INBO.R.2013.12

Wetenschappelijke instelling
van de Vlaamse overheid

AFLP fingerprinting of elms

Investigation of clonality and self-pollination of *Ulmus* spp.

Karen Cox and An Vanden Broeck

inbo
Instituut voor
Natuur- en Bosonderzoek

Auteurs:
Karen Cox and An Vanden Broeck
Instituut voor Natuur- en Bosonderzoek

Het Instituut voor Natuur- en Bosonderzoek (INBO) is het Vlaams onderzoeks- en kenniscentrum voor natuur en het duurzame beheer en gebruik ervan. Het INBO verricht onderzoek en levert kennis aan al wie het beleid voorbereidt, uitvoert of erin geïnteresseerd is.

Vestiging:
INBO Geraardsbergen
Gaverstraat 4, B-9500 Geraardsbergen
www.inbo.be

e-mail:
karen.cox@inbo.be

Wijze van citeren:
Cox K., Vanden Broeck A. (2013). AFLP fingerprinting of elms. Investigation of clonality and self-pollination of *Ulmus* spp. Rapporten van het Instituut voor Natuur- en Bosonderzoek 2013 (12). Instituut voor Natuur- en Bosonderzoek, Brussel.

D/2013/3241/055
INBO.R.2013.12
ISSN: 1782-9054

Verantwoordelijke uitgever:
Jurgen Tack

Druk:
Managementondersteunende Diensten van de Vlaamse overheid

Foto cover:
Samaras of *Ulmus laevis* (Kristine Vander Mijnsbrugge)

AFLP fingerprinting of elms

Investigation of clonality and self-pollination of *Ulmus* spp.

K. Cox & A. Vanden Broeck

D/2013/3241/055
INBO.R.2013.12

Acknowledgements

We would like to thank Eric Collin (Irstea, France) for providing valuable information. Also, we are grateful to our lab staff, Sabrina Neyrinck, Leen Verschaeve, An Van Breusegem, David Halfmaerten and Nancy Van Liefferinge, for providing the necessary AFLP fingerprints.

Index

List of figures and tables

1. Introduction

This study entails a few specific questions. Each question was of a different nature, concerning the species *Ulmus laevis* Pall., *U. minor* Mill. and its hybrid with *U. pumila* L., and covering the investigation of clonality and selfing using genetic markers. Genetic fingerprints of several French samples were also provided for the purpose of marketing clones/cultivars as forest reproductive material.

2. Material and methods

1.1. Samples and DNA extraction

Samples were delivered by Eric Collin (Irstea, EFNO - Ecosystèmes forestiers, Nogent-sur-Vernisson, France) in 2007 (Table 1). Total DNA was extracted from ground leaf samples that were stored on silica gel, with DNeasy Plant Mini Kit (Qiagen) on 20 mg of dried leaf tissue. The integrity and quantity of DNA were assessed on 1.5% agarose gels and spectrophotometrically with the ND-1000 Nano-Drop (NanoDrop Technologies), respectively. Table 1 lists the different questions and the samples involved. Each question is then answered in the following paragraphs.

Table 1: List of samples. If not otherwise indicated, the species of the individuals is *U. minor*.

Purpose	**Sample**	**Remark**
Paternity analysis (selfing or outcrossing?)	52,01	Offspring of CEM052
	52,02	Offspring of CEM052
	52,04	Offspring of CEM052
	52,05	Offspring of CEM052
	52,06	Offspring of CEM052
	52,07	Offspring of CEM052
	140,02	Offspring of CEM140
	140,03	Offspring of CEM140
	140,05	Offspring of CEM140
	140,06	Offspring of CEM140
	140,07	Offspring of CEM140
	140,08	Offspring of CEM140
	307	Offspring of CEM052
	315	Offspring of CEM052
	320	Offspring of CEM052
	403	Offspring of CEM052
	413	Offspring of CEM052
	601	Offspring of CEM140
	619	Offspring of CEM140
	809	Offspring of CEM052
	816	Offspring of CEM052
	906	Offspring of CEM140
	907	Offspring of CEM140
	920	Offspring of CEM140
	CEM052	Mother tree
	CEM140	Mother tree
Verification of ramet-to-ortet or ramet-to-ramet relationship	61120h	Ortet 1
	FRA.US.0041	Ramet 1
	61120i	Ortet 2
	FRA.US.0042	Ramet 2
	61160b	Ortet 3

Purpose	Sample	Remark
	FRA.US.0044	Ramet 3
	61160j	Ortet 4
	FRA.US.0048	Ramet 4
	61160k	Ortet 5
	FRA.US.0049	Ramet 5
	61200d	Ortet 6
	FRA.US.0053	Ramet 6
	61500a	Ortet 7
	FRA.US.0072	Ramet 7
	61500A	Ramet 7
	FRA.US.0043	Ramet 8
	61150A	Ramet 8
	CEM052	Ramet 9
	FRA.US.0052	Ramet 9
	CEM140	Ramet 10
	FRA.US.0140	Ramet 10
	CEM186	Ortet? 11
	FRA.US.0186	Ramet 11 (compare DNA source: CEM186)
	CEM188	Ortet? 12
	FRA.US.0188	Ramet 12 (compare DNA source: CEM188)
	CEM193	Ortet? 13
	FRA.US.0193	Ramet 13 (compare DNA source: CEM193)
Identification of clones	Aunay 14260 A	Neighbour of Aunay 14260 B
	Aunay 14260 B	Neighbour of Aunay 14260 A
	Le Vey Q	Neighbour of Le Vey N and in same village as J, U and R
	Le Vey N	Neighbour of Le Vey Q and in same village as J, U and R
	Le Vey J	In the same village as Q, N, U and R
	Le Vey U	In the same village as Q, N, J and R
	Le Vey R	In the same village as Q, N, J and U
	TUS.008	Neighbour of TUS.009
	TUS.009	Neighbour of TUS.008
	TUS.027	Neighbour of TUS.028/29
	TUS.028	Neighbour of TUS.027/29
	TUS.029	Neighbour of TUS.027/28
Species identification	CEM350	*Ulmus minor* x *U. pumila*?
	Blismes	*Ulmus minor*?
Deliver clone ID (for marketing of cultivars)	61100XA	
	FRA.US.0032	
	FRA.US.0041	
	FRA.US.0043	
	FRA.US.0072	
	FRA.US.0077	
	FRA.US.0083	
	FRA.US.0115	
	FRA.US.0205	
	FRA.US.0250	
	FRA.US.0351	

1.2. AFLP analysis

The two samples of *U. laevis* were analysed together with the samples of the same species mentioned in the report of Cox et al. (2012) in chapter one. The *U. minor* samples were

included in the set of samples of the *U. minor- U. glabra* complex and several other species, including *U. pumila*, mentioned in chapter two of the same report. Amplified Fragment Length Polymorphism (AFLP) fingerprints were generated according to Vos et al. (1995), but with restriction and ligation conducted in one single step. We used the following three primer combinations for *U. laevis*: EcoRI-ACA(fam)/MseI-CAC, EcoRI-ACC(ned)/MseI-CAC and EcoRI-ACC(ned)/MseI-CTG. For the second set including *U. minor* and *U. pumila*, two primer combinations were selected: EcoRI-AGC(ned)/MseI-CTG (PC1) and EcoRI-ACC(ned)/MseI-CTG (PC2). AFLP fragments were separated by electrophoresis on an ABI 3500 Genetic Analyzer (Applied Biosystems). The electropherograms were visualized with 3500 Data Collection Software v 1.0 and GeneMapper v 4.1 (Applied Biosystems). The latter was used to adjust the analysis method of the electropherograms and produce an inputfile for RawGeno v 2.0 R CRAN package of Arrigo et al. (2009) for automated scoring. Rawgeno also checks for potential homoplasy, by assessing the correlation between AFLP band size and frequency among samples as recommended by Vekemans et al. (2002). 191 and 394 polymorphic loci were retained for the total *U. laevis* and the total *U. minor-U. glabra* dataset, respectively. Also, 14 and 17% of the samples were blindly replicated, respectively.

Table 2 shows samples of *U. minor* with poor quality AFLP profiles. These samples could therefore not be included in the data analysis.

Table 2: Samples of *Ulmus minor* with poor quality AFLP profiles.

Sample	PC1	PC2
61120h	x	x
413	x	
920	x	

3. Data analysis and results

3.1. Paternity analysis

The question of interest concerning the offspring of the mother trees CEM052 and CEM140 was if the seedlings were obtained through self-pollination or outcrossing. Parentage assignment was performed using Colony v2.0.1.9 (Jones and Wang 2009). The program uses a maximum likelihood method based on the individual multilocus genotypes (MLG), in this case, at 394 dominant marker loci. It can also incorporate scoring errors. We selected the full likelihood approach which is considered the most accurate (Wang 2004, Wang and Santure 2009) and selected the inbreeding model.

Because we do not know all the genotypes that are planted in the seed orchard and consequently do not know all of the potential fathers, we included the following French multilocus genotypes: UPM111, TUS.029, TUS.028, TUS.027, TUS.009, TUS.008, FRA.US.0351, FRA.US.0250, FRA.US.0205, FRA.US.0188, FRA.US.0186, FRA.US.0140, FRA.US.0115, FRA.US.0083, FRA.US.0077, FRA.US.0072, FRA.US.0052, FRA.US.0049, FRA.US.0048, FRA.US.0044, FRA.US.0043, FRA.US.0042, FRA.US.0041, FRA.US.0032, CEM386, CEM350, CEM340, CEM339, CEM330, CEM328, CEM280, CEM276, CEM196, CEM190, CEM188, CEM186, CEM140, CEM052 and CEM275. All these individuals are of the same species, *U. minor*, except the first (UPM111), which was identified as an *U. pumila*. More information on these samples can be found in Cox et al. (2012).

As a result, Colony assigned four different fathers to the offspring, none of which were included in the list of candidate fathers (Table 3, Fig. 1). Consequently, selfing seems unlikely.

Table 3: The best configuration with the maximum likelihood of sibship structures for the *Ulmus minor* samples, obtained with Colony v2.0.1.9. For each individual (Offspring) the known mother ID and the potential father ID is given. Because the inferred fathers are not among the candidates, they are given an index preceded by *.

Sample ID	Father ID	Mother ID
307	#1	CEM052
315	#1	CEM052
320	#1	CEM052
403	#1	CEM052
413	#2	CEM052
809	#2	CEM052
816	#2	CEM052
52-01	#4	CEM052
52-02	#2	CEM052
52-04	#4	CEM052
52-05	#4	CEM052
52-06	#4	CEM052
52-07	#4	CEM052
601	#3	CEM140
619	#3	CEM140
906	#3	CEM140
907	#3	CEM140
920	#3	CEM140
140-02	#4	CEM140
140-03	#4	CEM140
140-05	#4	CEM140
140-06	#3	CEM140
140-07	#3	CEM140
140-08	#3	CEM140

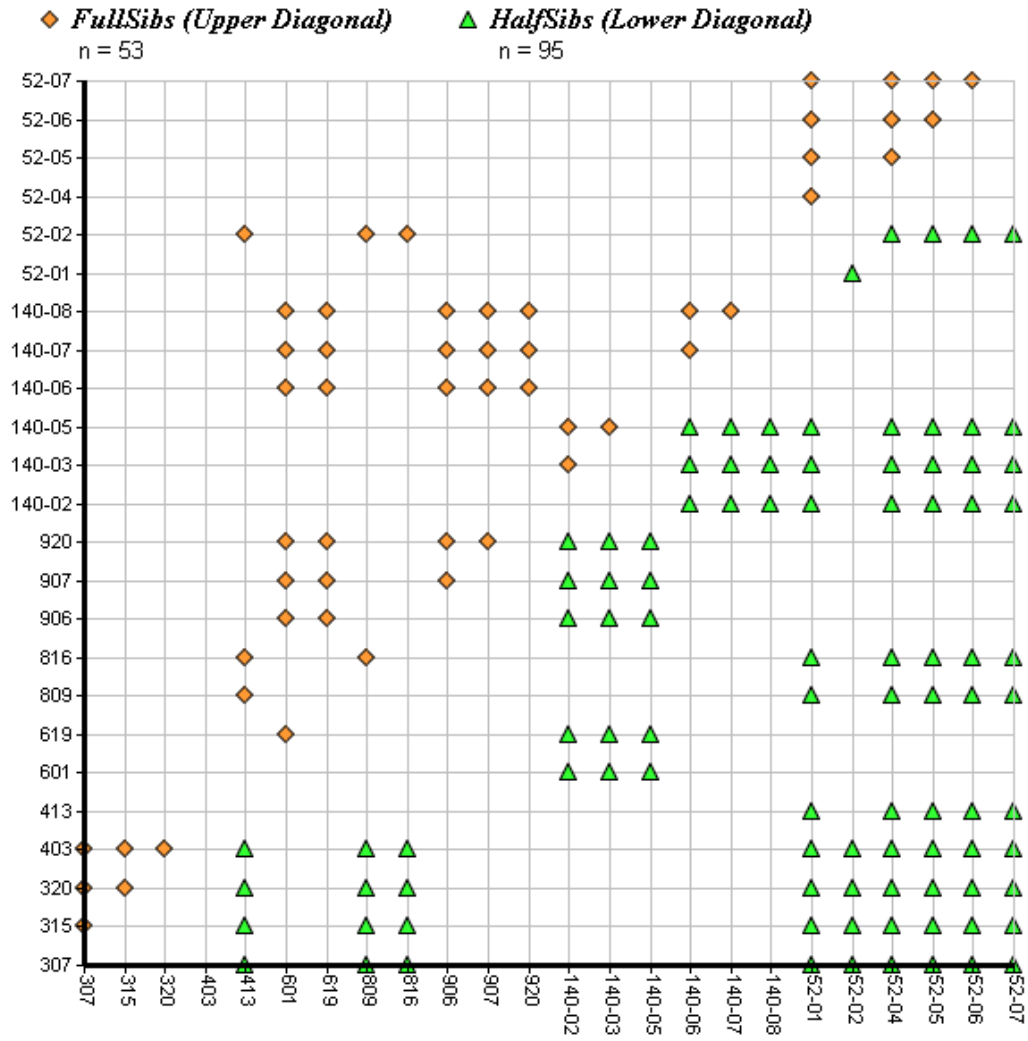

Figure 1: Best Maximum Likelihood Sibship Assignment plot of the sibship structure for the *U. minor* samples, obtained with Colony v2.0.1.9.

3.2. Verification of ramet-to-ortet or ramet-to-ramet relationship

Because DNA-markers are prone to scoring and sometimes technical errors, it becomes difficult to identify clones among a set of samples. In addition, mutations are possible.

To infer clonal identity we used Genotype (Meirmans and Van Tienderen 2004) with a threshold of 0.94 Dice similarity, which is equal to the mean Dice similarity of 0.94 calculated for the duplicate samples of the entire dataset (Cox et al. 2012). CEM193 and FRA.US.0193 are *U. laevis* and are therefore included in the *U. laevis* dataset. For this dataset, a threshold of 3 differences using the infinite allele model was used (Cox et al. 2012). 61120h and FRA.US.0041 could not be compared with each other because of the poor quality of the AFLP profiles of 61120h for both the primer combinations.

None of ortet-ramet or ramet-ramet pairs has a similarity of 0.94 (Table 4). Although mutations could have occurred, the high number of differences between comparisons is

unexpected. We checked the quality of the AFLP profiles again and compared them visually of each ramet-ramet or ramet-ortet pair. Some profiles, for one or both primer combinations, showed lower quality compared to the other sample(s). Because of the heterogeneity in quality of the AFLP profiles across the whole dataset, it is difficult to obtain a reliable threshold that can be used on all samples. Creating a subset of these samples might result in another error rate and therefore another, more reliable threshold. However, there are not enough replicates among these samples to do this. Nonetheless, there are samples, supposedly ramets, which seem very different from each other, which might be caused by mislabelling in the field or in the lab.

Table 4: Dice similarity and number of markers different between ramets or between ortet and ramet.

ortet	ramet	Dice similarity	number of different markers
61120i	FRA.US.0042	0.93	17
61160b	FRA.US.0044	0.91	19
61160j	FRA.US.0048	0.78	45
61160k	FRA.US.0049	0.87	30
61200d	FRA.US.0053	0.91	23
61500a	FRA.US.0072	0.81	43
61500a	61500A	0.83	37
CEM186*	FRA.US.0186	0.84	39
CEM188*	FRA.US.0188	0.93	17
CEM193*	FRA.US.0193	0.97**	7**
ramet	ramet	Dice similarity	number of different markers
FRA.US.0072	61500A	0.81	43
FRA.US.0043	61150A	0.84	37
CEM052	FRA.US.0052	0.92	16
CEM140	FRA.US.0140	0.93	16

*: uncertain if this is an ortet or a ramet.

**: based on the *U. laevis* dataset, i.e. other markers and different number of markers.

Additionally, a Neighbour Joining tree was constructed using Treecon v1.3b (Van de Peer and De Wachter 1994) based on genetic distances according to Nei and Li (1979) (100 bootstraps). The tree was rooted with sample UPM111 which is an *U. pumila*. Other French samples were included as well as duplicate samples. Only bootstrap values above 50 are shown in Fig. 2. The following pairs do not seem to cluster together: 61160j vs. FRA.US.0048, 61500a vs. FRA.US.0072 vs. 61500A and 61150A vs. FRA.US.0043. Although other ortet/ramet-ramet pairs showed a lower Dice similarity, they are clearly related in the NJ tree with bootstrap values ≥ 78%.

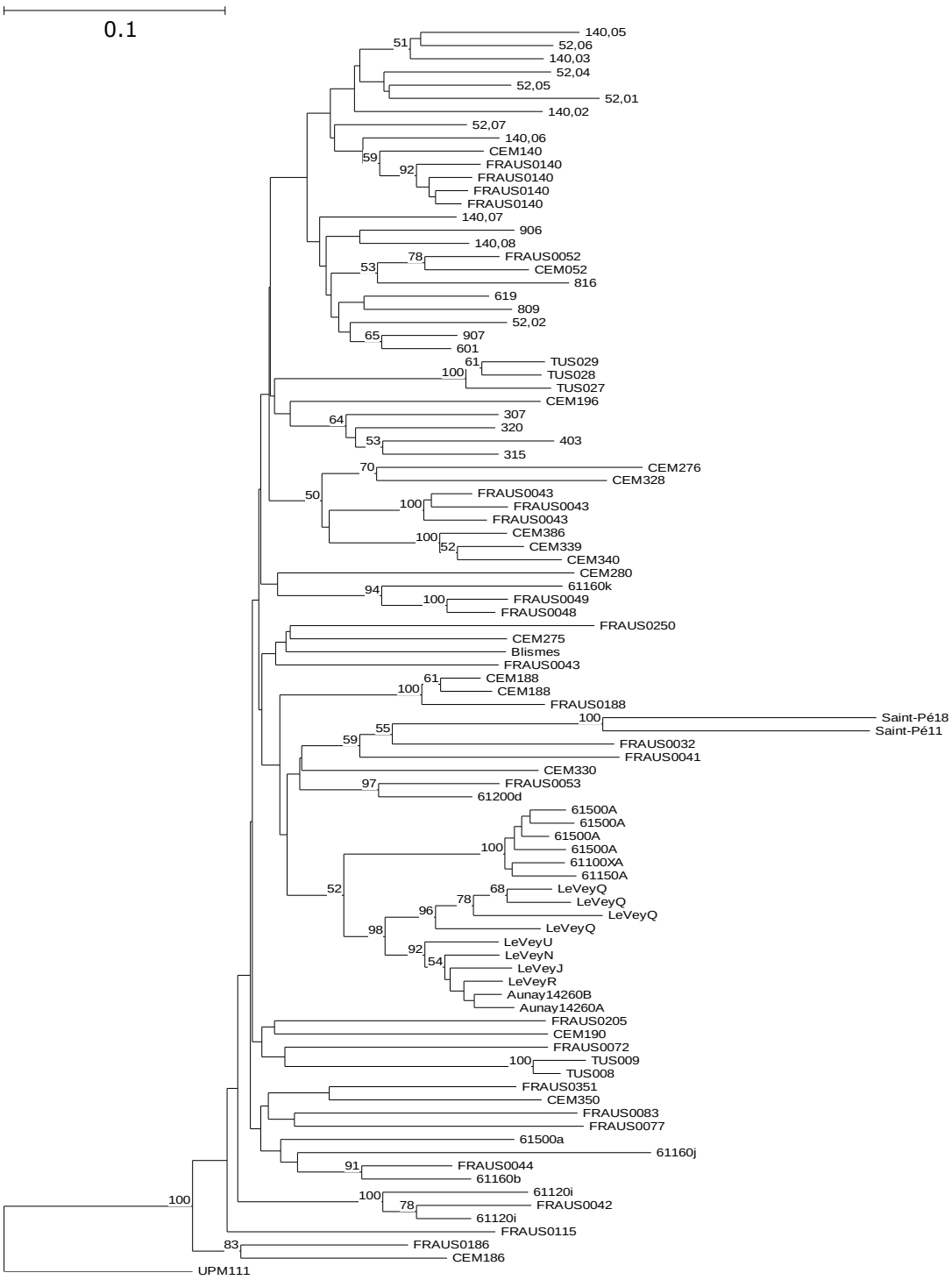

Figure 2: Neighbour Joining tree based on genetic distances according to Nei and Li (1979) (100 bootstraps).

3.3. Clonality

As mentioned above, we used GenoType to define genets among neighbouring trees of *U. minor* (Table 5). Le Vey Q and TUS.027 are the only trees that show a similarity below 0.95 with their neighbours, but are at least very closely related to the surrounding trees. Furthermore, there is a Dice similarity between Le Vey R and Aunay (14260 A and B), between the Le Vey trees N, J, U and Aunay, and between Le Vey Q and Aunay of 0.97, 0.95 and 0.92, respectively. The locations are about 16 km apart from each other according to the available coordinates. As suggested by Cox et al. (2012) root suckers might have been translocated, or ramets of the same ortet were planted on both locations. However, since Le Vey holds more than one MLG, the first scenario seems more likely, with a translocation of plant material from Le Vey to Aunay.

The same relationships are shown in the NJ tree (Fig. 2).

Table 5: Dice similarity and number of markers different between neighbouring individuals of *U. minor*.

Field Code 1	Field Code 2	Dice similarity
Aunay 14260 A	Aunay 14260 B	0.97
Le Vey Q	Le Vey N	0.92
Le Vey Q	Le Vey J	0.92
Le Vey Q	Le Vey U	0.92
Le Vey Q	Le Vey R	0.92
Le Vey N	Le Vey J	0.95
Le Vey N	Le Vey U	0.95
Le Vey N	Le Vey R	0.95
Le Vey J	Le Vey U	0.95
Le Vey J	Le Vey R	0.95
Le Vey U	Le Vey R	0.95
TUS.008	TUS.009	0.96
TUS.027	TUS.028	0.93
TUS.027	TUS.029	0.93
TUS.028	TUS.029	0.95

3.4. Species identification

The species of two trees proved to be difficult to determine solely based on their morphology.

We constructed a PCoA plot using the Nei genetic distance matrix with data standardization with Genalex 6.4 (Peakall and Smouse 2006) (Fig. 3). Only samples of *U. minor*, *U. pumila*, *U. minor* x *U. pumila* and the hybrid cultivar *Ulmus* 'Den Haag', which is derived from a crossing of *U. pumila* and *U.* x *hollandica* 'Belgica', were included. These samples were selected from the set of Cox et al. (2012), where their species were checked genetically. The plot shows that both samples, CEM350 and Blismes, are well embedded in the *U. minor* cluster. This was also the case in the PCoA of the total dataset, including samples of *U. glabra*, *U. procera*, *U.* x *hollandica* and several cultivars (results not shown). Consequently, both samples were visibly differentiated from *U. glabra* and *U.* x *hollandica*. This confirms the findings of Cox et al. (2012).

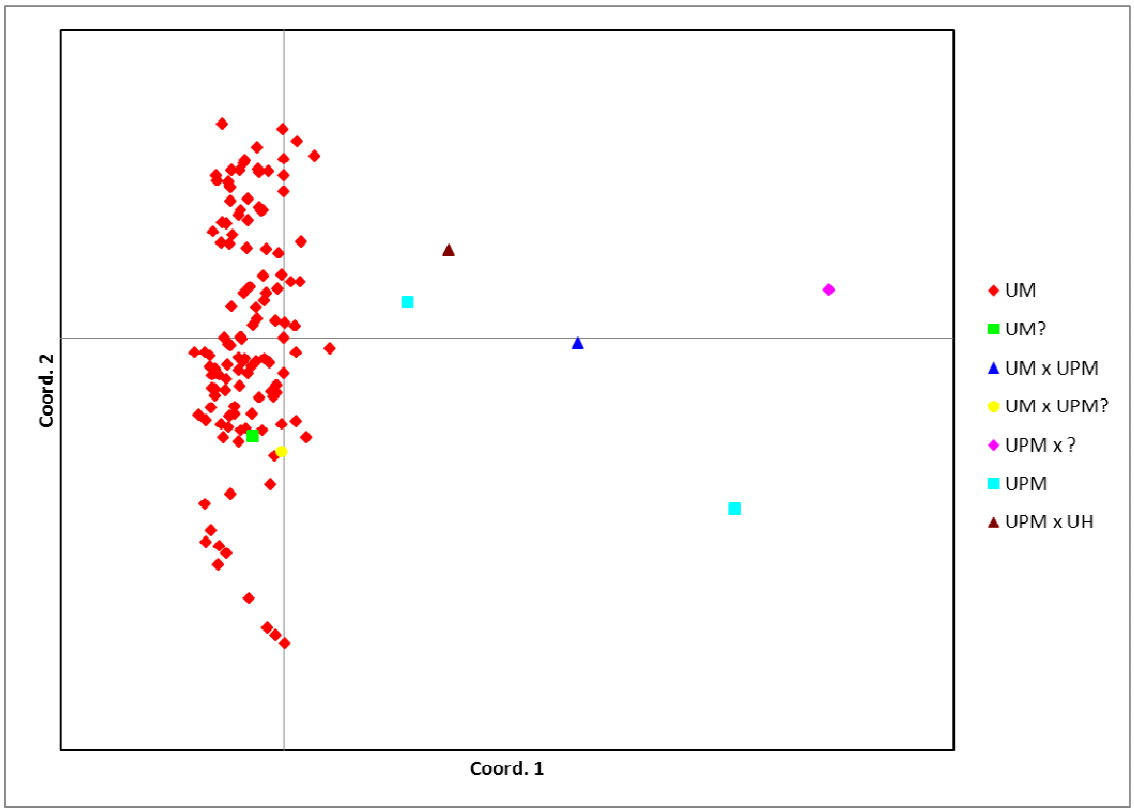

Figure 3: PCoA plot using Nei genetic distance matrix with data standardization. 33.18 % and 19.24 % of the variation is explained by the first and second axis, respectively. UM: *Ulmus minor*; UM?: Blismes; UM x UPM: *U. minor* x *U. pumila*; UM x UPM?: CEM350; UPM x ?: 73P (open pollinated *U. pumila*); UPM: *U. pumila*; UPM x UH: hybrid cultivar 'Den Haag'.

3.5. Clone IDs for marketing of forest reproductive material

The requested AFLP fingerprints are listed in a separate excel file (Irstea_Ulmus_AFLP2012.xlsx) as supplementary information.

Considering the high dissimilarity between the following samples, we recommend repeated analyses to make sure no technical mistakes occurred: 61160j vs. FRA.US.0048, 61500a vs. FRA.US.0072 vs. 61500A and 61150A vs. FRA.US.0043.

Remark: The PCoA plot of the total dataset mentioned in section 3.5, showed that FRA.US.0032 and FRA.US. 0041 appear to be more like hybrids (not like pure *U. minor*). Potentially, they are backcrosses with *U. minor*. This was not investigated further.

References

Cox K., Vanden Broeck A., Vander Mijnsbrugge K. (2012). Genetic variation in European elms: Genetic characterisation of populations of *Ulmus laevis* and of the *U. minor-U. glabra* complex. Rapporten van het Instituut voor Natuur-en Bosonderzoek 2012 (54). Instituut voor Natuur- en Bosonderzoek, Brussels.

Arrigo, N., J. W. Tuszynski, D. Ehrich, T. Gerdes, and N. Alvarez. 2009. Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. Bmc Bioinformatics **10**.

Jones, O. R. and J. Wang. 2009. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources **10**:551-555.

Meirmans, P. G. and P. H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Molecular Ecology Notes **4**:792-794.

Nei, M. and W.-H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences of the United States of America **76**:5269-5273.

Peakall, R. and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes **6**:288-295.

Vekemans, X., T. Beauwens, M. Lemaire, and I. Roldan-Ruiz. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. Molecular Ecology **11**:139-151.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Vandelee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. Aflp - a New Technique for Dna-Fingerprinting. Nucleic Acids Research **23**:4407-4414.

Wang, J. 2004. Estimating pairwise relatedness from dominant genetic markers. Molecular Ecology **13**:3169-3178.

Wang, J. and A. W. Santure. 2009. Parentage and Sibship Inference From Multilocus Genotype Data Under Polygamy. Genetics **181**:1579-1594.