Identification of three large full-sib families of *Quercus rubra* for genetic mapping in an isolated planting outside the species' range in Germany

Tania Dominguez-Flores[1], Victor Chano[1,2] and Oliver Gailing[1,2,3]*

[1] Department of Forest Genetics and Forest Tree Breeding, Georg-August University of Göttingen, Büsgenweg 2, 37077 Göttingen, Germany.
[2] Center for Integrated Breeding Research (CiBreed), Georg-August University of Goettingen, 37073 Göttingen, Germany
[3] Centre of Biodiversity and Sustainable Land Use, Georg-August University of Göttingen, 37073 Göttingen, Germany

* Corresponding author: Oliver Gailing, E-mail: ogailin@gwdg.de

Abstract

Forest tree breeding faces challenges due to long generation times and high costs. To address these issues, the concept of "breeding without breeding" (BwB) offers the opportunity to accelerate breeding cycles by using naturally generated half-sib and full-sib families. An isolated population of trees may provide an ideal environment to identify full-sib families for genetic mapping without the need for controlled pollination. In this study, paternity analysis based on genotyping of 16 SSR markers in 1232 seedlings identified three large full-sib families of *Quercus rubra* (Qr_N - Qr_W with 576 siblings, Qr_N - Qr_S with 175 siblings and 124 siblings in the family Qr_P - Qr_11). Paternity assignments were effectively corroborated by segregation analysis, resulting in large F_1 full-sib families for further experiments such as Quantitative Trait Locus (QTL) mapping.

Keywords: *Northern red oak; Breeding without Breeding; parentage analysis; microsatellites; QTL*

Introduction

Forest tree breeding is hampered by long generation times, breeding cycles, and associated high costs. But the current changing environment together with the growing demand of wood and timber are increasing the pressure on tree breeders to select individuals presenting desirable traits like higher quality or resistance to pests and diseases (Gauthier et al., 2015). For these reasons, technical enhancement in tree breeding is now essential, and accelerating breeding cycles can be achieved by implementing the "breeding without breeding" (BwB hereafter) concept (El-Kassaby and Lstibůrek, 2009). Given the technical difficulties and costs of performing controlled crosses in forest trees, the BwB method aims to work with naturally generated half-sib and full-sib families after the reconstruction of pedigree structures by paternity assignment, and this material can be later used for quantitative genetic analyses like Quantitative Trait Locus (QTL) or linkage mapping (El-Kassaby and Lstibůrek, 2009). Examples of the successful identification of full-sib families in isolated forest patches can be already found in trees, like in European common ash (Krautwurst et al., 2023), honey locust (Owusu et al., 2016; Gailing et al., 2017) or northern red oak in North America (Konar et al., 2017).

In our previous work, we studied the effect of geographic isolation on gene flow in a small isolated group of *Quercus rubra* L. trees outside their native range in Germany (Dominguez-Flores et al., in press). Paternity analysis based on genotyping of 16 microsatellite markers of 365 seedlings from four seed parents revealed that most of the pollen (84.89 %) came from trees located within the site, and three larger full-sib families were identified (Qr_N - Qr_W with 145 siblings, Qr_N - Qr_S with 64 siblings and Qr_P – Qr_11 with 51 siblings) (Dominguez-Flores et al., in press).

We used paternity assignment on additional progeny (1232 seedlings in total) and segregation analysis in each of the three tentative full-sib families to identify and validate large full-sib families that can later be used for genetic mapping. These 1232 offspring from four *Q. rubra* seed parents (864 from Qr_N, 57 from Qr_S, 81 from Qr_W and 230 from Qr_P) planted in the arboretum of the University of Göttingen, Germany

DOI:10.2478/sg-2024-0008
edited by the Thünen Institute of Forest Genetics

(51°33'26.9"N 9°57'26.1) (Supplementary material *Fig. S1*) were used for the paternity analyses.

After DNA extraction using the commercial DNeasy® 96 Plant Kit (Qiagen, Hilden, Germany), the adult trees and the 1232 seedlings were genotyped using 16 nuclear SSR markers (nSSRs) (Supplementary material *Table S1*), that in earlier studies demonstrated reliable amplification in *Q. rubra* (Dominguez-Flores et al., in press).

Polymerase chain reactions (PCRs) were performed as described in Dominguez-Flores et al., in press). PCR products were separated on an ABI PRISM® 3130 Genetic Analyser and visualised using GeneMapper™ v4.1 software (Applied Biosystems™) and the R package Fragman (R Core Team, 2021; Covarrubias-Pazaran et al., 2016) (Supplementary material *Fig. S2*).

Parentage was assessed using four chloroplast SSRs (Deguilloux et al., 2003; Götz and Gailing, 2022) and 16 nuclear SSRs (Aldrich et al., 2002; Durand et al., 2010; Sullivan et al., 2013) (for more details see Dominguez-Flores et al., in press). In short, the 15 adults found within the arboretum were genotyped and considered as potential parents. Paternity analyses were performed using the 16 nSSRs based on the maximum likelihood approach implemented in CERVUS v3.0.7 (Marshall et al., 1998), and COLONY v2.0.6.8 (Jones and Wang, 2010) was later used for an independent validation of full-sib families. The input parameters were the same as described in Dominguez-Flores et al., in press with the proportion of genotyped loci being 0.9659 for the 16 loci (Supplementary material *File S1*).

Seedlings that appeared in both CERVUS and COLONY as full-siblings and with less than 3 mismatches were considered members of the same full-sib family.

The R package OneMap (Margarido et al., 2007) was used for segregation analyses to check for null alleles and to validate full-sib families. The analysis was carried out on the three largest putative full-sib families, and results are shown in Table 1.

As shown in Figure 1, from the total of 1232 plants analysed, 576 were offspring of neighbouring trees Qr_N and Qr_W, for which Qr_N was the seed parent for 509 seedlings while Qr_W was the seed parent of the remaining 67 seedlings. The second largest family with 175 seedlings was Qr_N - Qr_S, of which 140 and 35 seedlings had Qr_N and Qr_S as seed parent, respectively. The third family consisted of 124 seedlings whose seed parent was Qr_P and the pollen donor of all of them was Qr_11.

Thirty-two of the performed tests showed the expected Mendelian segregation patterns based on χ^2 tests ($p < 0.05$). Moreover, a significant deviation from expected segregation was found in two cases (loci 2P24 and FIR031, in family Qr_N – Qr_W) and no null allele was detected in any of the loci. Four markers, 1PC10, quru-GA-0E09, quru-GA-1F07 and quru-GA-0C11, showed stutter bands, complicating accurate interpretation of the presence of alleles when allelic variations only diverged by two base pairs (Supplementary material *Fig. S3*). Without parental information, discerning the authenticity of these alleles becomes challenging.

Obtaining full-sib families for quantitative genetic analyses in forest tree species such as *Q. rubra* is costly and difficult to perform by controlled crosses, which are hampered by factors like the variation in timing and duration of flowering, the interference of external pollen, or the long time required to reach reproductive maturity and seed maturation. Methods like BwB aim to overcome these difficulties by selecting elite offspring individuals generated by natural mating after reconstruction of pedigree structure (El-Kassaby and Lstibůrek, 2009). Currently, one full-sib family and genetic linkage map are available for *Q. rubra*. This full-sib family was generated from two isolated neighbouring trees (SM1 and SM2) on the campus of the University of Purdue, Indiana (EEUU) and propagated in 2013 after applying paternity analyses on open pollinated progeny (Konar et al., 2017).

Previously we genotyped 365 seedlings, identifying three large full-sib families distributed in two clusters of neighbouring trees (Dominguez-Flores et al., in press). After the analysis of 1232 seeds, we identified 576, 175 and 124 full-sibs for each family (expected numbers based on proportions in the 365 seedlings: 489, 212 and 172, respectively) (Fig. *2*).

Therefore, the isolation of the planting site from other conspecific plantations and the spatial configuration of the trees (clusters of trees with large canopies and ample seed production), provide an ideal environment to implement BwB strategies and segregation analyses to validate parent genotypes and identify full-sibs, which are a valuable resource for high-density linkage maps, QTL mapping, and marker-assisted selection.

Moreover, segregation analyses allow for the detection of possible events of segregation distortion/meiotic drive (Buckler et al., 1999). Only two markers showed significant segregation distortion in one family (2P24 and FIR031) that might indicate gametic or postzygotic selection (Møller, 1997; Kuang et al., 1999). In Bodénès et al. (2016), high-density linkage mapping in two intraspecific and two interspecific full-sib families of *Q. petraea* and *Q. robur* was used to detect patterns of segregation distortion in chromosomal regions indicative of reproductive isolation between species due to different causes, like chromosome loss, chromosome rearrangements, genetic load, or pre/post zygotic selection. Additionally, it has been observed that acorn production in *Quercus* is not uniform, and stress conditions might favour selection of genotypes (seeds) with greater viability in terms of genetic or phenotypic quality, and ultimately resulting in selection on certain genes (Sork and Bramble, 1993; Díaz et al., 2003; Bodénès et al., 2016; Gręda et al., 2022).

In summary, we applied paternity assignment and segregation analysis to identify and validate three large full-sib families, Qr_N - Qr_W (576), Qr_N - Qr_S (175) and Qr_P - Qr_11 (124), from open pollinated progeny. The largest full-sib family can be used to generate high-density linkage maps for QTL mapping, but with the other two smaller families (with less than 500 siblings), the effects of QTL might be greatly overestimated due to the Beavis effect (Beavis, 1994). However, the availability of multiple full-sib families would increase the power to detect QTL due to a larger number of segregating loci, so more seedlings should be collected and genotyped to increase the size of both families for future analyses.

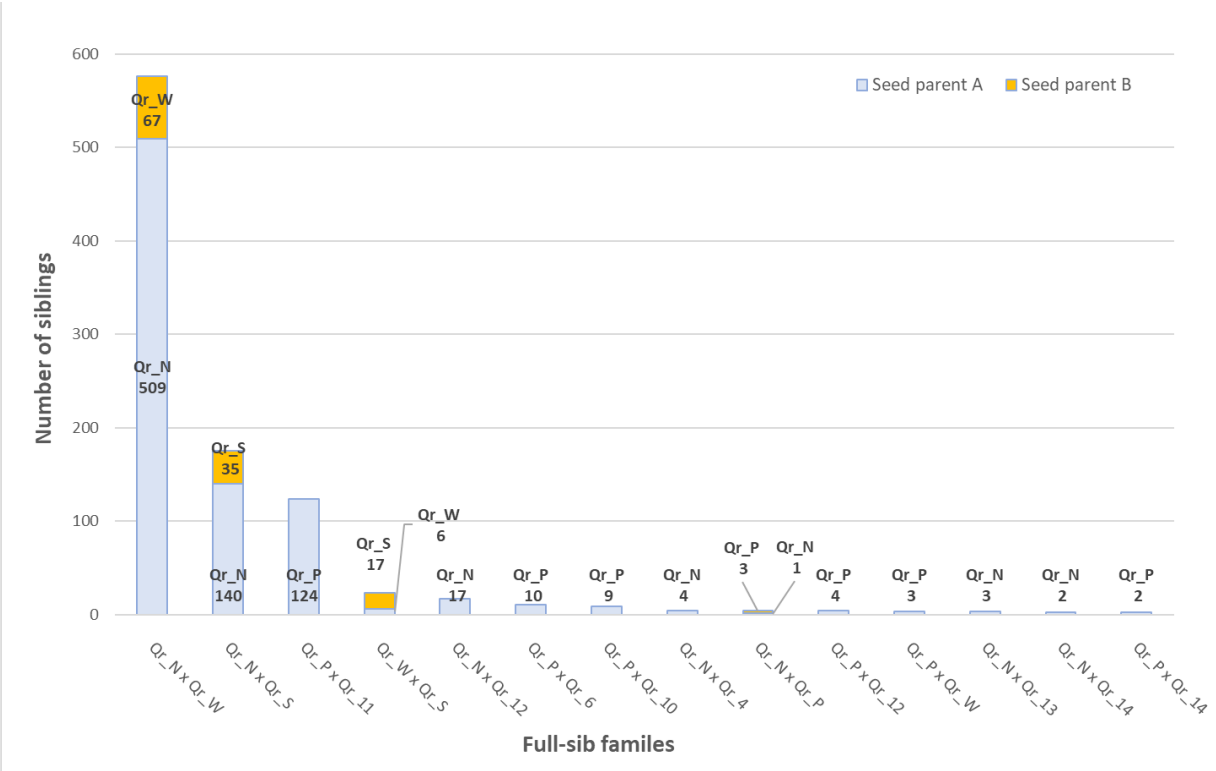

Figure 1
The number of full-sibs in each family and the seed parents of each *Quercus rubra* family are indicated in the graph.

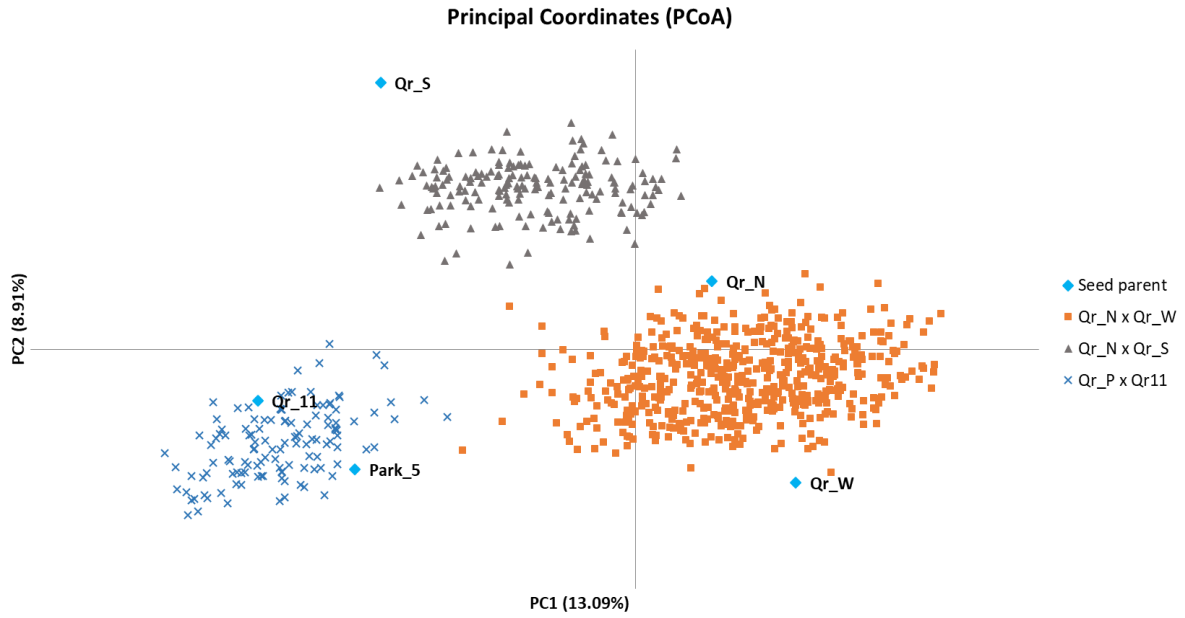

Figure 2
Principal Coordinates Analysis (PCoA) based on 16 nSSRs characterized in the three full-sib families.

Table 1
Inheritance analysis for the nuclear SSRs in the three full-sib families.

Locus	Family	No. siblings	Parent genotype	Progeny genotype	Observed segregation	No. of siblings genotyped	Expected progeny genotype	Expected segregation	Chi-square	p-value
1P10	Qr_N Qr_W	576	250-256 x 250-256	250-250/250-256/256-256	129:284:121	534	250-250/250-256/256-256	1:2:1	2.40	3.0E-01
	Qr_N Qr_S	175	250-256 x 242-244	242-250/242-256/244-250/244-256	33:44:35:52	164	242-250/242-256/244-250/244-256	1:1:1:1	5.61	1.3E-01
	Qr_P Qr_11	124	250-250 x 250-256	250-250/250-256	68:53	121	250-250/250-256	1:1	1.86	1.7E-01
2P24	Qr_N Qr_W	576	154-158 x 150-160	150-154/150-158/154-160/158-160	99:140:129:186	554	150-154/150-158/154-160/158-160	1:1:1:1	28.22	3.3E-06
	Qr_N Qr_S	175	154-158 x 144-146	144-154/144-158/146-154/146-158	42:52:32:42	168	144-154/144-158/146-154/146-158	1:1:1:1	4.76	1.9E-01
	Qr_P Qr_11	124	148-152 x 150-152	148-150/148-152/150-152/152-152	38:28:26:29	121	148-150/148-152/150-152/152-152	1:1:1:1	2.80	4.2E-01
3A05	Qr_P Qr_11		150-154 x 150-154	150-150/150-154/154-154	29:54:36	119	150-150/150-154/154-154	1:2:1	1.84	3.9E-01
3D15	Qr_N Qr_W	576	224-238 x 222-224	222-224/222-238/224-224/224-238	145:133:141:119	538	222-224/222-238/224-224/224-238	1:1:1:1	2.94	4.0E-01
	Qr_N Qr_S	175	224-238 x 222-228	222-224/222-238/224-228/228-238	41:37:52:35	154	222-224/222-238/224-238/238-238	1:1:1:1	4.44	2.2E-01
	Qr_P Qr_11	124	216-224 x 224-230	216-224/216-230/224-224/224-230	37:28:19:35	119	216-224/216-230/224-224/224-230	1:1:1:1	6.68	8.3E-02
FIR013	Qr_N Qr_S	175	138-141 x 141-141	137-141/141-141	81:92	173	138-141/141-141	1:1	0.70	4.0E-01
FIR028	Qr_N Qr_W	576	216-216 x 218-220	216-218/216-220	263:239	502	216-218/216-220	1:1	1.15	2.8E-01
	Qr_N Qr_S	175	208-216 x 216-216	208-216/216-216	83:80	163	208-216/216-216	1:1	0.06	8.1E-01
	Qr_P Qr_11	124	216-220 x 216-222	216-216/216-220/216-222/220-222	26:33:21:19	99	216-216/216-220/216-222/220-222	1:1:1:1	4.72	1.9E-01
FIR035	Qr_N Qr_W	576	146-148 x 146-148	146-146/146-148/148-148	151:273:148	572	146-146/146-148/148-148	1:2:1	1.21	5.5E-01
	Qr_N Qr_S	175	146-146 x 146-148	146-146/146-148	88:84	172	146-146/146-148	1:1	0.09	7.6E-01
FIR031	Qr_N Qr_W	576	141-141 x 143-151	141-143/141-151	319:179	498	141-143/141-151	1:1	39.36	3.5E-10
	Qr_N Qr_S	175	141-141 x 159-161	141-159/141-161	74:77	151	141-159/141-161	1:1	0.06	8.1E-01
	Qr_P Qr_11	124	141-141 x 137-169	137-141/141-169	60:49	109	137-141/141-169	1:1	1.11	2.9E-01
quru-GA-0E09	Qr_N Qr_W	576	204-208 x 210-224	204-210/204-224/208-210/208-224	139:127:137:116	519	204-210/204-224/208-210/208-224	1:1:1:1	2.58	4,61E-01
	Qr_N Qr_S	175	188-200 x 204-208	188-204/188-204/200-204/200-208	50:43:41:34	168	188-204/188-204/200-204/200-208	1:1:1:1	3.10	3.8E-01
	Qr_P Qr_11	124	190-192 x 196-234	190-196/190-234/192-196/192-234	27:25:34:16	102	190-196/190-234/192-196/192-234	1:1:1:1	6.47	9.1E-02

Table 1: Continued

Locus	Family	No. siblings	Parent genotype	Progeny genotype	Observed segregation	No. of siblings genotyped	Expected progeny genotype	Expected segregation	Chi-square	p-value
quru-GA-1F07	Qr_N Qr_W	576	311-331 x 323-327	311-323/311-327/323-331/327-331	129:142:141:117	529	311-323/311-327/323-331/327-331	1:1:1:1	3.14	3.7E-01
	Qr_N Qr_S	175	305-315 x 323-327	305-323/305-327/315-323/315-327	34:42:43:50	169	305-323/305-327/315-323/315-327	1:1:1:1	3.05	3.8E-01
	Qr_P Qr_11	124	321-331 x 327-329	321-327/321-329/327-331/329-331	24:24:28:36	112	321-327/321-329/327-331/329-331	1:1:1:1	3.43	3.3E-01
GOT040	Qr_N Qr_S	175	233-239 x 239-239	233-239/239-239	85:88	173	233-239/239-239	1:1	0.05	8.2E-01
	Qr_P Qr_11	124	233-239 x 239-241	233-239/233-241/239-239/239-241	21:33:28:38	120	233-239/233-241/239-239/239-241	1:1:1:1	5.27	1.5E-01
PIE125	Qr_N Qr_W	576	154-157 x 157-160	154-157/154-160/157-157/157-160	139:157:138:130	564	154-157/154-160/157-157/157-160	1:1:1:1	2.77	4.3E-01
	Qr_N Qr_S	175	154-157 x 154-157	154-154/154-157/157-157	26:99:47	172	154-154/154-157/157-157	1:2:1	9.06	1.1E-02
	Qr_P Qr_11	124	151-154 x 154-157	151-154/151-157/154-154/154-157	36:30:24:34	124	151-154/151-157/154-154/154-157	1:1:1:1	2.71	4.4E-01
VIT023	Qr_N Qr_W	576	116-116 x 116-119	116-116/116-119	307:256	563	116-116/116-119	1:1	4.45	3.5E-02
quru-GA-0C11	Qr_N Qr_W	576	213-217 x 205-221	205-213/205-217/213-221/217-221	150:146:120:99	515	205-213/205-217/213-221/217-221	1:1:1:1	13.29	4.1E-03
	Qr_N Qr_S	175	213-217 x 203-225	203-213/203-217/213-225/217-225	48:42:40:32	157	203-213/203-217/213-225/217-225	1:1:1:1	3.23	3.6E-01
	Qr_P Qr_11	124	209-211 x 205-213	205-209/205-211/209-213/211-213	29:38:24:28	119	205-209/205-211/209-213/211-213	1:1:1:1	3.52	3.2E-01

Acknowledgments

We thank Alexandra Dolynska for her assistance in the laboratory and Jeremias Götz, Gudrun Diederich and Gerold Dinkel for their technical support collecting and planting the seeds. We also thank Dr. Katharina B. Budde for her comments and suggestions.

References

Aldrich PR, Michler CH, Sun W, Romero-Severson J (2002) Microsatellite markers for northern red oak (Fagaceae: *Quercus rubra*). Molecular Ecology Notes 2(4):472-474. https://doi.org/10.1046/j.1471-8286.2002.00282.x

Beavis W (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. Proceedings of the forty-ninth annual corn and sorghum industry research conference 250:266.

Bodénès C, Chancerel E, Ehrenmann F, Kremer A, Plomion C (2016). High-density linkage mapping and distribution of segregation distortion regions in the oak genome. DNA Research 23(2):115-124. https://doi.org/10.1093/dnares/dsw001

Buckler ES IV, Phelps-Durr TL, Buckler CSK, Dawe RK, Doebley JF, Holtsford TP (1999) Meiotic drive of chromosomal knobs reshaped the maize genome. Genetics 153(1):415-426. https://doi.org/10.1093/genetics/153.1.415

Covarrubias-Pazaran G, Diaz-Garcia L, Schlautman B, Salazar W, Zalapa J (2016) Fragman: an R package for fragment analysis. BMC Genetics 17:1-8. https://doi.org/10.1186/s12863-016-0365-6

Deguilloux MF, Dumolin-Lapègue S, Gielly L, Grivet D, Petit RJ (2003) A set of primers for the amplification of chloroplast microsatellites in Quercus. Mol Ecol Notes 3:24–27. https://doi.org/10.1046/j.1471-8286.2003.00339.x

Díaz M, Møller AP, Pulido FJ (2003) Fruit abortion, developmental selection and developmental stability in *Quercus ilex*. Oecologia 135:378–385. https://doi.org/10.1007/s00442-003-1202-y

Dominguez-Flores T, Budde KB, Carlson JE, Gailing O (accepted) Distance dependent mating but considerable pollen immigration in an isolated *Quercus rubra* planting in Germany. European Journal of Forest Research. https://doi.org/10.1007/s10342-024-01704-x

Durand J, Bodénès C, Chancerel E, Frigerio JM, Vendramin G, Sebastiani F, Buonamici A, Gailing O, Koelewijn HP, Villani F, Mattioni C, Cherubini M, Goicoechea PG, Herrán A, Ikaran A, Cabané C, Ueno S, Alberto F, Dumoulin PY, Guichoux E, de Daruvar A, Kremer A, Plomion C (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11:1-13. https://doi.org/10.1186/1471-2164-11-570

El-Kassaby YA, Lstibůrek M (2009) Breeding without breeding. Genetics Research 91(2):111-120. https://doi.org/10.1017/S0016672309000070X

Gailing O, Staton ME, Lane T, Schlarbaum SE, Nipper R, Owusu SA, Carlson JE (2017) Construction of a framework genetic linkage map in *Gleditsia triacanthos* L. Plant Molecular Biology Reporter 153(1):415-426. https://doi.org/10.1007/s11105-016-1012-0

Gauthier S, Bernier P, Kuuluvainen T, Shvidenko AZ, Schepaschenko DG (2015) Boreal forest health and global change. Science 349(6250):819-822. https://doi.org/10.1126/science.aaa9092

Götz J, Gailing O (2022) A novel set of polymorphic chloroplast microsatellite markers for northern red oak (*Q. rubra* L.). Plant Genet Resour Charact Util 20: 174–177. https://doi.org/10.1017/S1479262122000156

Gręda A, Woziwoda B, Dyderski MK, Jagodziński AM, Frelich LE (2022) Acorn production of introduced *Quercus rubra* is more strongly impacted by the weather than by the forest site (a case study from Poland). Agricultural and Forest Meteorology 327:109228. https://doi.org/10.1016/j.agrformet.2022.109228

Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10(3):551-555. https://doi.org/10.1111/j.1755-0998.2009.02787.x

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/s12864-017-3765-8

Krautwurst M, Past F, Kersten B, Bubner B, Mueller NA (2023) Identification of full-sibling families from natural single-tree ash progenies based on SSR markers and genome-wide SNPs. bioRxiv 2023–07. https://doi.org/10.1101/2023.07.18.549475

Kuang H, Richardson T, Carson S, Wilcox P, Bongarten B (1999) Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. I. Genetic map with distorted markers. Theoretical and Applied Genetics 98:697-703. https://doi.org/10.1007/s001220051123

Margarido GRA, Souza AP, Garcia AA (2007) OneMap: software for genetic mapping in outcrossing species. Hereditas 144(3):78-79. https://doi.org/10.1111/j.2007.0018-0661.02000.x

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7(5):639-655. https://doi.org/10.1046/j.1365-294x.1998.00374.x

Møller AP (1997) Developmental selection against developmentally unstable offspring and sexual selection. Journal of Theoretical Biology 185(4):415-422. https://doi.org/10.1006/jtbi.1996.0332

Owusu SA, Schlarbaum SE, Carlson JE, Gailing O (2016) Pollen gene flow and molecular identification of full-sib families in small and isolated population fragments of *Gleditsia triacanthos* L. Botany 94(7):523-532. https://doi.org/10.1139/cjb-2015-0244

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Sork VL, Bramble JE (1993) Prediction of acorn crops in three species of North American oaks: *Quercus alba*, *Q. rubra* and *Q. velutina*. Annales des Sciences Forestières 50:128s–136s. https://doi.org/10.1051/forest:19930712

Sullivan AR, Lind JF, McCleary TS, Romero-Severson J, Gailing O (2013) Development and characterization of genomic and gene-based microsatellite markers in North American red oak species. Plant Molecular Biology Reporter 31:231–239. https://doi.org/10.1007/s11105-012-0495-6