Supplementary Appendix A1: sampling and genotyping details

This is the supplementary appendix of the manuscript: "Genetic consequences of recent contraction and expansion of European beech populations" by T. Lander, E.K. Klein3, A. Roig, and S. Oddou-Muratorio.

1. Sampling design, sample collection and measurements

In this study, we combined three data sets consisting in:

A. Three "intensively" studied plots (West-N2, West-384 and West-257-2), within the NW_REF region, with exhaustive sampling of adult trees (579 trees in total). This data set is previously described and analyzed in Oddou-Muratorio, Gauzere, Bontemps, Rey, & Klein (2018) and in Lander, Oddou-Muratorio, Prouillet-Leplat, & Klein, (2011).
B. A set of 48 plots where a total of 1353 adult trees were sampled non-exhaustively (see detail protocol below), previously described and analysed in Lander, Oddou-Muratorio, Prouillet-Leplat, & Klein, (2011).
C. A set of 20 plots where a total of 600 adult trees were sampled non-exhaustively specifically for this study, following the protocol of Lander, Oddou-Muratorio, Prouillet-Leplat, & Klein (2011).

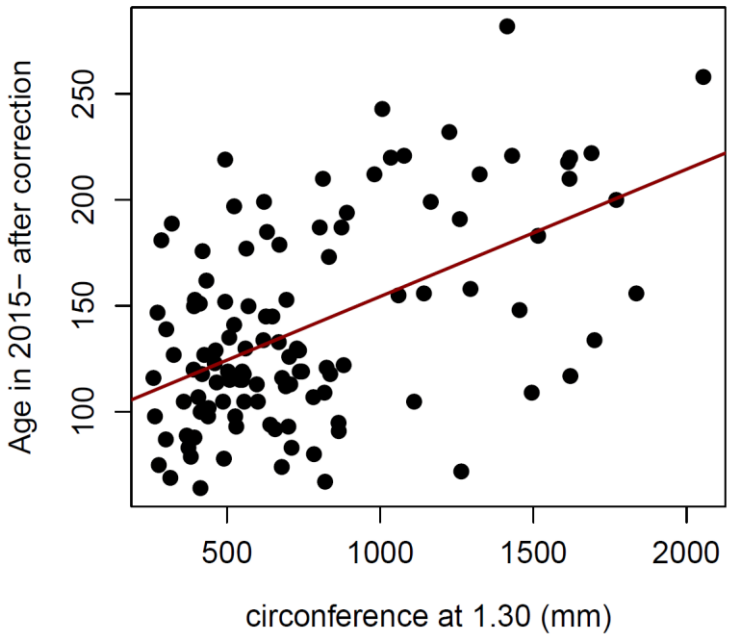
For data sets B and C, the same sampling protocol was applied. We selected 28 adult beech trees on average (up to a maximum of 40 individuals) in an area of ~50 m radius so that all trees were separated by at least 3 meters. Circumference at breast height and coordinates were recorded for each tree. Moreover, half of the trees were chosen because they had the largest circumference in the stand ("Old" trees) and half had the smallest circumference (but > 160 mm; "Young" trees).

In the three intensive study plots of data set A, all adult trees were exhaustively sampled, and their circumference at breast height and coordinates were recorded. This intensive sampling effort allowed us to estimate the probability of vegetative reproduction. Indeed, beech is known to have the ability to reproduce vegetatively through resprouting around cut or fallen trees. Note that in data set A, if a tree obviously had multiple stems, only the largest stem was sampled. After careful elimination of clonal individual (see paragraph 2 below), on each of the three plots, we selected 20 small ("Young") and 20 large ("Old") trees based on their circumference (40 trees in total per plot), ensuring a minimum distance between them >3 m.

For data sets A and B leaf samples were collected in 2008 and 2009 and stored at INRA Avignon laboratory at –20°C. For data set C, leaf samples were collected in 2015 and stored at the INRA Avignon laboratory at ambient temperature (using desiccant for conservation).

The maximal age of a tree within each plot was estimated based on two approaches. In 44 plots among the 71 studied, two trees per plot were selected to be cored at 1.30 m (targeting the largest and the smallest one). Cores were read to estimate tree age from ring profile. The oldest age was considered as a rough estimate of the maximal age of a tree in 2015 in a given plot. In the other plots, this maximal age of a tree was estimated based on detailed dendroecological studies (F Jean & P Dreyfus, pers comm) which combined tree ring profile, size distribution of trees, and historical record of management operation within each plot. We examined the relationship between circumference and age to validate these estimations (Figure A1.1 below): using a simple linear model, age significantly increased with circumference (coeff= 0.06, p-value <0.001).

Figure A1.1: Relationship between circumference at 1.30 m and age estimated from cored tree on the different studied plot.



2. Genotyping and identification of clones

All individuals were genotyped using 13 microsatellite markers: FS1-15 and FS3-04 (Pastorelli et al., 2003); sfc0007-2, sfc0018, sfc0036, sfc1063, and sfc1143 (Asuka, Tani, Tsumura, & Tomaru, 2004); mfs7 (Vornam, Decarli, & Gailing, 2004); csolfagus-5, csolfagus-6, csolfagus-25, csolfagus-29, and csolfagus-31 (Vendramin, pers. comm.).

Total DNA was extracted from 50 mg wet weight frozen leaf material using the protocol for the DNeasy 96 Plant Kit (QIAGEN). The concentration and purity of the DNA were estimated by measuring the absorbance at 260nm and 280nm in a spectrophotometer and by using pulse-field gel electrophoresis on agarose gel. Extracted DNA was stored at -20°C. Samples were amplified using the Type-it Microsatellite PCR Kit (QIAGEN). The PCR program was: 94°C for 15min, 30 cycles of 94°C for 30sec, 60°C for 90sec , 72°C for 60sec, final extension 72°C for 30min. Some PCR products were analysed on a MegaBACE 1000 sequencer and scored with MegaBACE Genetic Profiler software (Amersham Biosciences 2003 version 2.2) against an internal size standard (ET400 DNA size markers). Other PCR products were analysed on an ABI3930 sequencer and scored with the software against an internal size standard (ET400 DNA size markers). Automatic allele assignment was checked and revised manually twice to ensure consistency of genotyping. The different data sets obtained on different sequencer were assembled and harmonized.

Among the 2531 sampled trees, 46 clone groups made up of 97 individuals were identified (mostly on the densely sampled plots West-N2, 384 and 257-2). The probability that two individuals that are not genetically identical would show the same genotype (the probability of identity) with the 13 loci employed for this study was 8.1 x 10-14 (GenAlEx, Peakall & Smouse, 2006). Hence, it was assumed that the clone groups identified represented true vegetative reproduction. We selected a single individual for each clone group, leading to 2421 distinct genets.

We further excluded 379 genets in plots West-N2, 384 and 257-2 (to be consistent with the sampling protocol described in paragraph 1 above). Finally, we retained a maximum of 40 genets per plot and 2042 distinct genets in total.

3. Quality of the marker sets, and basic statistics.

Null allele : We first estimated Null Allele Frequencies (NAF) per locus and per population using the ML-null software (Kalinowski & Taper, 2006). We focused on three loci showing NAF > 0.05 in a large number of population, and overall large average NAF (Table A1.2). First, locus Sfc1063 showed NAF > 0.05 in 31 populations out of 71, with an average NAF=0.091. Then, locus Sfc0018 showed NAF > 0.05 in 21 populations out of 71, but an average NAF=0.04. Finally, locus mfc7 showed an average NAF=0.06, but NAF > 0.05 in "only" 13 populations out of 71.

→*We decided to remove Sfc1063, and to keep all the other 12 loci or genetic analyses.*

Linkage disequilibrium: Test for linkage disequilibrium was investigated using Genepop 4.0 (Rousset, 2008). Significant linkage disequilibrium was detected between all pairs of locus across all populations (66 tests). Among a total of 4686 tests for linkage disequilibrium between pairs of loci within each population, 357 (7.6%) were significant but they did not cluster by locus or population.

Missing data: the proportion of missing data ranged from 1.76% (locus csolfagus06) to 19.83% (locus mfc7) with a mean of 4.24% (Table A1.3)

HWE equilibrium per locus: We detected 45 cases of significant excess of homozygosity of the 852 possible plot × locus pairs (5.8%), 15 of which were obtained at locus Sfc0018. This is consistent with the prevalence of null allele suspected at this locus. We also detected 22 cases of significant excess of heterozygosity of the 852 possible plot × locus pairs (2.6%), but they did not cluster by locus or population.

Overall, four loci showed a significant deviation from HWE: locus Sfc0018 showed a significant excess of homozygosity, while FS3-04, sfc1143 and csolfagus-31 showed a significant deficit of homozygosity (Table A1.3)

Estimation of allelic richness (Ar) :

Figure A1 below summarizes the diversity and the quality of the marker set for the 2042 individuals and 12 markers.

Table A1.2: Estimated Null Allele Frequencies (NAF) per locus in the 71 studied plots. Only non-null NAF-values are shown. The two last lines show the number of plots with NAF>0.05, and the mean NAF averaged over plots (including null values).

Plot	Locus												
	FS1-15	**FS3-04**	**Sfc007-2**	**Sfc0018**	**Sfc0036**	**Sfc1063**	**Sfc1143**	**Csol-fagus25**	**Csol-fagus29**	**Csol-fagus31**	**Csol-fagus05**	**Csol-fagus06**	**mfc7**
E_1217						0.01				0.02	0.02		
E_1231	0.03	0.02	0.02			0.01	0.07		0.01	0.06	0		
E_21			0.02	0.03	0.05	0.1	0.03	0.05		0.01		0.02	0.02
E_28	0.03				0.02			0.07				0.01	
E_68		0.01	0.01	0.15		0.26			0.09	0.11		0.02	0.07
E_940		0.06	0.02	0.11	0.04	0			0.01				0.13
E_979		0.05	0.11		0.08	0.01		0.06		0.08			0.17
E_980	0.03	0.08		0.15		0.01			0.13				0.03
E_132		0.04		0.04		0.02	0.02				0.01	0.03	
E_168		0.02										0.01	
E_170			0.03				0.03		0.03				

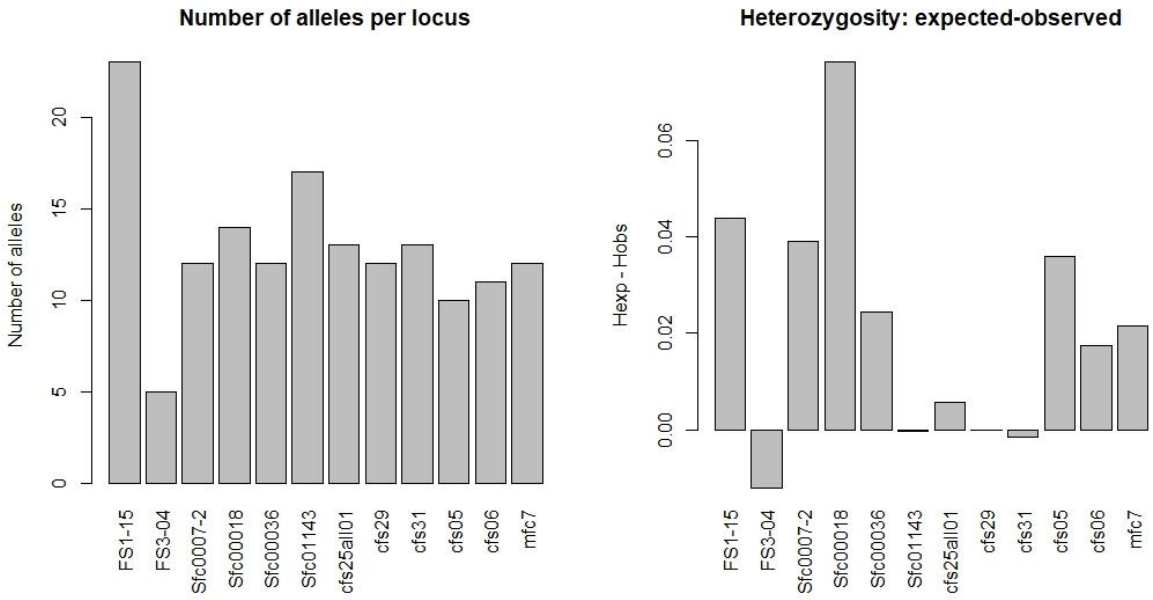
E_184		0.03							0.04		0.02		
E_186			0.03									0.08	0
E_188		0.04		0.03		0.03					0.05	0.02	
E_191	0.04	0.11	0.04		0.06	0.07					0.03		
E_203								0.08					
E_217			0.04										
E_224				0.06	0.01	0.36	0.01	0.14			0.01		
E_243				0.08	0.01			0.01	0.02				
E_256			0.06	0.02									
E_259	0.05			0.01	0.03			0.01		0.02			
E_267		0.07		0.08		0.05							0.06
E_274	0.03			0.03	0.01	0.02					0.01	0.03	
E_61	0.04			0.14		0.05	0			0.02		0.02	
E_76		0.14		0.03		0.07			0.03		0.03		
E_91	0.07				0.04								
E_E						0.25							
S_1101				0.07	0.04	0.01	0.02						0.97
S_1225				0.12		0.01			0.08		0.01		0.08
S_1250	0.01	0.06			0.04	0					0.04		0.98
S_1332			0.04	0.08		0.01	0.01						
S_1727	0				0.01	0	0.16		0.13		0.05		0.01
S_1739											0.02		0.98
S_1755	0.07		0.02			0.01	0.05				0.06	0.01	
S_1832						0	0.03		0.06			0	
S_1913	0.03		0.12	0.07	0.02	0.98			0.04		0.16		
S_2007				0.3		0.01				0.06			
S_2046					0.07	0	0.02	0.94	0	0.02	0.03		0.01
S_901	0.05	0.02	0.01	0.2							0.01		
S_905	0.05			0.03		0.01						0.03	0.06
S_985			0.02			0.01		0.06			0.01	0	
S_S03				0.08		0.18		0.12		0.02	0.02		0.03
S_23	0.03		0.07	0.05		0.1						0	
S_93	0.01			0.29		0.18					0.05	0.01	
S_01			0.08	0	0.04	0.33		0.02		0.01	0.02	0.01	0.06
S_05			0.05	0.05		0.18		0.07		0.02	0.05		
W_N2	0.1			0.05		0.02		0.01		0.02		0.01	0.01
W_N3			0.05			0.27		0.17			0.03	0.01	
W_10	0.05	0.05			0.05	0.03	0.01				0.07		0.02
W_A	0.02		0.02	0.08		0.01				0.03		0.03	
W_B		0.03				0.05			0.03		0.02		0.09
W_C	0.03		0.05	0.01	0.03						0.01		
W_D				0.02		0.27		0.03	0.01				
W_122		0.1		0.12		0.02			0.05				0.07
W_142	0.02		0.04	0.03		0.07	0.02				0.04	0.04	
W_164	0.08		0.09			0.12			0.07				
W_206			0.08	0.04		0.05	0.04		0.05		0.13		

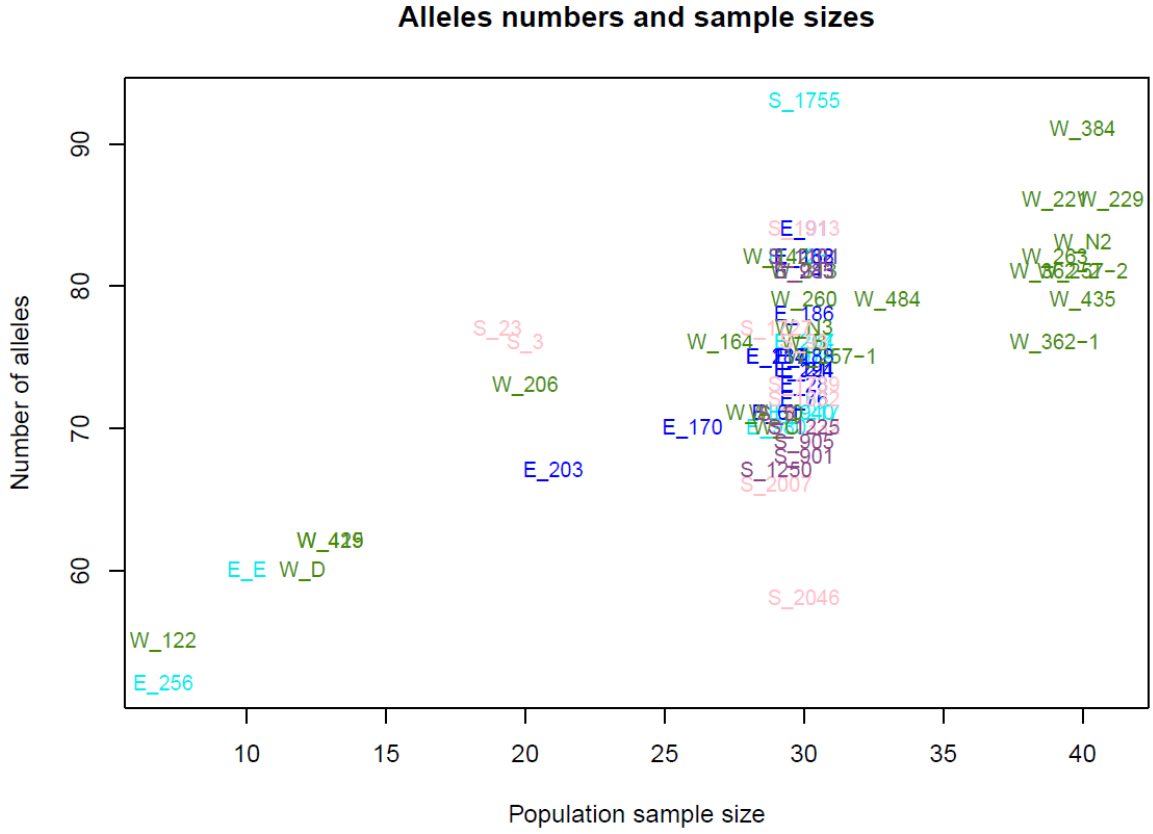
W_221	0.01		0.05			0.27							
W_229						0.04		0.03		0.01			
W_257-1	0.1		0.04	0.13		0.09						0	
W_257-2	0.08		0.02	0.08	0.02	0.07		0.01				0	
W_263			0.04	0.05		0.19			0.05	0.04			
W_313				0.05	0.05	0.06	0	0.19	0.08		0.12		
W_362-1			0.01	0.02		0.1	0.01						
W_362-2				0.08	0.05	0.08	0.02			0.04			
W_384	0		0	0.02		0.14	0.02				0.14		
W_260		0.03	0.01	0.02	0.04	0.24					0.14	0.06	
W_415				0.04	0.03	0.27					0.01		0.07
W_429				0.05		0.27							
W_435					0.01	0.3		0.01	0.04		0.13		0.02
W_484	0.09				0.02	0.17	0.03				0.11		
#pop$_{NAF>0.05}$	8	8	8	21	4	31	2	10	8	4	9	2	13
Mean NAF	0.02	0.01	0.02	0.04	0.01	0.09	0.01	0.03	0.01	0.01	0.02	0.01	0.06

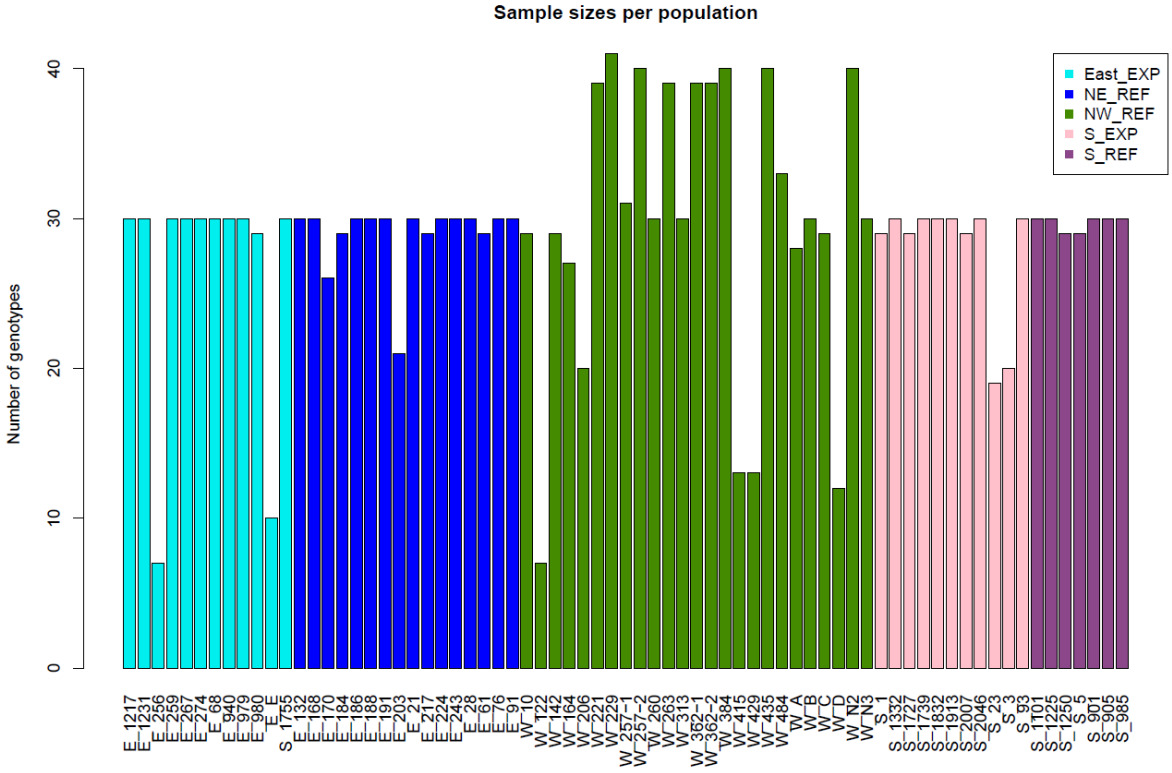
Table A1.3: Average diversity values per locus

Locus	Na	Missing data	Ho	Hs	Ht	F_{ST}	F_{IS}
FS1.15	23	2.45%	0.825	0.833	0.865	0.037	0.010
FS3.04	5	3.67%	0.503	0.475	0.488	0.027	-0.060
Sfc0007.2	12	3.62%	0.739	0.760	0.784	0.031	0.027
Sfc0018	14	3.87%	0.659	0.716	0.735	0.025	0.080
Sfc0036	12	2.64%	0.661	0.661	0.684	0.034	0.001
Sfc01143	17	1.86%	0.816	0.793	0.817	0.029	-0.029
csolfagus25	13	4.06%	0.527	0.519	0.537	0.032	-0.015
csolfagus29	12	2.20%	0.549	0.536	0.552	0.029	-0.025
csolfagus31	13	2.45%	0.756	0.728	0.754	0.035	-0.039
csolfagus05	10	2.55%	0.718	0.731	0.753	0.028	0.018
csolfagus06	11	1.76%	0.841	0.834	0.857	0.027	-0.008
mfc7	12	19.83%	0.587	0.602	0.625	0.036	0.025

Figure A1: Summary statistics of the data sets






Sample sizes per population
Number of genotypes
0
10
20
30
40
East_EXP
NE_REF
NW_REF
S_EXP
S_REF
E_1217
E_1231
E_256
E_259
E_267
E_274
E_68
E_940
E_979
E_980
E_E
S_1755
E_132
E_168
E_170
E_184
E_186
E_188
E_191
E_203
E_21
E_217
E_224
E_243
E_28
E_61
E_76
E_91
W_10
W_122
W_142
W_164
W_206
W_221
W_229
W_257_1
W_257_2
W_260
W_263
W_313
W_362_1
W_362_2
W_384
W_415
W_429
W_435
W_484
W_A
W_B
W_C
W_D
W_N2
W_N3
S_1
S_1332
S_1727
S_1739
S_1832
S_1913
S_2007
S_2046
S_23
S_3
S_93
S_1101
S_1225
S_1250
S_5
S_901
S_905
S_985

References

Asuka, Y., Tani, N., Tsumura, Y., & Tomaru, N. (2004). Development and characterization of microsatellite markers for Fagus crenata Blume. *Molecular Ecology Notes*, *4*(1), 101–103. doi:10.1046/j.1471-8286.2003.00583.x

Kalinowski, S. T., & Taper, M. L. (2006). Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics*, *7*(6), 991–995. doi:10.1007/s10592-006-9134-9

Lander, T. A., Oddou-Muratorio, S., Prouillet-Leplat, H., & Klein, E. K. (2011). Reconstruction of a beech population bottleneck using archival demographic information and Bayesian analysis of genetic data. *Molecular Ecology*, *20*(24), 5182–5196. doi:10.1111/j.1365-294X.2011.05356.x

Oddou-Muratorio, S., Gauzere, J., Bontemps, A., Rey, J. F., & Klein, E. K. (2018). Tree, sex and size: Ecological determinants of male vs. female fecundity in three Fagus sylvatica stands. *Molecular Ecology*, *27*(15), 3131–3145. doi:10.1111/mec.14770

Pastorelli, R., Smulders, M. J. M., VAN'T Westende, W. P. C., Vosman, B., Giannini, R., Vettori, C., & Vendramin, G. G. (2003). Characterization of microsatellite markers in Fagus sylvatica L. and Fagus orientalis Lipsky. *Molecular Ecology Notes*, *3*(1), 76–78. doi:10.1046/j.1471-8286.2003.00355.x

Peakall, R., & Smouse, P. E. (2006). genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*(1), 288–295. doi:10.1111/j.1471-8286.2005.01155.x

Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106. doi:10.1111/j.1471-8286.2007.01931.x

Vornam, B., Decarli, N., & Gailing, O. (2004). Spatial Distribution of Genetic Variation in a Natural Beech Stand (Fagus sylvaticaL.) Based on Microsatellite Markers. *Conservation Genetics*, *5*(4), 561–570. doi:10.1023/B:COGE.0000041025.82917.ac