Pre-disease Levels of Genetic Diversity and Differentiation among Common Ash (*Fraxinus excelsior* L.) Seedlots in Austria

BERTHOLD HEINZE[1]* AND BARBARA FUSSI[2]

[1] *Austrian Federal Research Centre for Forests (BFW), Department of Forest Genetics, Seckendorff-Gudent Weg 8, 1130 Vienna, Austria*

[2] *Bavarian Office for Forest Seeding and Planting (ASP), Forstamtsplatz 1, 83317 Teisendorf, Germany*

** Corresponding author: berthold.heinze@bfw.gv.at, tel. +43 1 878382219*

Heinze, B. and Fussi, B. 2017. Pre-disease Levels of Genetic Diversity and Differentiation among Common Ash (*Fraxinus excelsior* L.) Seedlots in Austria. *Baltic Forestry* 23(1): 198-208.

Abstract

Ash is an important component of forestry in Austria; its loss due to the dieback disease would be a great challenge for many forest owners. We investigated seed material that was harvested in 2001, just prior to the onset of the disease. Seeds from at least ten (allegedly) separate trees per stand were obtained from commercial harvest lots, from six different stands in Austria. The separate sampling from at least ten seed-bearing trees of ash is a legal requirement in Austria. Levels of genetic differentiation on the basis of six microsatellite markers were low, but somewhat higher than in other typical European forest trees. Stands along the Danube river seemed to share more genetic similarity with each other than with two stands in the Alps. In comparison, within the stands, most single tree seed lots were highly differentiated and they mostly fitted to the stands of origin with their genetic patterns. An attempt was made at reconstructing the unknown genotypes of the mother trees of the seeds from the offspring data. This led to the presumable identification of cases where these mother trees shared more alleles than expected, and their seed lots were closer genetically than on average. It also revealed cases where single seeds did not fit into their lot genetically (as defined by Mendelian rules). The data reported here confirm that detailed information on the genetic background of seed can be obtained from such structured samples, supporting law enforcement. It further confirms that harvesting from a minimum of ten trees leads to seed that more comprehensively reflects levels of genetic diversity in the whole stand. The data presented can be used as a baseline for investigating any genetic effects of the progressing disease in the future.

Keywords: *Fraxinus excelsior*, *Hymenoscyphus fraxineus*, microsatellite markers, genetic diversity, seed, genetic differentiation, forest reproductive material

Introduction

Common ash (*Fraxinus excelsior*) in Europe, the subject of this volume, has some particular features that distinguish it from other hardwood forest trees. In contrast to the more continuously distributed, stand-forming deciduous trees like beech (*Fagus sylvatica*) and the oaks (*Quercus* sp.), ash has a more ecotypic pattern of distribution with some local abundance, but in general, its more scattered distribution (Gömöry et al. 2012) is correlated with that of nutrient-rich soils with good levels of water availability, especially in spring (Weiser 1995). Therefore, the patterns of genetic diversity of this species across the landscape have received some interest in the scientific community. Especially after the publication of several microsatellite DNA markers for this species (Brachet et al. 1999, Lefort et al. 1999), such studies have addressed patterns across countries or regions (e.g. Heuertz et al. 2001, Morand et al. 2002, Ferrazzini et al. 2007, Sutherland et al 2010, Gömöry et al. 2012, Fussi and Konnert 2014, Beatty

et al. 2015) or even across Europe (Heuertz et al. 2004). The peculiar flowering system of the species, with predominantly male, female, or hermaphrodite trees, has also raised attention genetically (e.g. Morand-Prieur et al. 2003, Heuertz et al. 2003), as well as its possible hybridization with the sister species, *Fraxinus angustifolia* (Morand-Prieur et al. 2002, Gerard et al. 2006, 2013; Heuertz et al. 2006, Lexer et al. 2004, Thomasset et al. 2011, 2013). Gene flow in these species has been assessed, *i.a.*, by Heuertz et al. (2003), Bacles et al. (2006); Bacles and Ennos (2008); Gömöry et al. (2012), and Thomasset et al. (2014).

The progress of the disease (see other contributions in this volume) makes it necessary to discern effects of e.g. gene flow and hybridization on the population genetics of this species. For example, it would be interesting to know whether progressive fragmentation (because of tree mortality in affected regions) decreases, or rather, even increases (Bacles et al. 2006; Bacles and Ennos 2008) gene flow; or whether hybridization with *F. angustifolia* affects disease tolerance. For all these questions, comparing genetic analyses before and after the onset of the disease in a particular region would provide valuable data.

The FRAXBACK Cost Action FP1103 (www.fraxback.eu) has brought together scientists with an interest, *i.a.*, in these genetic questions. In the frame of a previous project, 'RAP – Realising Ash' Potential' (contract QLK5-CT-2001-00631 of the EC's Fifth Framework Programme), genetic diversity and relatedness data were obtained also for Austria. The main rationale for this work was an assessment of seed harvesting practices, and whether they would conform to national regulations. These regulations in Austria state that seeds in ash have to be harvested from at least ten different trees in an approved seed stand. A sample of a handful of seeds from each of these trees has to be sent to the BFW research station. For checking conformity with seed harvest regulations, such single seed lots from six different stands were analysed with microsatellite markers. Exactly at the end of the investigation mentioned above, any interest in planting ash stopped more or less completely in Austria (Heinze et al. 2017), and the results have so far remained unpublished. However, the steady progress of the disease now opens new possibilities for presenting the data and re-analysing it in order to provide a baseline for investigating any genetic effects that the disease has on the tree species, now and in the future. Genetic distances between the seed lots, and an attempt at reconstructing the (unknown) maternal genotypes, served for this purpose and are presented here.

Material and Methods

Seed and DNA extraction

Samples of *Fraxinus excelsior* were collected in seed lots in different regions of Austria at the time of harvest (autumn of year 2001) by local forest personnel (Table 1). Seeds from 10 to 13 trees (adult stage – 'mother trees', in total 65 trees) from each of the six seed lots were sent to the BFW research station separately. Seeds from each tree were imbibed in water for softening. Eight to ten seeds per tree (with very few exceptions of only four or seven, or up to 12 seeds) were used in DNA extraction, resulting in a total of 573 seeds (seed stage) investigated genetically. Total DNA was extracted from excised embryos using the SIGMA GenElute^{TM} Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, Vienna, Austria).

Table 1. Origin of seed (harvest year: 2001) for this study

Stand	District (province)	Field name	Coordinates (WGS84)	Elevation (m)	Number of mother trees
ASL02	Eferding (OÖ)	Polsing	E 14°05'09.38'' N 48°16'06.68''	279	13
ASL23	Hartberg (STMK)	Riegler-viertel	E 15°49'19.11'' N 47°29'15.71''	928	10
ASL27	Klagenfurt (KTN)	Loiblthal	E 14°15'28.85'' N 46°27'55.73''	775	10
ASL50	Melk (NÖ)	Aggsbach	E 15°30'31.53'' N 48°18'26.40''	466	10
ASL65	Tulln (NÖ)	Langen-schönbichl	E 15°58'51.36'' N 48°19'47.86''	180	10
ASL98	Lilienfeld (NÖ)	Keeramt	E 15°30'04.80'' N 47°49'47.28''	821	12
(total number of trees)					65

Genetic markers

For molecular analysis six primer pairs of microsatellite loci were chosen which showed high polymorphism in initial studies (FEMSATL4, 11, 12, 16, 19, and M2-30; Lefort et al. 1999; Brachet et al.1999). These markers were amplified by polymerase chain reactions (PCR) performed in a mix containing 2 mM MgCl2, 0.5 units Taq Polymerase (Platinum; in Platinum Puffer, Thermofisher, Vienna, Austria) per reaction, 0.2 µM of each primer, 0.2 mM dNTPs, 1 µl of template DNA (approx. 10-50 ng/µL) in a total reaction volume of 20 µl. The thermal cycling profile consisted of an initial denaturing step for 3 min at 94 °C, 10 cycles of 50 s at 94 °C and 1 min at 70 °C, followed by 35 cycles of 30 s 94 °C, 55 °C (annealing temperature) for 50 s, and 2 min at 70 °C. PCR products were separated by electrophoresis in 8 % polyacrylamide gels for 2 h at 40 W and visualized by SYBRGold (Molecular Probes, ThermoFisher) staining on a standard UV–transilluminator (Heinze et al. 2014). Ladders of size standards (10 and 100 bp DNA ladders) were used for sizing bands. Selected alleles were loaded side by side on every gel for additional calibration. Pictures of the gels were taken and alleles recorded with the program Gene Profiler 4.03 (Scanalytics, Inc., Fairfax, VA, USA).

Data analyses

Parameters of genetic variation were calculated for each of the six seed lots separately, and for the total, by the Fstat software (Goudet 1995): (1) the average number of observed alleles A, (2) gene diversity or expected heterozygosity H_E (Nei 1987), (3) Wright´s inbreeding coefficient F_{IS}, (4) observed heterozygosity H_O (for each locus over all populations in seed samples), (5) Weir & Cockerham's F-statistics for describing genetic diversity within (F_{SP}) and between seed lots (F_{ST}). Hierarchical F-statistics were established by computing F_{IT} for comparing individual samples to all samples (individuals to total), along with F_{ET}, calculating the relationship of subpopulations (seeds from single trees) to the total data set, and finally F_{IS}, differentiating individuals within subpopulations. (6) Departure from Hardy-Weinberg-equilibrium, as implemented in the Fstat software (Goudet 1995), was also assessed.

With the program STRUCTURE (vers. 2.1 and 2.3.4, Prichard et al. 2000, Falush et al. 2003, 2007), which applies a widely used model-based clustering method, we attempted to assign individual multi-locus genotypes (the seeds) to a user-defined number K of clusters or gene pools. STRUCTURE (ver. 2.1) was run for 10^4 iterations after a burn-in period of 10^4 on the total dataset of 573 individuals; trial runs for 10^6 iterations and burn-in period runs did not result in substantial differences to the 10^4 finally used for our dataset. The number of clusters were set as K = 1, 2, …, 80, and 10 runs were done at each K. We used the model without admixture in the runs initially (assuming that the analysed stands are situated too far apart from each other for any significant contemporary pollen flow), and repeated these with the admixture model as well. No prior information on the population origin of the individuals was used in these run, and in all runs with STRUCTURE, the "correlated allele frequencies" model was employed.

The runs were repeated separately for each seed lot (in STRUCTURE ver. 2.3.4), this time using 10^6 runs both for burn-in runs and iterations (K = 1 to 20 or K = 2 to 12, 10 runs per K). A similar set of conditions was employed in analysing the whole data set with "population info" set to seed stands and using the "start at population info – locprior" option in STRUCTURE (K = 5 to 15, 10 runs per K). In these cases, both the admixture and no admixture models were used. The most likely number of all these pre-defined K clusters in the various sets of runs was assessed by Pritchard et al.'s (2000) "informal" original suggestion (highest ln(K) with still low variability), and by the more formal criteria of Evanno et al. (2005), as implemented in Earl and von Holdt's (2012) STRUCTURE HARVESTER web service (http://taylor0.biology.ucla.edu/structureHarvester/#).

For assessing similarities versus differences (genetic distances) among entities (seed stands and individual seed lots, respectively), the Q matrices of the STRUCTURE output at the selected K (which give a proportion of ancestry of each actual population in each assumed ancestral cluster) were utilized for calculating simple Euclidian distances between each entity in the analyses (as averages of the ten runs at the same K value).

Reconstruction of maternal genotypes

Lexer et al. (1999) proposed a procedure to infer maternal alleles on the basis of an array of maternal half-sib families (of *Quercus robur*, in their case) with a minimum of eight to ten seeds per tree, using the rules of codominant Mendelian inheritance (Mendel 1866). The same methods are applied to the half-sib families of *Fraxinus excelsior* here. The "rules" how to infer maternal genotypes at a locus, for our specific case, can be described as follows: if a homozygous (single-locus) genotype appears in the seeds, this allele is one of the maternal ones; if there is no such homozygous allele, two alleles are selected so that each of them is present in at least one seed of the lot; if there are more than two homozygous genotypes, the two most frequent alleles among them are selected; if a single allele is present in all seed genotypes, and partially in homozygous condition, no second allele is selected (the maternal genotype is assumed homozygous for the frequent allele); and finally, if a single allele is present in all seed genotypes, but only as a heterozygote, the second most frequent allele is selected as well. An example with two loci is shown in Table 2. However, uncertainties in this inference remain because of technical errors, possible mix-ups of seeds or DNA samples, and stochasticity.

Table 2. Reconstruction of the maternal genotype at an example of 2 loci in population ASL23

Individual	FEMSATL4		FEMSATL11	
23_02_01	192	182	196	190
23_02_02	182	166	196	186
23_02_03	182	166	208	186
23_02_04	182	166	196	186
23_02_05	168	168	196	186
23_02_06	182	168	208	196
23_02_07	240	168	186	178
23_02_08	168	168	196	186
23_02_09	174	168	196	186
23_02_10	202	168	208	186

Offspring individuals of one mother tree (23_02) are arranged in rows, the genotypes at the two loci are listed in columns as allele sizes in base pairs. Maternal alleles as inferred from the offspring are shaded in bright and dark grey

With the inferred data, detection of possibly identical mother trees was done by calculating a mean proportion of shared alleles (POSA) for each seed lot using the pro-

gram MSA (Dieringer and Schlötterer 2003). The same parameter POSA was also calculated for "suspiciously similar" inferred maternal multi-locus genotypes. Because of the uncertainties associated with inference of these genotypes, mother trees belonging to the same seed lot and with POSA 66.7 % and higher were registered as possibly identical.

Results

Basic genetic parameters

Allele size ranges and other basic genetic parameters are listed in Tables 3 and 4, along with data from previous studies in *F. excelsior* which have employed similar sets of microsatellites. On a per locus basis, we observed between seven and 60 alleles. Gene diversity H_E ranged from 0.520 to 0.962 and F_{IS} from 0.165 to 0.608 (not shown). Higher values for H_E and F_{IS} in FEMSATL12 indicate null alleles at this locus, but amplification success was not strikingly different from other markers, and there were not too many homozygous individuals; consequently, it was included in further calculations. Locus FEMSATL16 showed unusually low values in gene diversity and number of alleles, together with high F_{IS}, revealing that this locus is not as polymorphic as the others, with a higher degree of homozygosity.

Other basic genetic parameters on a population basis are given in Table 3. Within populations (seed stage), gene diversity H_E ranged from 0.755 to 0.815, with an excess of homozygotes and a mean inbreeding coefficient (F_{IS} = 0.245) significantly deviating from zero (Table 3). When omitting FEMSATL12 and FEMSATL16 from the calculation of F_{IS}, there was still a highly significant positive mean value (F_{IS} = 0.106 +/- 0.019).

Hierarchical genetic differentiation among seed stands and single tree harvests

Certain alleles strongly differentiated some of the seed lots. This was most pronounced in FEMSATL4 and FEMSATL11, and for seed stand ASL27. In general, differentiation among seed stands was low (F_{ST} = 0.057) but significant (Table 5); *i.e.*, only 5.7 % of the total genetic diversity (H_T = 0.843) arises from among-population (seed stand) differentiation. Among single tree harvests, compared to the total, we calculated a differentiation of 16.9 %. We further assessed the differentiation of single tree harvests within the seed stands, resulting in a slightly lower F_{SP} (between 0.109 and 0.159). Finally, F_{IS} (individual seeds within single tree lots) ranged from -0.116 to 0.514 in the six loci; this means that when collecting seed from a different tree in the stand, or when raising a seedling from a different seed within the same single tree lot (same mother), genetic differentiation is already quite high - 13.3 % and 14.1 % on average, respectively.

The model-based clustering approach implemented in the software STRUCTURE indicated some population structure in the data. The runs of the whole data set with the "no admixture" model (which gave a more meaningful curve of ln(K)), gave "peaks" (according to Pritchard et al.'s 2000, and Evanno et al.'s 2003 criteria) at K = 2 and K = 10, which are close to the actual number of six seed

Table3. Genetic diversity statistics within seed lots and across all seed lots

Population	n	A	H_O	H_E	F_{IS}
ASL02	96	17.8	0.634	0.809	0.234***
ASL23	96	19.7	0.584	0.755	0.235***
ASL27	94	15.3	0.630	0.799	0.227***
ASL50	95	18.7	0.581	0.815	0.294***
ASL65	96	18.0	0.612	0.793	0.237***
ASL98	96	18.3	0.602	0.805	0.257***
mean populations	95.5	18.0 ± 1.45	0.607	0.796	0.245 ± 0.093
all populations	573	33.7	0.603	0.843	0.248***

n, sample size; A, average number of alleles per locus; H_O, average proportion of heterozygotes; H_E, average gene diversity; F_{IS} Wright´s inbreeding coefficient. Deviation from Hardy-Weinberg genotypic proportions: ***P < 0.001

Table 4. Comparison of allele size ranges in various studies using similar sets of microsatellite markers

	Brachet et al. 1999, Lefort et al. 1999		Heuertz et al. 2001		This study		Fussi and Konnert 2014		Gömöry et al. 2012		Ferrazzini et al. 2007	
Locus	A_t	size	A_t	size	A_t	size	A_t	size	A_t	size	A_t	size
M2-30	18	182-248	59	182-294	49	178-282	58	175-301	54	174-290	n.a.	n.a.
FEMSATL4	9	164-228	50	158-251	32	164-243	49	154-254	54	152-268	32	157-205
FEMSATL 11	11	180-226	32	176-266	30	174-238	35	181-255	31	180-242	42	161-234
FEMSATL 12	9	180-262	18	181-264	24	178-265	31	175-229	n.a	n.a.	39	147-261
FEMSATL 16	4	180-200	10	176-204	7	178-203	15	178-208	5	184-200	9	184-214
FEMSATL 19	12	174-214	33	142-229	60	145-243	37	143-223	n.a.	n.a.	55	142-238

A_t, total number of alleles; size of alleles in base pairs; n.a., not analysed

Table 5. Hierarchical F-statistics in the six microsatellite loci

	Overall inbreeding	Differentiation among:			
		seed lots	single trees	single trees within populations	seeds within trees
Loci	F_{IT}	F_{ST}	F_{ET}	F_{SP}	F_{IS}
FEMSATL4	0.256	0.086	0.208	0.159	0.049
FEMSATL11	0.070	0.071	0.158	0.110	-0.116
FEMSATL12	0.680	0.062	0.187	0.149	0.514
FEMSATL16	0.591	0.072	0.192	0.133	0.488
FEMSATL19	0.176	0.034	0.153	0.133	0.024
M2-30	0.173	0.028	0.290	0.109	0.047
Multilocus	0.291	0.057	0.169	0.133	0.141
Significance (*p*)	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Significances of the multilocus estimates are computed by permutation tests

stands. "Secondary peaks" were found at K = 30, and a minor one also at K = 65, the actual number of seed mother trees. A "peak" of K = 29 was most pronounced in the runs with the admixture model. Some structure was visible in all these cases from the STRUCTURE bar plots of Q for individuals, in the sense that individuals belonging to the same entity (stand seed lot or mother tree) showed similar Q proportions (for an example, see Figure 1A). When the resulting colour bars were sorted according to the stand seed lots, they indicated distinct patterns for each stand (Figure 1B-C). Within the stands, however, patterns showed also some differentiation (in Figure 1B, the first bars in ASL02 are probably due to many missing data points for the individual seeds). Using the "no admixture" and "location prior" models, the structure was much more apparent, and mostly conformed with seed stand origin (Figure 1C) and single tree lot, respectively.

In order to illustrate the relatedness expressed in these calculations, we present here pie charts of Q values for the seed stands with progressive values for K assumed in STRUTURE (using "no admixture" and "location prior"; Figure 1D-F): for K = 2, the six studied populations separated into two groups that match geographical patterns (Figure 1D). Seed stands ASL27 and ASL98 (blue pies) are situated in the northern and southern outer chains of the Alps, respectively, whereas ASL02, ASL50 and ASL65 (red pies) are stands along the river Danube. The latter ones formed a second group, also comprising ASL 23 in the region of the easternmost edge of the Alps. For K = 4, the three stands along the Danube still share the same (blue) colour, while all others are differentiated from those, and from each other (Figure 1E). At K = 6, only the two closest stands on the Danube retain some relationship (same colour – purple segments in charts; Figure 1E).

In a more formalized way, Euclidian distances based on the Q proportions of the most likely K, calculated between seed stands, gave a similar picture (Table 6). Stand ASL02 and ASL50 along the Danube came out closest in all runs. ASL50 and ASL65 were still reasonably close. The highest distances were obtained for ASL27 and ASL65 (without "location prior") and ASL65 and ASL98 (using "location prior"). Both latter cases involved the same stand on the Danube and one of the Alpine stands.

In order to better evaluate differentiation within the stand seed lots with the STRUCTURE approach, K values between one and 20 were tested in each lot, and values between K = 3 and K = 6 were most meaningful. Distinct colour patterns were visible for the single tree lots of different mother trees, and calculations of Euclidian distances mostly confirmed these patterns (see also further below), although single seeds sometimes did not completely fit to their lot. In some instances, they even seemed to better fit to other single tree lots in the same stand.

Table 6. Euclidian distances between seed stands, averages based on "Q" proportions of ten runs each at K = 6, from STRUCTURE runs using the "no admixture" model

Seed stands	ASL02	ASL23	ASL27	ASL50	ASL65	ASL98
ASL02	-	0.7379	0.8616	0.4030	0.7827	0.7705
ASL23	1.3548	-	0.9429	0.7314	0.8728	0.8196
ASL27	1.3720	1.3825	-	0.8731	0.9901	0.8948
ASL50	1.0908	1.2653	1.2739	-	0.5323	0.7290
ASL65	1.3639	1.3960	1.3996	1.1239	-	0.8874
ASL98	1.3854	1.3945	1.3861	1.2860	1.4115	-

Above diagonal, distances obtained from STRUCTURE runs without "location prior" information; below diagonal, using "location prior"

Reconstructed maternal multilocus genotypes

In addition, multilocus genotypes were reconstructed for the 65 mother trees. On the basis of 3124 successfully assessed single-locus genotypes in 573 seeds (over all loci), a total of 390 single-locus maternal genotypes (65 trees times six loci) was inferred by the method described in Lexer et al. (1999). The majority of single-locus seed genotypes (2863 out of 3124, 91.65 %) fitted the inferred, most likely maternal genotypes. Regarding single seeds, about one third of the seeds in each lot had at least one incompatible single-locus genotype; however, the cases of seeds having two or three such incompatible genotypes were rare (only five cases had three incompatible single-locus genotypes, which is less than 1 %).

The inferred maternal multi-locus genotypes differed from each other at varying degrees. In three out of six seed lots, the results gave some indications that the actual number of genetically distinct seed mothers was lower than

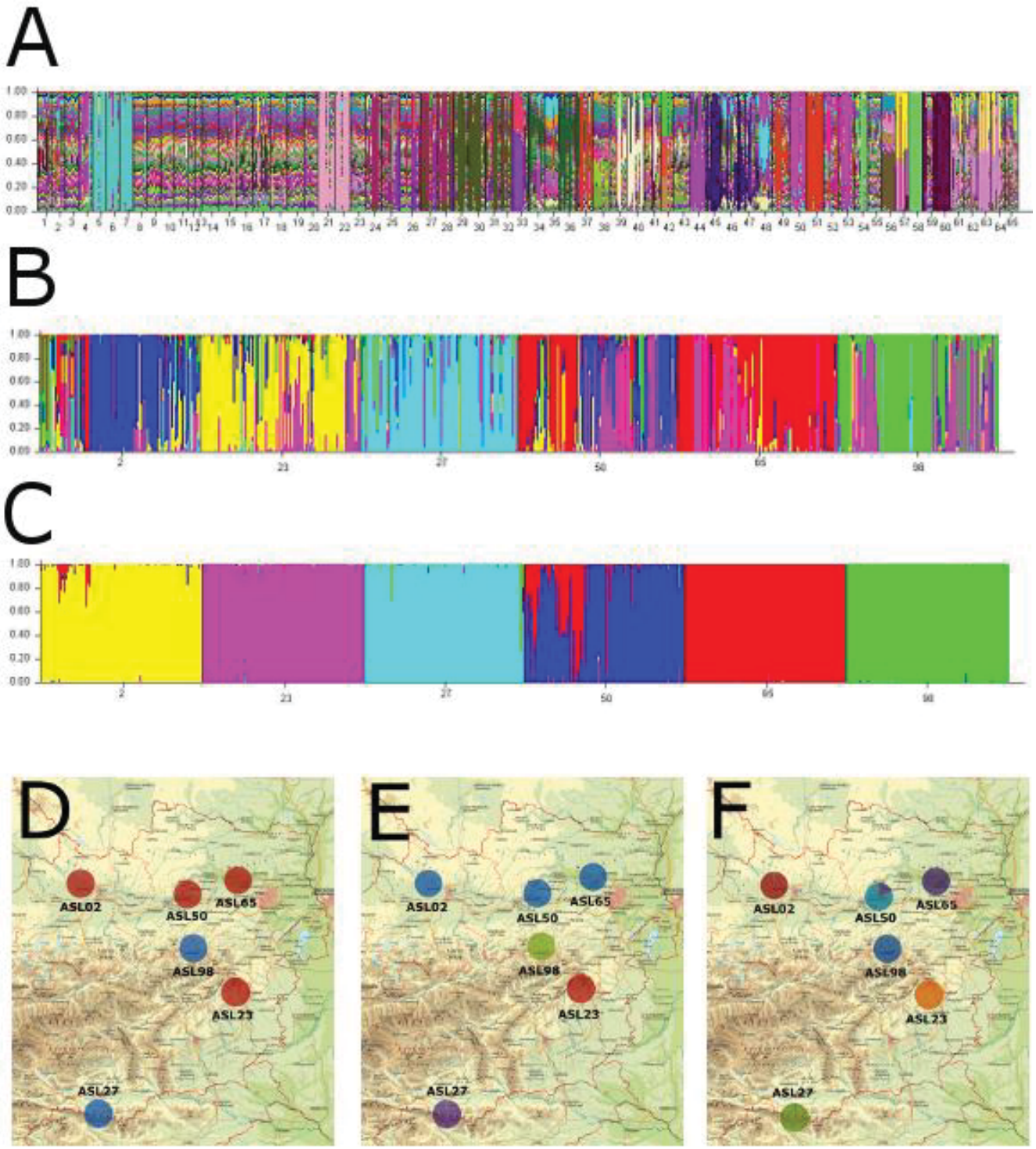

Figure 1. A: Sample bar plot of Q from STRUCTURE runs of the complete data set, at K = 65. Numbers below plot indicate single tree seed lots - ASL02: 1-13; ASL23: 14-23; ASL27: 24-33; ASL50: 34-43; ASL65: 44-53; ASL98: 54-65. Figure 1B-C: Sample bar plots of Q from STRUCTURE runs of the complete data set, sorted according to stand seed lots, at K = 6. Numbers below plot indicate stand seed lots ASL02, ASL23, ASL27, ASL50, ASL65 and ASL98, respectively, at K = 6. 1B, no prior population location info used; 1C, with location prior option (seed stand affiliations as starting points for iterations) Figure 1D-F: Progression of differentiation of stand seed lots at increasing K values assumed in STRUCTURE. 1D, Q proportions in each stand seed lot at K = 2; 1E, at K = 4; 1F, at K = 6

written in the documentation. For example, in stand seed lot ASL23, some inferred maternal multi-loci genotypes were quite similar (Table 7). The degree of allele sharing was higher overall in ASL23 than in the other two seed lots ASL27 and ASL98; 41.3 % versus 26.7-39.4 %, respectively. Given the technical and methodical sources of error, and the greater differentiation among the other reconstructed genotypes, it is fair to assume that the single tree seed lots that arrived in the BFW laboratory may have been derived from identical trees. For the stand seed lot ASL23, this would mean that seed was only harvested possibly from as few as seven individual trees. Additionally, proportions of allele sharing for possibly identical mother trees in the three seed lots ranged from 66.7 to 83.3 %, while other stand seed lots only gave highest POSA at 50.0 to 58.0 %.

Identical mother trees should give fairly identical seed lots (if the seeds are drawn from a well-mixed "bag"). This was tested by comparing Euclidian distances, based on the Q proportions obtained with a likely value for K, among the single tree seed lots. In stand seed lot ASL23, this distance was exceptionally low for single tree lots 6 and 10 (0.01598 +/- 0.003524) which also shared a high proportion of their alleles (Table 7). Lots 8 and 9 had the second lowest distance value (0.05632 +/- 0.009385), but lots 1 and 2, though sharing a lot of alleles in the inferred mothers as well, were much more distant (0.3453 +/- 0.003956). In contrast, in a stand seed lot that did not attract attention for including possibly identical inferred maternal multi-locus genotypes, ASL02, at K = 5, distances between single tree seed lots 5 and 7 were unusually low (0.04233 +/- 0.0245), and this was also the case for trees 9 and 11 (0.07511 +/- 0.01519; while all other 66 pair-wise distances between single tree seed lots were greater than 0.1 in this stand lot). Mother trees 5 and 7 did not share a high proportion of alleles (33.3 %), and 9 and 11 shared only 41.7 %. These values are much lower than for seed stand ASL23, but are among the highest within this seed lot ASL02.

In the two other cases of stand seed lots with possibly identical maternal genotypes, these comparisons appeared as follows: In ASL27, mother trees 1 and 3 shared nine out of 12 alleles, and their Euclidian distance was at 0.1290 +/- 0,01545, the second lowest in this lot. The lowest was between trees 8 and 9 (0.1218 +/- 0.01629), sharing 50 % of the inferred maternal alleles. Another pair, mother trees 6 and 7, shared eight out of 12 alleles (66.6 %). The Euclidian distance between their single tree seed lots was 0.1486 +/- 0.004983). All other distances in this lot were above 0.2. In ASL98, mother trees 10 and 11 shared the highest proportion of alleles (75 %), and also their Euclidian distance was the lowest in the lot (0.1671 +/- 0.003235).

The remaining stand seed lots had single cases of "suspicious" low distances with relatively low allele sharing levels (ASL50: 6 and 8 – 0.1030 +/- 0.02207 – 33.3 % allele sharing; ASL65: 6 and 8 – 0.1064 +/- 0.002228 – 33.3 %, and 7 and 10 – 0.1647 +/- 0.01483 – 50 %).

Discussion

The primers used in this study have a long record of usage (Table 4). Although they are not always "ideal" in a sense that unambiguous allele scoring is easily possible, the data from multiple laboratories (Table 4) show a remarkable synchronicity in allele size ranges, something that is often confounded by the usage of various fluorescent labels, DNA polymerases, and (capillary) electrophoresis apparatus in different laboratories. Sutherland et al. (2010) as well as Beatty et al. (2015) suggested modifications to the primer sequences in order to minimize null allele problems. Although the original primer sequences were used in this study, such null allele problems were not apparent. Sequencing of the *Fraxinus excelsior* genome (see e.g. www.ashgenome.org) will reveal whether the currently used primers may be replaced by more "robust" ones, but sequence data from multiple populations across the range of the species would be necessary for an assessment.

Table 7. Inferred (hypothetical) mother trees of seed lot 23

Individual	FEMSATL4		FEMSATL11		FEMSATL12		FEMSATL16		FEMSATL19		M2-30	
23_1	182	168	196	186	200	184	182	182	201	199	258	212
23_2	182	168	196	186	200	184	182	182	239	201	222	212
23_3	168	166	188	186	200	184	186	182	198	146	218	212
23_4	188	168	200	186	200	184	182	182	198	185	246	198
23_5	206	166	206	196	200	184	182	182	185	183	246	224
23_6	168	166	206	186	200	184	182	182	201	197	220	216
23_7	174	174	200	196	188	184	182	182	195	183	200	216
23_8	184	174	208	196	200	184	182	182	193	183	244	226
23_9	184	174	208	208	200	196	186	182	193	183	226	220
23_10	168	166	206	186	202	200	186	182	219	197	220	216

Identical alleles of possibly identical inferred maternal multi-locus genotypes (mother trees) in identical colour shading

Stand seed lots analysed in this study were within a narrow range of basic genetic parameters (Table 3): only ASL27 (the southernmost, Alpine stand) had somewhat lower allele numbers, but this was not reflected in lower heterozygosities. Although inbreeding coefficients were significantly different from zero in all stands (Table 3), this may either be a remaining effect of a few null alleles present, or of the general tendency of forest trees to tolerate some pollination among relatives in the seed stands. Selection usually removes such effects: adult trees often show the opposite, i.e. heterozygote advantage.

The level of differentiation among seed lots (Table 5) is comparatively high (slightly above 5 %); this is more than could be expected from a typical European forest tree in such a limited area. In studies of ash in regions of comparable sizes, higher values were reported by Heuertz et al. (2001) in Bulgaria (F_{ST} = 0.087), and slightly lower ones for Bavaria (Southern Germany, F_{ST} = 0.046, Fussi and Konnert 2014). Beatty et al. (2015) reported similar low values of differentiation (though they calculated different parameters), in their study of a similar, comparatively small region. In their study, these low values of differentiation were coupled with indications for a strong long-distance gene flow component. Average F_{ST} values were lower in Great Britain (0.025; Sutherland et al. 2010), on an even larger geographical scale than in our study region. It appears that the stands along the Danube river in our study are closer to each other genetically than the rest (Table 6, Figure 1D). There is also some possible relatedness among the Alpine stands (ASL27 and ASL98), although these are quite distant geographically. Moreover, ash distribution in the Alps is more scattered, especially in the inner Alps (Heinze et al. 2017). The remaining genetic connectivity of these two stands, and their differentiation from "floodplain" stands (along the Danube), may hint at a general genetic "divide" among ecotypes in mountain valleys and the lowlands in Austria, but this remains to be investigated in much greater detail. Influences of hybridisation (introgression) with *F. angustifolia*, a sister and neighbour species in the Pannonian basin, could also play a role here. Fussi and Konnert (2014) included two stands from Austria in their analysis of material from Bavaria, one of which was situated very nearby ASL23, and the other somewhat upstream of ASL02. The Austrian stand near ASL23 in their study showed relatively high differentiation from the rest of the populations. As there was a large sampling gap in between, this would hint at typical clinal variation patterns of this species in this wider region.

What is remarkable are the high differentiation values for the single-tree seed lots, within their stands as well as compared to the total (Table 5). This far exceeds usual values obtained in comparison of trees among stands. One possible explanation could be pollen clouds of restricted diversity that fertilize single female flowers. The seeds for our study may have been picked from single clusters of seeds during harvesting.

It could be imagined that colonisation events with seed from just one or a few seed parents may lead to high differentiation in this species. However, ash seed is long-lived. If such colonisation events are repeated over various years, from various parents, differentiation effects may decrease, and this is what we apparently see in this species at the adult stage. For artificial regeneration with nursery-grown plant material, however, it follows that using seed from just a handful of parents may lead to unwanted differentiation effects. Such plants may not represent the entire genetic make-up of the stand they derived from.

Have the persons who harvested the seed for this study taken account of the detrimental effects of limiting the level of genetic diversity? When seeds are harvested by climbers, costs increase significantly with every tree climbed. Also when picking seed from felled trees, it takes much longer to collect a similar amount of seed from many trees, as opposed to just from a few. In Austria, forest seed harvesting regulations require a minimum number of trees to be used, in order to avoid effects of genetic bottlenecks, as outlined above. The data obtained in this study allow an assessment of how this measure was accepted in practice. We have some hints that may indicate issues in this respect. The "suspiciously" similar single tree seed lots could actually stem from identical trees. They showed only a small amount of differentiation among them. This should not be the case for well-mixed seed from the same tree, but seeds picked from different parts of the crown may have been derived from slightly different pollen clouds. We have analysed bigger numbers of seed from a single tree (96 seeds; Heinze et al. 2017 and unpublished data), but in this example, seeds from the whole crown were mixed. A more structured sampling in different parts of a crown may reveal if there is genetic differentiation of pollen clouds in different parts of the crown.

Another possibility for explaining only slightly differentiated seed lots would be clonal trees, e.g. trees grown from sprouts of an identical root system. Ash often grows from woodstocks, but is not known to form large clones of adult trees. It is not so likely that even only two large, seed-bearing trees develop from a single woodstock. Genetically closely related mother trees, e.g. full-sib mothers, may be another explanation for the results obtained.

An alternative method investigated here was the reconstruction of the genotypes of the mother trees. The procedure used has some drawbacks. Technical errors in allele sizing, single mixed-in seeds from other trees, and simple stochasticity effects may blur the determination of the true maternal alleles. It is possible for several tree species to analyse maternal tissue, e.g. seed coats (Ziegenhagen et al. 2003), though this "dead" tissue is more challenging to analyse. The re-constructed maternal genotypes in our

study show reasonable agreement with genetic distances of "questionable" seed lots, so they may come close to the true alleles. Analysing somewhat larger half-sib families (single tree seed lots) may also alleviate the problem. In any case, the methods employed here provide a first possibility of investigating conformity with legal requirements for seed harvests in Austria, and confirm that the suggestions originally made by Lexer et al. (1999) and Heinze and Lexer (2000) are a practicable approach.

The generally great differentiation among single-tree seed lots was also evident from the STURCTURE runs (and the derived Euclidian distances among them). This raises the question of how large a sample of seeds or plants must be in order to "truly" represent the genetic make-up of the stand. This question was answered empirically when the seed regulations were drawn up in Austria. It was decided at the time to demand harvesting from at least ten mother trees for this and other "scattered" hardwood species. The German law for forest reproductive material requires even twenty trees for harvesting seeds in ash. In contrast, in several other European countries no similar regulations for seed harvesting in ash are in place. The data presented here provide evidence that the choice of a minimum of ten trees was reasonable in ash, as the differentiation of whole seed stand lots was in the range of previous genetic studies that have analysed adult trees from different populations, while single tree seed lots showed very large differentiation values. This means that stand seed lots consisting of only few single mother tree seed lots would be atypical for the whole genetic diversity in the stand. This would lead to the effect that seed lots from different years (and different mother trees in each year) would not resemble each other much. Nurseries that would grow plants from such material would have to deal with quite variable traits, at least as far as genetic markers are concerned, but possibly also in growth traits. The reconstructed maternal genotypes were also quite variable (an attempt at inferring genetic clusters among the 65 reconstructed mother trees with STRUCTURE did not show any; data not shown). It should be investigated whether the minimum number of ten harvested trees also holds in similar "scattered hardwood" species like wild cherry (*Prunus avium*) or sycamore maple (*Acer pseudoplatanus*), which are insect-pollinated.

Sustaining a high level of genetic variability will be especially challenging in the wake of the progressing ash dieback disease. Less and less ash stands are utilized for seed harvesting since the onset of the disease (Heinze et al. 2017). If the same stands are utilized in different years, this may lead to a still narrower genetic basis of nursery plants. Contrary to that, it may be desirable to harvest from even more stands, as there may be different levels of disease tolerance in different stands. It is questionable whether this is still feasible, as the disease seems to progress steadily (Heinze et al. 2017). Disease-damaged trees are also dangerous for climbing.

Mixing plants for re-forestation from lots within the same region of provenance (and the same altitudinal zone) may be much desirable. Stored seed may help to overcome any developing bottlenecks in seed supply.

There is still a need to investigate whether trees with varying degrees of disease symptoms pass on any such "tolerance" to their offspring (McKinney et al. 2012, Pliura et al. 2011, Fussi and Konnert 2014), and whether it is durable. Pliura et al. (2014) reported that over time, disease incidence in all trees investigated increased, and that few remained relatively healthy (under 50 % damage) in later observations, but that heritabilities increased with time. Pliura et al. (2016), in a different set of experiments, also stated that none of their (half-sib) families stayed completely symptom-free. It may be concluded from such data that if there is any resistance, it is rather partial or quantitative, and that time series observations over several years, of individual trees, are desirable. "Survival rates" over several years may be better indicators of increased levels of disease tolerance than single "degree of damage" assessments in single years (see also Heinze et al. 2017).

We are currently investigating correlations of health levels between adult trees and their seedlings in forest situations, using microsatellites for establishing the parent-offspring relationships (A. Wohlmuth and B. Heinze, manuscript submitted).

Acknowledgement

Research in the BFW lab was financed by the EU-funded research project "RAP – Realising Ash' Potential" – contract QLK5-CT-2001-00631 of the EC's Fifth Framework Programme. Further funding from COST Action "Fraxinus dieback in Europe: elaborating guidelines and strategies for sustainable management - FRAXBACK" – FP1103 for meeting attendance is gratefully acknowledged. Prof. Dusan Gömöry (Technical University, Zvolen) is ackowledged for sharing data. Irene Nanista, Agnieszka Koziel, and Ingrid Gerstl assisted in the laboratory during parts of these investigations.

References

Bacles, C.F.E., Lowe, A.J. and Ennos, R.A. 2006. Effective seed dispersal across a fragmented landscape. *Science* 311(5761): 628.

Bacles, C.F.E. and Ennos, R.A. 2008. Paternity analysis of pollen-mediated gene flow for *Fraxinus excelsior* L. in a chronically fragmented landscape. *Heredity* 101(4): 368-380.

Beatty, G.E., Brown, J., Cassidy, E., Finlay, C.V., McKendrick, L., Montgomery, W.I., Reid, N., Tosh, D. and Provan, J. 2015. Lack of genetic structure and evidence for long-distance dispersal in ash (*Fraxinus excelsi-*

or) populations under threat from an emergent fungal pathogen: implications for restorative planting. *Tree Genetics & Genomes* 11(3): 1-13.

Brachet, S., Jubier, M.F., Richard, M., Jung-Muller, B. and Frascaria-Lacoste, N. 1999. Rapid identification of microsatellite loci using 5' anchored PCR in the common ash *Fraxinus excelsior*. *Molecular Ecology* 8(1): 160-163.

Dieringer, D. and Schlötterer, C. 2000. microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 3(1): 167-169.

Earl, D. and von Holdt, B. 2012 . STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4(2): 359-361.

Evanno, G., Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14(8): 2611-2620.

Falush, D., Stephens, M. and Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164(4): 1567-1587.

Falush, D., Stephens, M. and Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* 7(4): 574-578.

Ferrazzini, D., Monteleone, I. and Belletti, P. 2007. Genetic variability and divergence among Italian populations of common ash (*Fraxinus excelsior* L.). *Annals of Forest Science* 64(2): 159-168.

Fussi, B. and Konnert, M. 2014. Genetic analysis of European common ash (*Fraxinus excelsior* L.) populations affected by ash dieback. *Silvae Genetica* 63(5): 198-212.

Gerard, P.R., Fernandez-Manjarres, J.F. and Frascaria-Lacoste, N. 2006. Temporal cline in a hybrid zone population between *Fraxinus excelsior* L. and *Fraxinus angustifolia* Vahl. *Molecular Ecology* 15(12): 3655-3667.

Goudet, J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* 86:485-486.

Gömöry, D., Paule, L., Krajmerova, D., Romsakova, I. and Piecka, J. 2012. Gene exchange across a postglacial contact zone in *Fraxinus excelsior* L. *Silvae Genetica* 61(1-2): 18-27.

Gérard, P.R., Temunovic, M., Sannier, J., Bertolino, P., Dufour, J., Frascaria-Lacoste, N. and Fernández-Manjarrés, J.F. 2013. Chilled but not frosty: understanding the role of climate in the hybridization between the Mediterranean *Fraxinus angustifolia* Vahl and the temperate *Fraxinus excelsior* L. (Oleaceae) ash trees. *Journal of Biogeography* 40: 835-846.

Heinze, B. and Lexer, C. 2000. Oak seedlots and forest seed trade regulations: an investigation using molecular markers. *Glasnik Za Šumske Pokuse - Annales Experimentis Silvarum Culturae Provehensis (Universitas Studiorum Zagrebensis Facultas Forestalis)* 37: 347-360.

Heinze, B., Koziel-Monte, A. and Jahn, D. 2014. Analysis of variation in chloroplast DNA sequences In: Besse, Pascale, ed.: Molecular Plant Taxonomy. Humana Press, New York, pp. 85-120.

Heinze, B., Tiefenbacher, H., Litschauer R. and Kirisits T. 2017. Ash dieback in Austria - history, current situation and outlook. In: Vasaitis, R. and Enderle, R. (Editors), Dieback of European Ash (*Fraxinus* spp.): Consequences and Guidelines for Sustainable Management. Uppsala, Sweden, Swedish University of Agricultural Sciences, p. 33-52. https://www.slu.se/globalassets/ew/org/inst/mykopat/forskning/stenlid/dieback-of-european-ash.pdf

Heuertz, M., Hausman, J.-F., Tsvetkov, I., Frascaria-Lacoste, N. and Vekemans, X. 2001. Assessment of genetic structure within and among Bulgarian populations of the common ash (*Fraxinus excelsior* L.). *Molecular Ecology* 10(7): 1615-1623.

Heuertz, M., Vekemans, X., Hausman, J.-F., Palada, M. and Hardy, O.J. 2003. Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology* 12(9): 2483-2495.

Heuertz, M., Hausman, J.-F., Hardy, O.J., Vendramin, G.G., Frascaria-Lacoste, N. and Vekemans, X. 2004. Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). *Evolution* 58(5): 976-988.

Heuertz, M., Carnevale, S., Fineschi, S., Sebastiani, F., Hausman, J.F., Paule, L. and Vendramin, G.G. 2006. Chloroplast DNA phylogeography of European ashes, *Fraxinus* sp. (Oleaceae): roles of hybridization and life history traits. *Molecular Ecology* 15(8): 2131-2140.

Lefort, F., Brachet, S., Frascaria-Lacoste, N., Edwards, K.J. and Douglas, G.C. 1999. Identification and characterization of microsatellite loci in ash (*Fraxinus excelsior* L.) and their conservation in the olive family (Oleaceae). *Molecular Ecology* 8(6): 1088-1090.

Lexer, C., Heinze, B., Steinkellner, H., Kampfer, S., Ziegenhagen, B. and Glossl, J. 1999. Microsatellite analysis of maternal half-sib families of *Quercus robur*, pedunculate oak: detection of seed contaminations and inference of the seed parents from the offspring. *Theoretical and Applied Genetics* 99(1-2): 185-191.

Lexer, C., Heinze, B., Alia, R. and Rieseberg, L.H. 2004. Hybrid zones as a tool for identifying adaptive genetic variation in outbreeding forest trees: lessons from wild annual sunflowers (*Helianthus* spp.). *Forest Ecology and Management* 197 (1-3): 49-64.

McKinney, L. V., Thomsen, I.M., Kjaer, E.D. and Nielsen, L.R. 2012. Genetic resistance to *Hymenoscyphus pseudoalbidus* limits fungal growth and symptom occurrence in *Fraxinus excelsior*. *Forest Pathology* 42(1): 69-74.

Mendel, G. 1866. Versuche über Pflanzen-Hybriden [Experiments on plant hybrids]. *Verhandlungen des naturforschenden Vereins in Brünn [Transactions of the Natural History Society in Brno]* 4: 1-47.

Morand, M.-E., Brachet, S., Rossignol, P., Dufour, J. and Frascaria-Lacoste, N. 2002. A generalized heterozygote deficiency assessed with microsatellites in French common ash populations. *Molecular Ecology* 11(3): 377-385.

Morand-Prieur, M. E., Vedel, F., Raquin, C., Brachet, S., Sihachakr, D. and Frascaria-Lacoste, N. 2002. Maternal inheritance of a chloroplast microsatellite marker in controlled hybrids between *Fraxinus excelsior* and *Fraxinus angustifolia*. *Molecular Ecology* 11(3): 613-617.

Morand-Prieur, M.-E., Raquin, C., Shykoff, J.A. and Frascaria-Lacoste, N. 2003. Males outcompete hermaphrodites for seed siring success in controlled crosses in the polygamous *Fraxinus excelsior* (Oleaceae). *American Journal of Botany*. 90(6): 949-953.

Nei, M 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, 512 pp.

Pliura, A., Lygis, V., Suchockas, V. and Bartkevicius, E. 2011. Performance of twenty four European *Fraxinus excelsior* populations in three Lithuanian progeny trials with a special emphasis on resistance to *Chalara fraxinea*. *Baltic Forestry* 17(1): 17–34.

Pliura, A., Marciulyniene, D., Bakys, R. and Suchockas, V. 2014. Dynamics of genetic resistance to *Hymenoscyphus pseudoalbidus* in juvenile *Fraxinus excelsior* clones. *Baltic Forestry* 20(1): 10-27.

Pliura, A., Lygis, V., Marciulyniene, D., Suchockas, V. and Bakys, R. 2016. Genetic variation of *Fraxinus excelsior* half-sib families in response to ash dieback disease following simulated spring frost and summer drought treatments. *iForest - Biogeosciences and Forestry* 9(1): 12-22.

Pritchard, J. K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155(2): 945-959.

Sutherland, B. G., Belaj, A., Nier, S., Cottrell, J.E., P Vaughan, S., Hubert, J. and Russell, K. 2010. Molecular biodiversity and population structure in common ash (*Fraxinus excelsior* L.) in Britain: implications for conservation. *Molecular Ecology* 19(11): 2196-2211.

Thomasset, M., Fernandez-Manjarres, J.F., Douglas, G.C., Frascaria-Lacoste, N., Raquin, C. and Hodkinson, T.R. 2011. Molecular and morphological characterization of reciprocal F1 hybrid ash (*Fraxinus excelsior* x *Fraxinus angustifolia*, Oleaceae) and parental species reveals asymmetric character inheritance. *International Journal of Plant Sciences* 172(3): 423-433.

Thomasset, M., Fernandez-Manjarres, J.F., Douglas, G.C., Bertolino, P., Frascaria-Lacoste, N. and Hodkinson, T.R. 2013. Assignment testing reveals multiple introduced source populations including potential ash hybrids (*Fraxinus excelsior* x *F. angustifolia*) in Ireland. *European Journal of Forest Research* 132(2): 195-209.

Thomasset, M, Hodkinson, T.R., Restoux, G., Frascaria-Lacoste, N. , Douglas, G.C. and Fernandez-Manjarres, J.F. 2014. Thank you for not flowering: conservation genetics and gene flow analysis of native and non-native populations of *Fraxinus* (Oleaceae) in Ireland. *Heredity* 112(6): 596-606.

Weiser, F 1995. Beitrag zur Existenz von Ökotypen bei gemeiner Esche (*Fraxinus excelsior* L.) [Contribution to the existence of ecotypes in common ash]. *Forstarchiv* 66:251-257.

Ziegenhagen, B., Liepelt, S., Kuhlenkamp, V. and Fladung, M. 2003. Molecular identification of individual oak and fir trees from maternal tissues of their fruits or seeds. *Trees-Structure and Function* 17(4): 345-350.