *Juglans cinerea*. Frontiers in Plant Science, 11:1-9. https://doi.org/10.3389/fpls.2020.580693.

Booth J (1882) Die Naturalisation ausländischer Waldbäume in Deutschland. Berlin, Germany: Springer-Verlag, 168 p, ISBN 978-3-642-51844-7

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. https://doi.org/10.1093/bioinformatics/btm308.

Carrière EA (1870) Juglans intermedia quadrangulata. Revue Horticole, 42, 493–495.

Dangl GS, Woeste K, Aradhya MK, Koehmstedt A, Simon C, Potter D, Leslie CA, McGranahan G (2005) Characterization of 14 Microsatellite Markers for Genetic Analysis and Cultivar Identification of Walnut. Journal of the American Society for Horticultural Science, 130(3): 348–354. https://doi.org/10.21273/jashs.130.3.348

Doležel J, Bartoš J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. Annals of Botany, 95(1): 99–110. https://doi.org/10.1093/aob/mci005.

Dong W, Xu C, Li W, Xie X, Lu Y, Liu Y, Jin X, Suo Z (2017) Phylogenetic resolution in Juglans based on complete chloroplast genomes and nuclear DNA sequences. Front Plant Sci 8:. https://doi.org/10.3389/fpls.2017.01148.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. https://doi.org/10.1007/s12686-011-9548-7.

Ebrahimi A, Mathur S, Lawson SS, LaBonte NR, Lorch A, Coggeshall MV, Woeste KE (2019) Microsatellite Borders and Micro-sequence Conservation in Juglans. Scientific Reports, 9(1): 1–10. https://doi.org/10.1038/s41598-019-39793-z.

Ebrahimi A, Lawson SS, Frank GS, Coggeshall MV, Woeste KE, McKenna JR (2018) Pollen flow and paternity in an isolated and non-isolated black walnut (*Juglans nigra* L.) timber seed orchard. PLoS ONE, 13(12):1-17. https://doi.org/10.1371/journal.pone.0207861.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365- 294X.2005.02553.x.

Ehring A, Arnold E, Frank R, Hein S (2011) Anbauversuch mit Französischen Hybridnussbäumen. FVA-einblick (1):18-22.

Emilia MM, Spada M (1997) Differentiation of walnut hybrids (*Juglans nigra* × *Juglans regia*) through RAPD Markers. Acta Horticulturae, 442:43–52. https://doi.org/10.17660/actahortic.1997.442.4

European Commission (2021) New EU Forest Strategy for 2030. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2021) 572 Final. European Commission: Brussels, Belgium.

Fabio ES, Leary CJ, Smart LB (2019) Tolerance of novel inter-specific shrub willow hybrids to water stress. Trees 33:1015–1026. https://doi.org/10.1007/s00468-019-01835-4.

Fady B, Ducci F, Aleta N, Becquey J, Diaz Vazquez R, Fernandez Lopez F, Jay-Allemand C, Lefèvre F, Ninot A, Panetsos K, Paris P, Pisanelli A, Rumpf H (2003) Walnut demonstrates strong genetic variability for adaptive and wood quality traits in a network of juvenile field tests across Europe. New Forests 25:211-225. https://doi.org/10.1023/A:1022939609548

Feng Z, Zhang Z, Zhang S,d Luo S (2011) Development of walnut EST-SSR markers and primer design. Journal of Anhui Agri. Sci., 39(36): 22217–22221.

Fjellstrom RG, Parfitt DE, Mcgranahan GH (1994) Genetic Relationships and Characterization of Persian Walnut (*Juglans regia* L.) Cultivars Using Restriction Fragment Length Polymorphisms (RFLPs). J. AMER. SOC. HORT. SCI 119(4). https://doi.org/10.21273/jashs.119.4.833

Fornari B, Malvolti ME, Taurchini D, Fineschi S, Beritognolo I, Maccaglia E, Cannata F (2001) Isozyme and Organellar DNA Analysis of Genetic Diversity in Natural/Naturalised European and Asiatic Walnut (*Juglans regia* L.) Populations. Acta Horticulturae, 544:167-192. https://doi.org/10.17660/actahortic.2001.544.23

Grimshaw JM (2003) Notes on the temperate species of Juglans. International Dendrology Society Yearbook 2003:107–130.

Hoban S, Romero-Severson J (2012) Homonymy, synonymy and hybrid misassignments in butternut (*Juglans cinerea*) and Japanese walnut (*Juglans ailantifolia*) nut cultivars. Genetic Resources and Crop Evolution, 59(7):1397–1405. https://doi.org/10.1007/S10722-011-9767-5/TABLES/3

Khorami SS, Arzani K, Karimzadeh G, Shojaeiyan A, Ligterink W (2018) Genome size: A novel predictor of nut weight and nut size of walnut trees. HortScience, 53(3):275–282. https://doi.org/10.21273/HORTSCI12725-17.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: A program for identifying clustering modes and packaging population

structure inferences across K. Mol Ecol Resour 15:1179–1191. https://doi.org/10.1111/1755-0998.12387.

Leslie C, McGranahan G (2009) Breeding Walnuts (Juglans regia). In: Jain S and Priyadarshan P (eds) Breeding Plantation Tree Crops: Temperate Species. New York, USA: Springer Science + Business Media, pp 249-273. https://doi.org/10.1007/978-0-387-71203-1_8

Leslie CA, McGranahan GH (2014) The California Walnut Improvement Program: Scion breeding and rootstock development. Acta Horticulturae, 1050:81–88. https://doi.org/10.17660/ActaHortic.2014.1050.9

Li B, Howe GT, Wu R (1998) Developmental factors responsible for heterosis in aspen hybrids (Populus tremuloides × P. tremula). Tree Physiology, 18:29–36. https://doi.org/10.1093/treephys/18.1.29

Loiko RE (1990) Apomixis of Walnut. Acta Horticulturae 284: 233-236. https://doi.org/10.17660/actahortic.1990.284.31

Luo X, Chen J (2020) Distinguishing Sichuan Walnut Cultivars and Examining Their Relationships with *Juglans regia* and *J. sigillata* by FISH, Early-Fruiting Gene Analysis, and SSR Analysis. Front Plant Sci 11:. https://doi.org/10.3389/fpls.2020.00027.

McGranahan GH, Voyiatzis DG, Catlin PB, Polito VS (2019) High Pollen Loads Can Cause Pistillate Flower Abscission in Walnut. Journal of the American Society for Horticultural Science, 119(3):505–509. https://doi.org/10.21273/jashs.119.3.505.

Mettendorf B (2008) Anbau von Juglans-Hybriden. AFZ-Der Wald, 63(16):858-861.

Mu YL, Xi RT (1988) Microsporogenesis studying and karyotype analysis of *Juglans regia* L. and *J. hopeiensis*. J Agric Univ Hebei 11:48–55.

Mu YL, Xi RT, Lv ZS (1990) Microsporogenesis observation and karyotype analysis of some species in genus Juglans L. J Wuhan Bot Res 8:301–310.

Nicese FP, Hormaza JI, Mcgranahan GH (1998) Molecular characterization and genetic relatedness among walnut (*Juglans regia* L.) genotypes based on RAPD markers. Euphytica, 101:199–206.

Nickravesh MH, Vahdati K, Amiri R, Woeste K (n.d.) Application of Microsatellite and Inter Simple Sequence Markers for Certifying Trueness to Type of Grafted Ramets and Clustering of Persian Walnut Varieties. https://doi.org/10.21203/rs.3.rs-542091/v1.

Nicolescu VN, Rédel K, Vor T, Bastien JC, Brus R, Bencat T, Dodan M, Cvjetkovic B, Andrasev S, la Porta N, Lavnyy V, Petokova K, Peric S, Bartlett D, Hernea C, Pástor M, Mataruga M, Podráský V, Sfecla V, Stefancik I (2020) A review of black walnut (*Juglans nigra* L.) ecology and management in Europe. Trees, 34:1087–1112. https://doi.org/10.1007/s00468-020-01988-7

Pollegioni P, Woeste K, Major A, Mugnozza GS, Malvolti ME (2009a) Characterization of *Juglans nigra* (L.), *Juglans regia* (L.) and Juglans × intermedia (Carr.) by SSR markers: a case study in Italy. Silvae Genet 58:68–78. https://doi.org/10.1515/sg-2009-0009.

Pollegioni P, Woeste K, Mugnozza GS, Malvolti ME (2009b) Retrospective identification of hybridogenic walnut plants by SSR fingerprinting and parentage analysis. Molecular Breeding, 24(4):321–335. https://doi.org/10.1007/s11032-009-9294-7

Pollegioni P, Woeste K, Olimpieri I (2010) Pollen Biology and Hybridization Process: open problem in walnut. In B. J. Kaiser (Ed.), Pollen: Structure, Types and Effect. New York, USA: Nova Science Publishers, Inc., pp 65-99, ISBN 978-1-61668-669-7

Pollegioni P, Woeste K, Olimpieri I, Marandola D, Cannata F, Malvolti ME (2011) Long-term human impacts on genetic structure of Italian walnut inferred by SSR markers. Tree Genetics and Genomes, 7(4):707–723. https://doi.org/10.1007/s11295-011-0368-4.

Pollegioni P, Olimpieri I, Woeste KE, de Simoni G, Gras M, Malvolti ME (2013) Barriers to interspecific hybridization between *Juglans nigra* L. and *J. regia* L species. Tree Genetics and Genomes, 9(1):291–305. https://doi.org/10.1007/s11295-012-0555-y

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x.

Qi JX, Hao YB, Zhu Y, Wu CL, Wang WX, Leng P (2011) Studies on germplasm of Juglans by EST-SSR markers. Acta Horticulturae Sinica, 38(3):441–448.

Schanderl H (1964) Untersuchungen über die Blütenbiologie und Embryonenbildung von *Juglans regia* L. Biologisches Zentralblatt 83: 71-103.

Shu Z, Zhang X, Yu D, Xue S, Wang H (2019) Natural Hybridization between Persian Walnut and Chinese Walnut Revealed by Simple Sequence Repeat Markers. Journal of the American Society for Horticultural Science, 141(2):146–150. https://doi.org/10.21273/jashs.141.2.146

Stevens KA, Woeste K, Chakraborty S, Crepeau MW, Leslie CA, Martínez-García PJ, Puiu D, Romero- Severson J, Coggeshall M, Dandekar AM, Kluepfel D, Neale DB, Salzberg SL, Langley CH (2018) Genomic Variation Among and Within Six Juglans Species. Genes|Genomes|Genetics 8:2153–2165. https://doi.org/10.1534/g3.118.200030.

Torokeldiev N, Ziehe M, Gailing O, Finkeldey R (2019) Genetic diversity and structure of natural *Juglans regia* L. populations in the southern Kyrgyz Republic revealed by nuclear SSR and EST- SSR markers. Tree Genetics and Genomes, 15(5). https://doi.org/10.1007/s11295-018-1311-8.

Utley C, Nguyen T, Roubtsova T, Coggeshall M, Ford TM, Grauke LJ, Graves AD, Leslie CA, McKenna J, Woeste K, Yaghmour MA, Cranshaw W, Seybold SJ, Bostock RM, Tisserat N (2012) Susceptibility of walnut and hickory species to Geosmithia morbida. Plant Disease, 97(5):601–607. https://doi.org/10.1094/PDIS-07-12-0636-RE.

Venables WN, Ripley BD (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Wang H, Pei D, Gu RS, Wang BQ (2008) Genetic diversity and structure of walnut populations in central and southwestern China revealed by microsatellite markers. Journal of the American Society for Horticultural Science, 133(2):197–203. https://doi.org/10.21273/jashs.133.2.197.

Woeste K, Burns R, Rhodes O, Michler C (2002) Thirty polymorphic nuclear microsatellite loci from black walnut. Journal of Heredity, 93(1):58–60. https://doi.org/10.1093/jhered/93.1.58.

Woeste K, Michler C (2011) Juglans. In: Kole C (eds) Wild Crop Relatives: Genomic and Breeding Resources. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-21250-5_5.

Yajima M, Nakamura H, Takahashi K, Watanabe Y, Saito S, Yokozawa Y (1997) Somatic Chromosome Numbers of Colchicine-Treated Shinano Walnut (*Juglans regia* L.) and Its F1 Seedlings. J. Jap. Soc. Hort. Sci, 65(4), 677–683. https://doi.org/10.2503/jjshs.65.677

Zhang ZY, Han JW, Jin Q, Wang Y, Pang XM, Li YY (2013) Development and characterization of new microsatellites for walnut (*Juglans regia*). Genetics and Molecular Research, 12(4):4723–4734. https://doi.org/10.4238/2013.October.18.10.

Zhao P, Zhou HJ, Potter D, Hu YH, Feng XJ, Dang M, Feng L, Zulfiqar S, Liu WZ, Zhao GF, Woeste K (2018) Population genetics, phylogenomics and hybrid speciation of Juglans in China determined from whole chloroplast genomes, transcriptomes, and genotyping-by-sequencing (GBS). Mol Phylogenet Evol 126:250–265. https://doi.org/10.1016/j.ympev.2018.04.014.

Zhao P, Zhou H, Coggeshall MV, Reid B, Woeste KE (2017) Discrimination and assessment of black walnut (*Juglans nigra* L.) nut cultivars using phenology and microsatellite markers (SSRs). Canadian Journal of Plant Science 98(3):616-627. https://doi.org/10.1139/cjps-2017-0214

Zhu T, Wang L, You FM, Rodriguez JC, Deal KR, Chen L, Li J, Chakraborty S, Balan B, Jiang CZ, Brown PJ, Leslie CA, Aradhya MK, Dandekar AM, McGuire PE, Kluepfel D, Dvorak J, Luo MC (2019) Sequencing a *Juglans regia* × *J. microcarpa* hybrid yields high-quality genome assemblies of parental species. Hortic Res 6:1–16. https://doi.org/10.1038/s41438-019-0139-1.