Newcastle University

ePrints

Logan SA, Phuekvilai P, Wolff K. Ancient woodlands in the limelight: delineation and genetic structure of ancient woodland species *Tilia cordata* and *Tilia platyphyllos* (Tiliaceae) in the UK.
***Tree Genetics & Genomics* 2015, 11: 52.**

Copyright:

The final publication is available at Springer via http://dx.doi.org/10.1007/s11295-015-0872-z

Date deposited:

13/05/2015

Embargo release date:

16 May 2016

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License

Newcastle University ePrints - eprint.ncl.ac.uk

Ancient woodlands in the limelight: delineation and genetic structure of ancient woodland species *Tilia cordata* and *T. platyphyllos* (Tiliaceae) in the UK

Abstract

Ancient woodlands in Europe are currently subjected to fragmentation, which leads to ecological deterioration and genetic isolation. This is likely to intensify with further climate change and increased land use. *Tilia* are keystone, ancient woodland indicators, represented in the United Kingdom by two species (*Tilia cordata* and *T. platyphyllos*) and their hybrid (*T. x europaea*). For management of the species we need to investigate genetic differentiation and population genetic structure. Morphology can usually separate the two species but the distinguishing features are not always observable, particularly in less optimum, shady, conditions. Here we test for the first time, 13 microsatellite markers to discriminate the species and hybrid, and to assess the population genetic diversity of the two species. The markers successfully distinguished the species and hybrid. The two species show significant genetic differentiation, indicating two distinct biological units. Sub-structure within both species is loosely related to geographic location. Because *Tilia* populations in the UK are highly fragmented future conservation should consider provenance issues.

Key words Genetic diversity - microsatellites - population genetic structure - species delineation - *Tilia*.

Introduction

Ancient woodlands are areas of land with continuous woodland cover since c1600 (Peterken 1977), and have been of considerable conservation concern within the UK and Europe (Hermy et al. 1999; Pigott 1969; Rackham 2008). They may be recognized by the presence of indicator species and are regarded as remnants of old European forest (Peterken 1974; Pigott 1969; Rackham 2003). Evidence of continuous cover at some sites can be found from pollen records (Pigott 2012). While some ancient woodlands within the UK are protected as features of statutory designated sites (Natural England, 2014), many are highly fragmented and under threat of further fragmentation (Rackham 2008). This can result in the loss of ecological interactions (Bailey et al. 2002; Petit et al. 2004; Riutta et al. 2014) and genetic integrity (Cottrell et al. 2009) and may intensify the effects of climate change on biodiversity (Opdam and Wascher 2004).

Fragmentation causes populations to breakup into smaller groups and can lead to isolation (Young et al. 1996). In remnant ancient woodlands, past genetic variation may therefore have been lost over time. However, Hamrick (2004), suggests that fragmentation may have a lesser effect on some tree species than on other organisms due to their long generation times, elevated genetic diversity and extensive pollen flow. Contrasting perspectives add an uncertainty about the genetic status of small fragmented woodlands. It is therefore important to understand the level of genetic diversity that remains in remnant stands and the level of gene flow between fragments.

The description and identification of species of plants has been primarily based on morphological traits. However, phenotypic plasticity and hybridization can confound identification and lead to misclassification (Duminil and Di Michele 2009; Meimberg

et al. 2010). Studies in Oak (*Quercus*) have focused on the effects of hybridization, and population divergence on species delineation and have shown some incongruities following morphological and molecular analyses (Aldrich et al. 2003; Bacilieri et al. 1995; 1996). Incorrect identification could potentially have negative effects on conservation or management (Barrett and Freudenstein 2009; Bateman et al. 2008).

Tilia spp. in the UK are ancient woodland indicator species, (Babington 1862; Rackham 2008). *Tilia* L.. (lime, linden, or basswood), is a genus of broad-leaved temperate forest trees (Elwes and Henry 1913), with an almost circumboreal distribution. Two species, *T. cordata* Mill. (small-leaved lime), and *T. platyphyllos* Scop. (large-leaved lime), are widely distributed and sympatric in central and western Europe, including England and Wales, where their hybrid *T. x europaea* L. (syn. *T. x vulgaris*) occurs naturally and has been widely planted in towns and parks (Pigott 1969; 1981). They are insect pollinated and seed dispersal can be wind facilitated (Pigott 2012).

Species within the genus have similar morphological traits, which led to ambiguity in species designation (Pigott 2012). Distinguishing features that differentiate *T. cordata* and *T. platyphyllos* may not always be easily observable from the ground and given that *Tilia* can often grow to 30 – 40m (Pigott 2012), and even taller in some parts of eastern Europe (Pigott 1975; Wesolowski and Rowinski 2006), collecting leaf samples from the upper canopy may be problematic. Moreover, considering the two UK species naturally hybridize and display intermediate characteristics (Pigott 1969), identifying features can become less useful. To confirm identification of the two species and hybrids, genetic techniques should be applied.

At present, there is a large amount of ecological information available regarding *Tilia* (Pigott 2012). In contrast, there is very little genetic information due to the limited number of useful molecular markers available. As the two species are important ecological components of ancient woodland in the UK and in woodland throughout Europe, we need to understand their genetic diversity as well as the level of hybridization between the two. As UK woodlands have been fragmented for some time, we expect that gene flow between populations will have been restricted, allowing populations to differentiate. Informative genetic markers will allow us to investigate this further.

Earlier genetic studies using chloroplast PCR-RFLP markers (Fineschi et al. 2003), RAPD markers, (Hosseinzadeh Colagar et al. 2013; Liesebach and Sinkó 2008); and the ITS regions (Yousefzadeh et al. 2012) were successful in determining various levels of genetic diversity. However, RAPDs are dominant markers and therefore restrict population genetic analyses, while chloroplast and ITS regions are generally conserved and may not differentiate the two UK species. Other authors have used isozyme markers (Fromm and Hattemer 2003; Maurer and Tabel 1995), and to some extent were able to determine species and clones. However, the need for fresh tissue for isozyme/allozyme studies may be a disadvantage over DNA based approaches where samples can be stored until required. Fewer loci and lower diversity has also been observed using isozymes compared to microsatellite markers in some tree species (Pfeiffer et al. 1997; Streiff et al. 1998; Sun et al. 2001). Microsatellites reveal high levels of diversity and allow quantification of population genetic structure.

Microsatellites have recently been developed for *T. platyphyllos* (Phuekvilai and Wolff 2013). Many of these markers can be successfully amplified in other *Tilia*

species, including *T. cordata*. Phuekvilai and Wolff (2013) reported that one locus (*Tc918*) amplified in *T. platyphyllos* but not in *T. cordata* and so could be useful in species delineation.

In this study we aim to test for the first time, the resolution of 13 variable nuclear microsatellite markers to discriminate between the two UK species and to distinguish them from the hybrid. More specifically, given the fragmented nature of many UK *Tilia* populations, we aim to test three hypotheses; (1) although hybridization is known to occur between the two species, *T. cordata* and *T. platyphyllos* are two separate evolutionary units and will show high genetic differentiation, (2) as remnant populations have undergone long-term fragmentation within the UK they will show a high degree of genetic structure and limited gene flow among populations, and (3) small, isolated populations will show low genetic diversity due to stochastic genetic effects and inbreeding. This will broaden our understanding of the molecular ecology of *Tilia*, in the UK and aid in any future restoration and conservation of the two autochthonous species.

Material and Methods

Study sites and sample collection

Leaf samples were collected from 27 locations (Table 1, Online Resource 1), spanning most of the UK range of the two species. In total 550 *Tilia* samples were collected from the two species and the hybrid across all locations. Population size ranged from five individuals in a small isolated Worcestershire site to 45 individuals from Chanstone Wood, a large mixed site. Most sites were documented to have *T. cordata* or *T. platyphyllos* present and all sites are considered to be of ancient and/or semi-natural origins. Many are presently designated as Sites of Special Scientific

Interest (SSSI) or National Nature Reserves (NNR, Natural England, 2014). Samples from some Worcestershire and Herefordshire sites are believed to be remnants of parish boundaries and old access routes from the pre-Enclosures Act (R. Roseff pers. comm.). Although all sites are considered to be of ancient or natural origin, they would have been managed in the past (*e.g.* coppiced).

Many of the *T. cordata* samples were kindly provided by Dr Paul Ashton (Edge Hill University). These were collected between June and August 2012. Additional *T. cordata* and all *T. platyphyllos* samples were collected in August 2012 and between June-July 2013. One leaf was taken from each tree, and a GPS point was recorded for most samples. For those that could not be given a GPS point due to canopy cover or topology, an eight digit grid reference was recorded. The grid references and GPS points were later transformed into a reference point and mapped in DIVA-GIS v7.5 (www.diva-gis.org). All samples were dried immediately at room temperature between tissues. Once dried all leaves were stored at -20°C until DNA extraction. In total 550 *T. platyphyllos, T. cordata* and putative hybrids were sampled from the 27 sites (Table 1).

Table 1 Sites sampled, code, putative species present, latitude and longitude coordinates

Population	Code	Species	Latitude (°N)	Longitude (°E)
Brignall Banks	BB	*T. cordata*	54.4968	-1.9114
Applegarth Scar	AS	*T. platyphyllos*	54.4093	-1.8165
Whitcliffe Wood	WW	*T. platyphyllos*	54.4097	-1.7792
Hayburn Wyke	HbW	*T. platyphyllos*	54.3600	-0.4518
Anston Stone Wood	AW	*T. platyphyllos*	53.3402	-1.1985
Hardy Gang Wood	HG	*T. cordata*	53.2607	-0.3614
Collyweston Wood	CGW	*T. cordata*	52.5974	-0.5176
Easton Hornstock	EaH	*T. cordata*	52.5912	-0.4981
Bedford Purlieus	BP	*T. cordata*	52.5833	-0.4646
Barton Hills	BH	*T. platyphyllos*	51.9562	-0.4220
Dowles Brooke	DB	*T. cordata*	52.3834	-2.3364
Shrawley Wood	ShW	*T. cordata*	52.2917	-2.2820
Lulsley	LU	Mixed	52.1914	-2.3888
Suckley Hills	SU	Mixed	52.1826	-2.3921
Sheepshill Coppice	SC	Mixed	52.1595	-2.3772
Knapp & Papermill	KP	Mixed	52.1615	-2.3713
Halesend Wood	HW	Mixed	52.1409	-2.3815
West Malvern	WM	Mixed	52.1189	-2.3581
Brockhill Wood	BW	Mixed	52.0928	-2.3538
Collin Park Wood	CPW	*T. cordata*	51.9463	-2.3695
Lady Park Wood	LP	Mixed	51.8245	-2.6568
Woolhope	WH	Mixed	52.0259	-2.6042
Eywas Harold	EH	Mixed	51.9596	-2.9136
Chanstone Wood	CW	Mixed	52.0111	-2.9422
Covenhope	CH	Mixed	52.2837	-2.8855
Roudsea Wood	RW	*T. cordata*	54.2332	-3.0255
Skelghyll Wood	SkW	*T. cordata*	54.4197	-2.9519

DNA extraction and amplification

Genomic DNA was extracted using the CTAB (Cetyl Trimethyl Ammonium Bromide) method as described in Morgan-Richards and Wolff (1999). Extracted DNA was dissolved in 80 µl of TE (Tris-EDTA) buffer and stored at -20° C until required for DNA amplification. Four multiplex Polymerase Chain Reaction (PCR) procedures were carried out to amplify 13 microsatellite regions using primers developed for *Tilia* (Phuekvilai and Wolff 2013). Amplification was carried out in a final volume of 10µl, consisting of 9µl of PCR master mix and 1µl (5ng) of DNA. Primer sequences, repeat motifs, allele size, fluorescent dyes, and primer concentration, as well as PCR conditions and parameters are described in Phuekvilai and Wolff (2013). PCR products were diluted to a 1/10 concentration with dH_2O, and stored at -20° C until required for genotyping.

Microsatellite genotyping

Electrophoresis was carried out with 1 µl of PCR product and a 10 µl mix of 10x Hi-Di formamide and 0.1x GeneScan™ 500 ROX size standard. Microsatellites were genotyped using an ABI 3130XL Genetic Analyser (Applied Biosystems), and allele fragment sizes were determined using GeneMapper® v4.0 (Applied Biosystems) software. Fragments were binned manually and checked for inconsistencies. To confirm consistent allele scoring, a selection of PCR products were re-scored and compared. We used GenAlEx v6.5 (Peakall and Smouse 2012) to identify individuals sharing multi-locus genotypes. When more than one identical genotype (clone) was detected, only one individual from each clone was kept in the dataset. Genotyping errors were assessed in MICRO-CHECKER v2.2.3 (Van Oosterhout et al. 2004). The program applies a Monte-Carlo approach to check for scoring errors due to

stuttering, and large allele dropout and uses Hardy-Weinberg Equilibrium (HWE) to detect the presence of null alleles.

Species demarcation and hybrid identification

To determine the power of the markers in grouping the samples, we carried out a Principal Coordinates Analysis (PCoA) based on individual pairwise genetic distances in GenAlEx v6.5. We used the Bayesian clustering program STRUCTURE v2.3.4 (Falush et al. 2003; Pritchard et al. 2000) to assign individuals to clusters within the dataset. The program assigns individuals to a predefined number of clusters (K), based on the allele frequencies at each locus. In this case we chose K = 2, with the assumption that the two species, *T. cordata* and *T. platyphyllos,* represent genetically distinct taxonomic units. STRUCTURE parameters were kept at the default settings, with a burn-in of 10^4 and MCMC iterations of 10^5. We also ran STRUCTURE with different values of K (ranging 1 – 5) with the default settings and same parameters. Each run was replicated ten times. We used the Evanno ΔK method (Evanno et al. 2005) in STRUCTURE HARVESTER (Earl and vonHoldt 2012) to observe if our assumption of K = 2 was justified and used DISTRUCT (Rosenberg 2004) to visualize the data.

Values of q, ranging from 0 – 1, describe how individuals are proportionally assigned to a particular cluster. For example, individuals with q = 0.50 mean that 50% of their genotype belongs to one cluster and 50% to another cluster, *i.e.* a hybrid. However, individuals can be unequally proportioned within clusters, so to identify parental species and hybrids a cut-off value is required. Other authors have reported threshold q-values of 0.10 (Neophytou 2014) and 0.20 (Duminil et al. 2006; Larcombe et al. 2014; Vähä and Primmer 2006). We used a threshold q-value of

0.20 to identify the parental species and hybrids as this is a balance between accuracy and efficiency (Vähä and Primmer 2006). Individuals with *q*-values between 0.20 – 0.80 were considered to be hybrids. They were removed from the dataset before population structure analyses. To be confident in our threshold of 0.20, a *q*-value of 0.10 was also tested and results compared.

Genetic diversity and differentiation

After identifying and removing hybrids and duplicated genotypes, standard population genetic diversity statistics were obtained from GenAlEx v6.5. Mean number of alleles (N_A), effective number of alleles (A_E), mean observed (H_O) and expected heterozygosity (H_E) were calculated from populations with five or more remaining individuals. The Inbreeding Coefficients (F_{IS}) and significance of each population was determined using FSTAT v2.9.3.2 (Goudet 1995).

Genetic differentiation was assessed using two measures, pairwise F_{ST} (Weir and Cockerham 1984) and D_{est} (Jost 2008), calculated in GENEPOP on the web v4.2 (Raymond and Rousset 1995) and SMOGD (Crawford 2010), respectively. D_{est} is an estimated value of actual differentiation (Jost 2008), and was calculated using 1000 bootstrap replicates to generate 95% confidence intervals (CI). We performed an Analysis of Molecular Variance (AMOVA) in the program Arlequin v3.5.1.3 (Excoffier and Lischer 2010), to determine the distribution of genetic variation (among species, among populations within species, and within populations). Statistical tests for differences between diversity measures were carried out in Minitab® v17.1 (2013 Minitab Inc.).

Population structure

We assessed within-species population structure using the program STRUCTURE v2.3.4. For *T. cordata* we set K to range from 1 – 18 and for *T. platyphyllos*, *K* ranging 1 – 10. We used both the default settings and the LOCPRIOR model (Hubisz et al. 2009) with a burn-in of 5 x 10^4 followed by 5 x 10^5 MCMC iterations and each run was replicated 20 times. We used the Evanno *ΔK* method in STRUCTURE HARVESTER to determine optimal *K*. Different runs for the same *K* were averaged using the program CLUMPP (Jakobsson and Rosenberg 2007) following the Greedy algorithm with 1000 repeats, and the data were visualised in DISTRUCT. To assess if genetic distance was correlated with geographic distance (Isolation by distance), a Mantel test was run with 10^4 permutations on each species, implemented in GenAlEx v6.5.

Results

In total 412 *Tilia* samples from 27 populations were analysed following the removal of 138 duplicated genotypes (clones). MICRO-CHECKER v2.2.3 revealed no evidence of scoring errors due to stuttering with the exception of locus *Tc31* in population TcBB and no large allele dropout. The program revealed homozygote excess at some loci suggesting the possible presence of null alleles. However, loci showing homozygote excess were different in different populations, with the exception of locus *Tc963* which showed homozygote excess in four *T. cordata* and three *T. platyphlyllos* populations (Online Resource 2). Therefore, we ran all analyses with and without this marker to see if the results differed. The effect was negligible and so results including the marker are shown.

Species demarcation and hybrid identification

Two clear groups with some intermediate individuals were revealed by PCoA analysis based on pairwise genetic distance (Fig. 1). The orange cluster consists of individuals which had previously been documented as being *T. platyphyllos* while the blue cluster consists of individuals thought to be *T. cordata*.

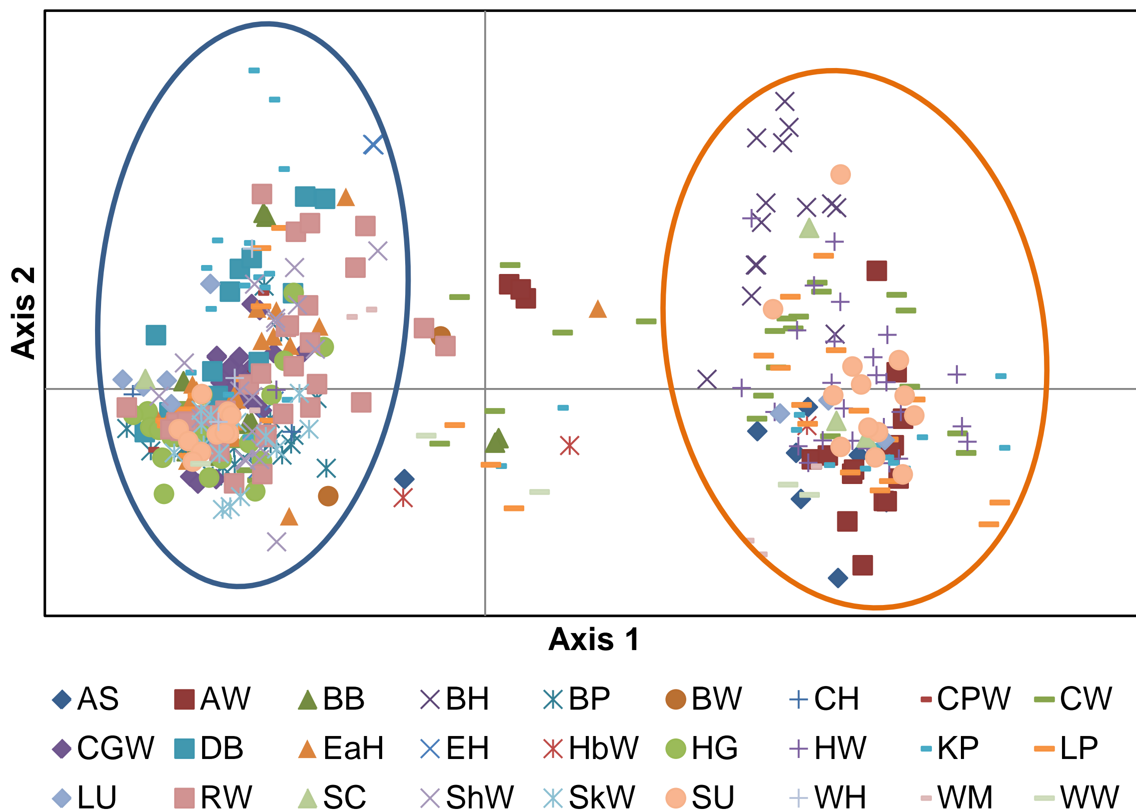


Fig. 1 PCoA of *Tilia* individuals from 27 populations. Blue cluster individuals are *T. cordata* and the orange cluster are *T. platyphyllos*. Intermediate points are putative hybrids. Axis 1 and 2 explains 41% of the genetic variation (36% and 5%, respectively)

Two groups ($K = 2$) representing the two species were observed in STRUCTURE v2.3.4 (Fig. 2). Evanno's ΔK method also revealed that $K = 2$ was optimal (Online

Resource 3). Using a threshold *q*-value of 0.10, 32 individuals (7.8%), were determined hybrids, while using a *q*-value of 0.20, indicated 25 individuals to be hybrids (6.1%, Online Resource 4). The 32 putative hybrid individuals had *q*-values for one parental species of between 0.105 and 0.866, while the 25 hybrid individuals had *q*-values for the same parental species of between 0.230 and 0.735 (Online Resource 5). All loci were polymorphic with the exception of locus *Tc8* which was monomorphic in *T. cordata*. This locus was fixed in *T. cordata* at 141bp and ranged from 156 – 170bp in *T. platyphyllos* suggesting this locus has species specific alleles. Of the 32 putative hybrids using the 0.10 *q*-value threshold, 22 had the expected combination of *T. cordata* specific and *T. platyphyllos* specific alleles. Ten individuals were not a combination of *T. cordata* and *T. platyphyllos* alleles at this locus. Seven were homozygote with the *T. cordata* allele (141bp) and three were heterozygote *i.e.* had two of the *T. platyphyllos* alleles (156, 158, 160, 162, 164, 166 or 170bp). Of the 25 putative hybrids using the 0.20 threshold, two were homozygote at this locus with only the *T. cordata* allele and one was heterozygote with two *T. platyphyllos* alleles (Online Resource 5). Two of the populations originally thought to be *T. platyphyllos* (WW and HbW), contained 19 (out of 20) and 11 (out of 13) *T. x europaea*, respectively.

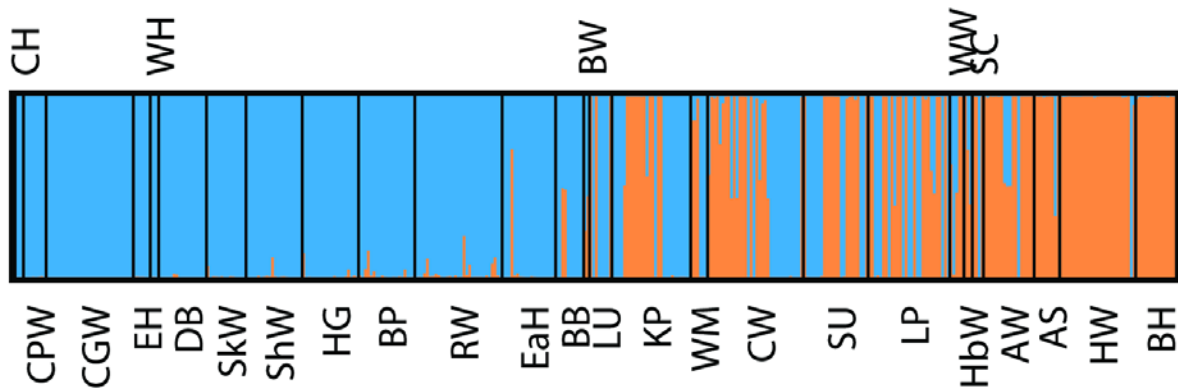


Fig. 2 Assignment of 412 individuals from 27 populations with $K = 2$ (*T. cordata* – blue cluster and *T. platyphyllos* – orange cluster) inferred by Bayesian clustering analysis implemented in STRUCTURE, visualized in DISTRUCT

Genetic diversity and differentiation

After removing hybrids, a total 380 (using *q*-value 0.10) and 387 (using *q*-value 0.20) *Tilia* individuals remained over 24 populations. The 13 microsatellite markers revealed high levels of polymorphism in both species, with the exception of locus *Tc8*. Diversity statistics were reported from populations with five or more individuals. The number of alleles per locus in the two species ranged from nine (loci *Tc8* and *Tc943*) to 33 (locus *Tc963*), a total of 206 alleles (Table 2). Overall *T. cordata* exhibited fewer alleles than *T. platyphyllos*, 127 and 149, respectively (Online Resource 6). The number of alleles ranged from one (locus *Tc8*) to 26 (locus *Tc963*) in *T. cordata*, and from five (*Tc943*) to 18 (*Tc915*) in *T. platyphyllos*. More than twice as many private alleles were observed in *T. platyphyllos* than in *T. cordata* (21 and nine, respectively, Online Resource 6). The average number of alleles per population (N_A) was significantly greater in *T. platyphyllos* (2 sample *t*-test, $P = 0.014$), and this was also reflected in the genetic diversity (H_E), which was higher in *T. platyphyllos* than in *T. cordata* (Table 3). Effective number of alleles (A_E) ranged from 1.93 (TcLu)

to 3.54 (TcEaH) in *T. cordata* while observed (H_O) and expected (H_E) heterozygosity ranged from 0.34 (TcLu) to 0.59 (TcShW) and 0.35 (TcLU) to 0.57 (TcRW), respectively (Online Resource 7). In *T. platyphyllos* these values were significantly higher ($P = <0.05$), with A_E ranging from 2.68 (TpAS) to 4.00 (TpSU), and H_O and H_E ranging from 0.66 (TpSU) to 0.78 (TpBH) and 0.60 (TpAS) to 0.73 (TpSU), respectively (Online Resource 8). Inbreeding coefficients (F_{IS}) ranged from -0.14 to 0.14 and were not significant (data not shown).

Genetic differentiation between populations within each species across all loci, measured as F_{ST} and D_{est} was higher in *T. platyphyllos* than in *T. cordata* (two-tailed Mann-Whitney *U* test, for F_{ST}: $P = 0.051$ and for D_{est}: $P = 0.006$, Table 2). Significant population pairwise $F_{ST,}$ among the two species, was revealed by GENEPOP ranging from 0.305 (TcShW and TpBH) to 0.454 (TcLU and TpAS, Online Resource 9).

There was high genetic differentiation between the two species, (28.54% of the total variation, AMOVA, $P = <0.001$). Among populations within species variation was 7.05% ($P = <0.001$) while the remaining variation (64.41%, $P = <0.001$) was found within populations (Online Resource 10). Based on between species F_{ST} values loci *Tc8* and *Tc927* displayed the highest difference, 0.734 and 0.625, respectively (Table 2). While these loci also showed high differentiation based on D_{est}, the range of variation across D_{est} values were smaller than those across F_{ST} values.

Table 2 Number of alleles (*A*), Observed heterozygosity (H_O), Expected heterozygosity (H_E), F_{ST} and D_{est} values within *T. cordata* & *T. platyphyllos*, and differentiation values between species

Locus	*T. cordata*					*T. platyphyllos*					*T. cordata and T. platyphyllos*		
	A	H_O	H_E	F_{ST}	D_{est}	*A*	H_O	H_E	F_{ST}	D_{est}	*A*	F_{ST}	D_{est}
Tc6	8	0.63	0.54	0.079	0.108	11	0.77	0.72	0.106	0.339	12	0.260	0.848
Tc937	6	0.44	0.41	0.077	0.054	9	0.68	0.61	0.121	0.253	13	0.404	0.935
Tc920	11	0.79	0.74	0.062	0.265	11	0.83	0.73	0.082	0.270	16	0.146	0.771
Tc8	1	0.00	0.00	0.000	0.000	8	0.71	0.63	0.087	0.194	9	0.734	0.987
Tc943	8	0.50	0.45	0.103	0.114	5	0.51	0.48	0.045	0.042	9	0.433	0.802
Tc31	6	0.27	0.33	0.104	0.055	10	0.65	0.61	0.142	0.313	13	0.416	0.890
Tc4	13	0.62	0.71	0.097	0.321	13	0.72	0.71	0.125	0.376	17	0.139	0.750
Tc927	4	0.03	0.03	0.533	0.128	15	0.74	0.72	0.089	0.282	17	0.620	0.978
Tc915	16	0.86	0.78	0.048	0.160	18	0.88	0.77	0.097	0.452	22	0.101	0.703
Tc963	26	0.77	0.86	0.046	0.439	15	0.65	0.69	0.133	0.486	33	0.088	0.816
Tc5	13	0.72	0.66	0.062	0.156	12	0.82	0.74	0.092	0.401	18	0.124	0.555
Tc951	7	0.56	0.53	0.095	0.131	9	0.52	0.67	0.093	0.267	11	0.234	0.687
Tc7	8	0.54	0.51	0.052	0.075	13	0.78	0.71	0.105	0.287	16	0.314	0.957
Mean	9.8	0.52	0.50	0.078	0.095	11.5	0.71	0.68	0.103	0.267	15.9	0.309	0.802

Table 3 Mean (and SE) values of diversity measures from 16 *T. cordata* populations and eight *T. platyphyllos* populations. N – number of samples; N_A – Average number of alleles; A_E – Effective number of alleles; H_O – Observed heterozygosity; H_E – Expected heterozygosity

Pop	N	N_A	A_E	H_O	H_E
T. cordata	15.38 (1.93)	4.36 (0.23)	2.90 (0.10)	0.52 (0.01)	0.51 (0.01)
T. platyphyllos	14.5 (1.84)	5.45 (0.33)	3.46 (0.15)	0.71 (0.01)	0.68 (0.15)

Population sub-structure

Bayesian clustering analysis revealed population structure within *T. cordata*. When the default settings (data not shown) and the LOCPRIOR model was used, STRUCTURE v2.3.4 analysis and Evanno's *ΔK* method suggested $K = 3$ as optimal (Fig. 3 and Online Resource 11). Populations CW, DB, EH, KP, LP, LU, ShW, and SU were assigned with large probability, to one group and population CPW also had high membership to this group. RW and SkW were assigned as a separate group with a high membership coefficient. Populations BP, EaH, and HG, although admixed, sharing membership with the first group, had a high membership coefficient with CGW, the third group. The remaining population, BB, is admixed with even probability of belonging to all three groups.

Likewise, there is clear population structure in *T. platyphyllos* (Fig. 4 and Online Resource 12). When both models were used, STRUCTURE and Evanno's *ΔK* method suggested optimal $K = 3$. Populations AS, BH, and HW were assigned to different groups. Individuals from the other populations (AW, CW, KP, LP, and SU), were not assigned to a single cluster.

GenAlEx v6.5 revealed a non-significant negative correlation between genetic and geographic distance (Pairwise F_{ST}) in *T. cordata* (R^2 = 0.0291, P = 0.141). In *T. platyphyllos*, a non-significant positive correlation was observed (R^2 = 0.192, P = 0.06). This suggests no significant isolation by distance occurs in UK *Tilia*.

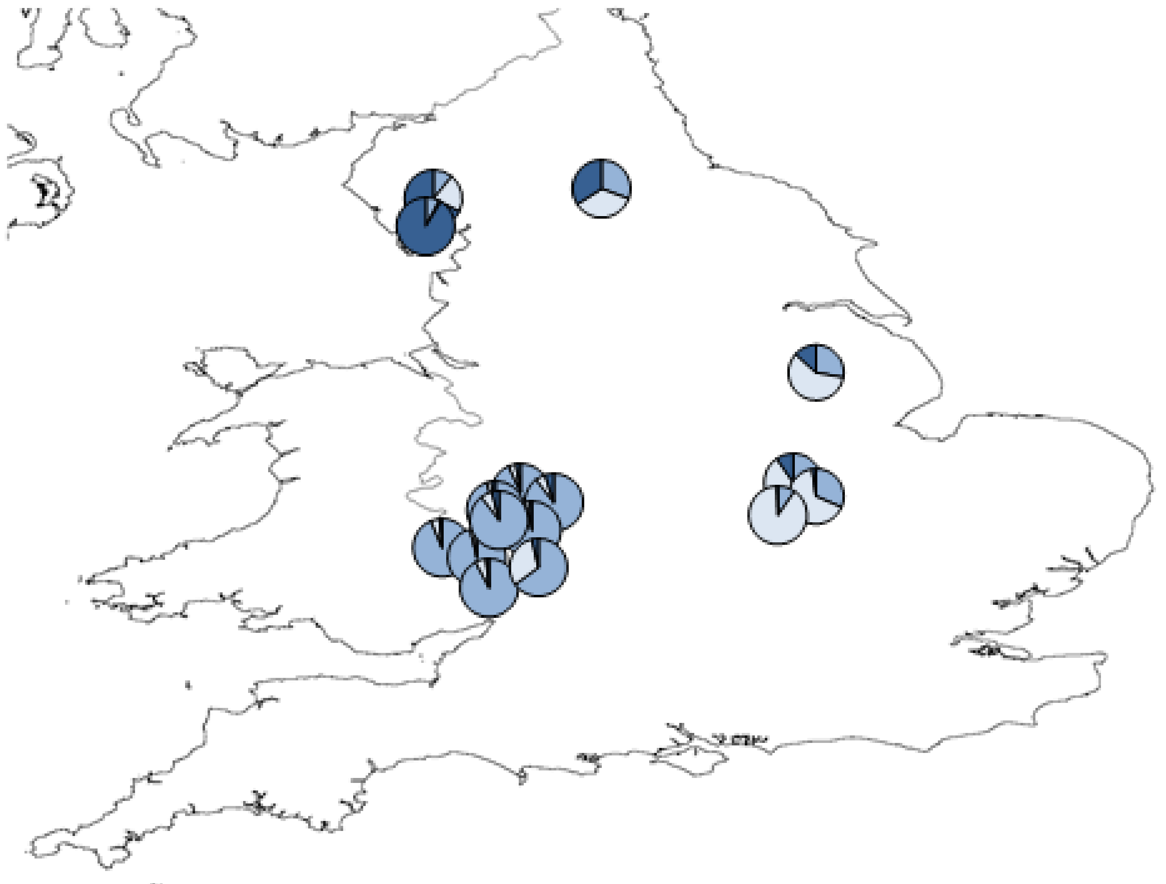


Fig. 3 Distribution of the genetic variation of 246 assigned individuals from 16 *T. cordata* populations with K = 3 inferred by Bayesian clustering analysis implemented in STRUCTURE, averaged in CLUMPP

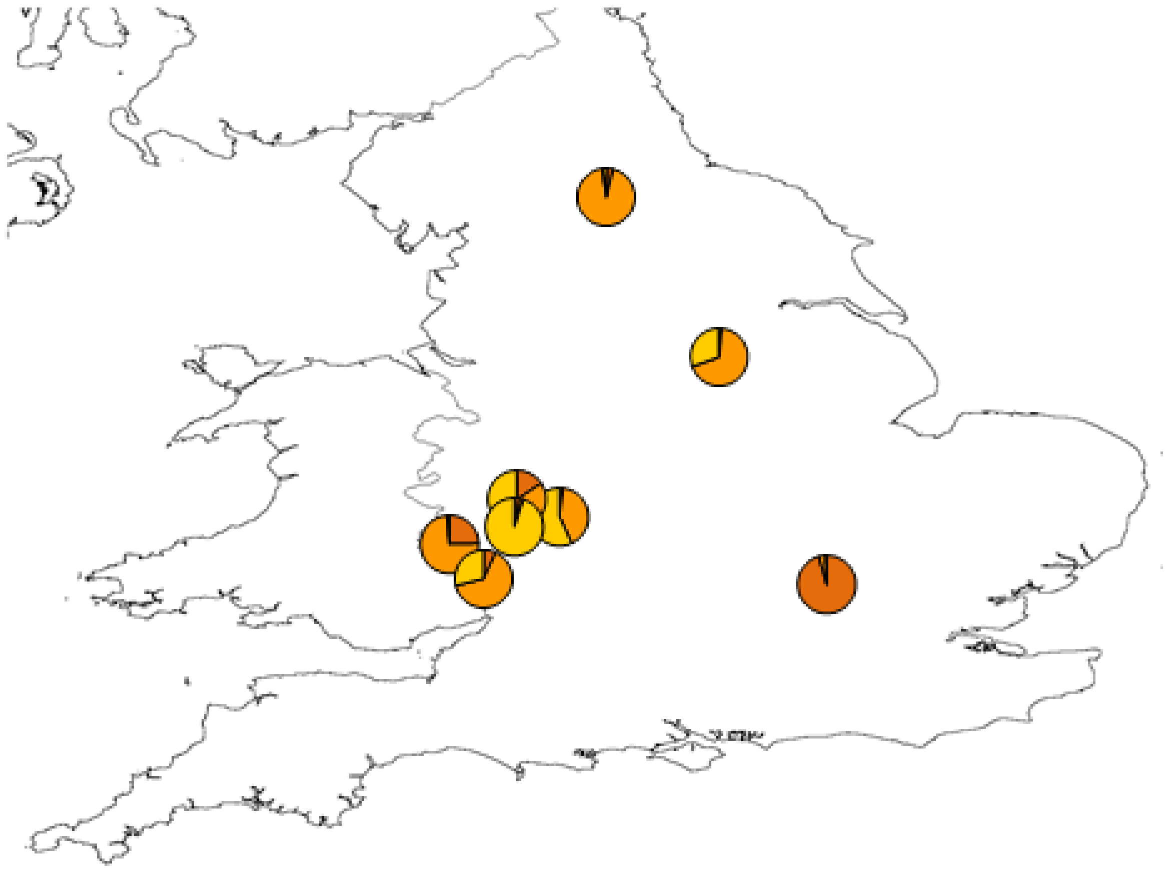


Fig. 4 Distribution of the genetic variation of 116 assigned individuals from eight *T. platyphyllos* populations when $K = 3$ inferred by Bayesian clustering analysis implemented in STRUCTURE, averaged in CLUMPP

Discussion

Genetic analysis of 246 *T. cordata* and 116 *T. platyphyllos* showed that the two species are two distinct biological units. Although phylogenetic analysis show that they are not closely related (Phuekvilai 2014), they naturally hybridise. This was clear in the PCoA and STRUCTURE analyses (Figs. 1 and 2), where we found several intermediate individuals.

Identification of pure species and hybrids from mixed populations

To determine whether an individual belongs to either species or is a hybrid we tested two optimal threshold *q*-values, 0.10 (Neophytou 2014) and 0.20 (Duminil et al. 2006; Larcombe et al. 2014) and found 92.2% and 93.9% of individuals could be assigned to either species (Online Resource 4). The remaining 7.8% and 6.1% were considered to be hybrids or introgressed individuals. Although the difference between the two thresholds in our study was negligible, a *q*-value of 0.20 might be a preferred option. Vähä and Primmer (2006), define three outcomes depending on threshold values chosen *i.e.* efficiency (highest proportion of correctly assigned pure species or hybrids), accuracy (the proportion of true species or true hybrids in the respective group) and performance (the result of considering both). The authors suggest that while *q*-values of 0.10 results in higher efficiency, when taking accuracy into consideration a *q*-value of 0.20 increases the overall performance.

STRUCTURE has generally been reported to do well in detecting hybrids from pure species (Duminil et al. 2006; Larcombe et al. 2014). For example, Larcombe et al. (2014), used a *q*-value of 0.20 and revealed 94-100% accuracy in detecting F_1 hybrids between six Eucalypts species. Similarly, Duminil et al. (2006), set a *q*-value of 0.20 to identify tropical rainforest trees (*Carapa* spp.) following a 'blind' survey *i.e.* without *a priori* information on taxonomy or morphology. STRUCTURE assigned over 90% of their samples to one of the *Carapa* species. To test the assignment power of STRUCTURE, Vähä and Primmer (2006), simulated a series of datasets consisting of different number of loci and divergence levels (F_{ST}). Their study revealed that hybrids can be identified with 12 – 24 loci and pairwise F_{ST} values of 0.21 – 0.12. We used 13 variable markers analysed between species and observed

F_{ST} values ranging from 0.305 to 0.454 (Online Resource 9) and therefore should be sufficient for detecting hybrids in this instance.

High genetic diversity and differentiation in UK Tilia

Tilia cordata and *T. platyphyllos* showed a high degree of genetic diversity within and between species (Table 2). All samples showed a maximum of two alleles per locus, confirming that *T. cordata* and *T. platyphyllos* are both diploid species, 2n=2x=82 (Pigott 2012). Diversity indices from the UK samples were marginally lower than those found in the two *Tilia* species from other European countries (Phuekvilai 2014). This might be expected given that some UK populations are nearing the edge of their ecological range. Range edge populations often have less within population genetic diversity due to genetic drift, limited gene flow, inbreeding and clonal reproduction (Arnaud-Haond et al. 2006; Caughley et al. 1988; Eckert 2001) and show more genetic differentiation among populations than their central range counterparts (Beatty et al. 2008).

With large distances between many of the fragmented populations, gene flow among populations is limited. Significant F_{ST} values confirm this. While the inbreeding coefficient (F_{IS}) values were not significant, suggesting little or no inbreeding, we can confirm a high level of clonal trees (25%). A pilot study has found greater clonal reproduction in *T. cordata* than in *T. platyphyllos* (S. A. Logan and K. Wolff, unpubl. data). *T. cordata* is thought to reproduce more freely from asexual methods than production from seed (Radoglou et al., 2009 and references therein).

Overall, *T. cordata* exhibited less genetic diversity than *T. platyphyllos* (Tables 2 and 3). The difference in diversity might be a result of ascertainment bias as the markers where designed from *T. platyphyllos* individuals (Phuekvilai and Wolff

2013). However, *T. cordata* does not freely regenerate from seed in many parts of its UK range and is rarely planted in native woodlands, while *T. platyphyllos* regularly has regenerated from seed and has been planted (Pigott 1981) and this may be an alternative reason why we see much lower diversity in *T. cordata* compared to *T. platyphyllos*. Another explanation is that cross-pollination from planted *T. platyphyllos* trees, originating from France and the Low Countries since c1500 (D. Pigott pers. comm.) may have increased diversity of UK *T. platyphyllos*. Furthermore, *T. platyphyllos* is thought to cross with planted *T. x europaea* as the two taxa flower simultaneously and share pollinators (D. Pigott pers. comm.).

Using the same set of markers, mean H_E in *T. platyphyllos* is slightly lower than that Phuekvilai and Wolff (2013) found in two French populations with values of 0.74 compared to 0.68 (Table 2). In their study, locus *Tc943* was monomorphic in one population and showed only two alleles in the other. In our study, locus *Tc943* had the fewest alleles in *T. platyphyllos*, with only five found across all eight populations. Mean number of alleles within UK populations of *T. platyphyllos* was 11.5 (Table 2), compared to a mean of 9.1 in the two French populations (Phuekvilai and Wolff 2013).

MICRO-CHECKER showed UK populations are in Hardy-Weinberg equilibrium (HWE). With the two species exhibiting high heterozygosity and non-significant F_{IS} values we can infer that both *Tilia* species within the UK are naturally outcrossing systems.

Outcrossing woody species show high within population genetic variation (Hamrick and Godt 1996). While the AMOVA analysis revealed high differentiation between the species, a small but significant proportion of the total variation was found among populations within species and not surprisingly, the largest proportion

of variation was found within populations (Online Resource 10). While within population variation is comparable, among population variation was lower and among species considerably higher than for *Quercus petraea* and *Q. robur* (Kelleher et al. 2005). While different markers reveal different estimates (Storfer et al. 2010), greater species differentiation might be because the two *Tilia* species diverged earlier from each other than the two *Quercus* species. Significant F_{ST} was found both between and within species (Online Resource 9) indicating two distinct biological units (Duminil et al. 2006). Likewise, per locus F_{ST} and D_{est} were higher between species than within species (Tables 2) revealing sufficient power of the markers in species delineation.

Phuekvilai and Wolff (2013), observed one locus (*Tc918,* not used in this study) did not amplify in *T.* cordata. So this marker is useful as a starting point in species delineation. However, amplification at this locus also occurs in the hybrids, therefore, it is less useful as a single identifying marker. We have revealed other potential loci that can be used to delineate the species. Based on F_{ST}, locus *Tc8* is an obvious candidate. This locus is monomorphic within UK *T. cordata* and exhibits up to eight alleles in *T. platyphyllos*. Locus *Tc927* is another potential candidate because there are very few alleles at this locus within *T. cordata* and F_{ST} is high. Conversely, within *T. platyphyllos* the locus exhibits 15 alleles and F_{ST} is relatively low (Table 2). Large numbers of alleles at this locus have also been observed in two French *T. platyphyllos* populations (Phuekvilai and Wolff 2013). Based on these values, loci *Tc8* and *Tc927* will be useful species discriminators. Other loci showing high between species variation are *Tc937, Tc943,* and *Tc31*, and will also be useful in distinguishing the two species.

Observed intra-specific structure in UK populations

Sub-structure within both *T. cordata* and *T. platyphyllos* was observed, ($K = 3$). All putative groups assigned by STRUCTURE within each of the species can loosely be split by geographic region and membership within each putative group is not unexpected due to the close proximity of the respective sites.

Within *T. cordata*, the first group (CPW, CW, DB, EH, KP, LP, LU, ShW, and SU) are west Midland sites, close to the Welsh border (Fig. 3 and Online Resource 1). These populations are interesting because, with the exception of CPW, there is very little admixture within each population. This might be related to the regeneration and dispersal ability of *T. cordata* in this area. While the species can produce fertile seeds in this part of England, the conditions of the wood must be favourable for seedlings to grow *i.e.* open canopy, little competition and limited predation (Pigott 2012).

The second group, (RW and SkW), are North West UK sites in relative isolation and there is evidence that RW has been managed as a coppiced wood for 300 years (Cottrell et al. 2003). While this could effectively restrict the potential for sexual recruitment, *Tilia* has frequently set seed and regenerated at this site. Conversely, *Tilia* within SkW, due to ecological conditions, has not reproduced from seed for at least a century (Pigott 2012). Populations RW and SkW might have ancestral genotypes that have been maintained due to isolation and although recruitment occurs at RW, the large number of *T. cordata* found there are continuing to maintain the ancestral genotypes and genetic drift does not have a significant effect on allele frequencies.

The remaining populations (BB, BP, CGW, EaH and HG) do not clearly constitute a single group. Three of these (BP, EaH, and CGW) are further to the east, but still

within the central UK range. They are remnants of the old Royal Rockingham Forest (Peterken and Welch 1975), and are within close proximity of one another. Population HG, which is part of the Bardney Limewoods complex, is also relatively close to the three sites, so genetic similarity is not unexpected. Populations CGW and EaH are adjacent woodlands and would be expected to have very similar genotypes and while F_{ST} between the two are non-significant (Online Resource 7); EaH is admixed with the other two groups. Many large maiden trees once occurred at EaH in the early 20th Century but have been destroyed (Pigott 1991). This may have provided the optimal conditions for young trees to grow and so may explain the difference in genetic composition.

Three interesting sites are clear from the *T. platyphyllos* analysis. Populations AS, BH, and HW represent the three assigned groups with very little admixture. All the other sites are assigned to different clusters. These sites are at the north (AS), west (HW) and central (BH) of the UK *T. platyphyllos* range (Fig. 4 and Online Resource 1).

AS, the first group, is a relatively isolated population compared to the other sites. It is the most northerly native UK population, and is the most north westerly natural European *T. platyphyllos* population (Pigott 2012). The limes are scattered across old coppice woodland on both the upper and lower part of a steep rocky limestone cliff (T. Laurie pers. comm.). Similar genotypes to those in AS were also observed in population CW which is close to the Welsh border. These may be ancestral genotypes that have been maintained within these populations.

BH, the second group, is old managed woodland. Pigott (2000), reported that a single *T. platyphyllos* tree at this site may be over 450 years old, and another over 200 years old. This population is more distinct than other *T. platyphyllos* populations

with more private alleles (Online Resource 6) and the highest F_{ST} values (Online Resource 7), explaining the greatest genetic differentiation among populations. Although there were younger individuals in this population, many of the trees at this site are large and multi-stemmed. Multiple stems from the base of *Tilia* are usually a sign of past coppicing, either naturally or managed. Coppicing was, and still is, an important part of the management of *Tilia* species (Pigott 2012), and may have resulted in the large present day differentiation from other *T. platyphyllos s*ites. Furthermore, BH is relatively isolated, much like AS, and so gene exchange with other populations is expected to be limited. However, sexual recruitment of *Tilia*, although restricted by roe deer, is more prevalent at southern populations *e.g.* BH, where temperatures during flowering can be up to 2°C higher than at northern populations *e.g.* AS (D. Pigott, pers. comm.) thus may be influenced by more optimum ecological conditions.

Population HW, the third group, is large woodland close to the Welsh border. Although documented as having both species and natural hybrids present (Natural England, 2014), all but one sample collected and genotyped from this population were *T. platyphyllos*. Other populations (KP, LP and SU), have a genetic affinity with HW and this is not unexpected due to their close proximity. However, population AW is further north, and while these could be ancestral genotypes, human assistance, due to the planting of the species in parts of its range might be a factor in the genetic similarity of these *T. platyphyllos* populations.

Conservation implications

Although some populations are small and isolated, genetic diversity remains high. However, if these ancient woodland sites further fragment, diversity could eventually

be eroded. Intra-specific structure, possibly from historical fragmentation, has resulted in three genetic groups within each species. At present, the two UK species are considered Least Concern on the IUCN Red List of Threatened Species. However, some sites are regarded as important UK *Tilia* populations, *e.g.* BH (Natural England, 2014) and *Tilia cordata* is currently listed as a high priority species on the SilviFuture Database (SilviFuture 2015). Therefore, efforts to maintain stands should be considered. The loss of ancient woodlands within the UK is a growing concern and many *Tilia* populations are facing further size reductions. In wake of future climate change, having an understanding of the status of the genus will now enable us to inform effective decision making regarding their conservation by concentrating on populations with the highest diversity, and detecting truly endemic genotypes. Conservation efforts to restore small populations should consider introducing genotypes from woodlands that group together. These efforts will ensure that the unique and ecologically important role of *Tilia* remains well established in the landscape.

Conclusion

The study showed, for the first time, the power of microsatellite markers in successfully distinguishing between the two UK *Tilia* species and estimating the genetic diversity and differentiation. Using 13 variable markers, the two species can be confidently defined and distinguished from the hybrid. Although, it was not the focus of this study to explain the extent of hybridisation in *Tilia*, we can confirm that some level of genetic introgression does occur, or at least has occurred in the past.

The genetic data confirm that *T. cordata* and *T. platyphyllos* are outcrossing and diploid. Both species have high genetic diversity. The two species show high genetic differentiation from one another, indicating two distinct biological units.

We have shown evidence of intraspecific structure and inter-population differentiation in both species. Based on the genetic markers used in this study, this structure is loosely related to location. Although no significant isolation by distance occurs, provenance issues and local adaptation should be taken into consideration regarding future conservation or management of certain sites.

Acknowledgements

We would like to thank Dr Paul Ashton for kindly providing *Tilia cordata* samples and Dr Rebecca Roseff and Tim Laurie for their assistance with sample collection and fieldwork. We would also like to thank the National Trust, Forestry Commission, Local Wildlife Trusts and the private landowners who granted permission to collect leaf material from their land. We must also thank Professor Donald Pigott and the anonymous reviewers for reviewing an earlier version of this manuscript and for providing comments and suggestions. This research was funded by the RB Cooke Studentship from Newcastle University (SL) and by the Thai Government (PP).

References

Aldrich PR, Parker GR, Michler CH, Romero-Severson J (2003) Whole-tree silvic identifications and the microsatellite genetic structure of a red oak species complex in an Indiana old-growth forest. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 33:2228-2237 doi:10.1139/x03-160

Arnaud-Haond S et al. (2006) Genetic structure at range edge: low diversity and high inbreeding in Southeast Asian mangrove (*Avicennia marina*) populations. Molecular Ecology 15:3515-3525 doi:10.1111/j.1365-294X.2006.02997.x

Babington CC (1862) Manual of British botany, containing the flowering plants and ferns arranged according to the natural orders. John Van Voorst,

Bacilieri R, Ducousso A, Kremer A (1995) Genetic, morphological, ecological and phenological differentiation between *Quercus petraea* (Matt) Liebl. and *Quercus robur* L. in a mixed stand of northwest of France. Silvae Genetica 44:1-10

Bacilieri R, Ducousso A, Kremer A (1996) Comparison of morphological characters and molecular markers for the analysis of hybridization in sessile and pedunculate oak. Annales Des Sciences Forestieres 53:79-91 doi:10.1051/forest:19960106

Bailey SA, Haines-Young RH, Watkins C (2002) Species presence in fragmented landscapes: modelling of species requirements at the national level. Biological Conservation 108:307-316 doi:10.1016/s0006-3207(02)00119-2

Barrett CF, Freudenstein JV (2009) Patterns of morphological and plastid DNA variation in the *Corallorhiza striata* Species Complex (Orchidaceae). Systematic Botany 34:496-504

Bateman RM, Smith RJ, Fay MF (2008) Morphometric and population genetic analyses elucidate the origin, evolutionary significance and conservation implications of *Orchis xangusticruris* (*O. purpurea* x *O. simia*), a hybrid orchid new to Britain. Botanical Journal of the Linnean Society 157:687-711 doi:10.1111/j.1095-8339.2008.00856.x

Beatty GE, McEvoy PM, Sweeney O, Provan J (2008) Range-edge effects promote clonal growth in peripheral populations of the one-sided wintergreen *Orthilia secunda*. Diversity and Distributions 14:546-555 doi:10.1111/j.1472-4642.2008.00469.x

Caughley G, Grice D, Barker R, Brown B (1988) The edge of the range. Journal of Animal Ecology 57:771-785 doi:10.2307/5092

Cottrell JE, Munro RC, Tabbener HE, Milner AD, Forrest GI, Lowe AJ (2003) Comparison of fine-scale genetic structure using nuclear microsatellites within two British oakwoods differing in population history. Forest Ecology and Management 176:287-303 doi:http://dx.doi.org/10.1016/S0378-1127(02)00289-X

Cottrell JE, Vaughan SP, Connolly T, Sing L, Moodley DJ, Russell K (2009) Contemporary pollen flow, characterization of the maternal ecological neighbourhood and mating patterns in wild cherry (*Prunus avium* L.). Heredity 103:118-128 doi:10.1038/hdy.2009.39

Crawford NG (2010) smogd: software for the measurement of genetic diversity. Mol Ecol Resour 10:556-557 doi:10.1111/j.1755-0998.2009.02801.x

DIVA-G.I.S. DIVA-GIS v7.5. www.diva-gis.org. Accessed May 2014

Duminil J, Caron H, Scotti I, Cazal S-O, Petit RJ (2006) Blind population genetics survey of tropical rainforest trees. Molecular Ecology 15:3505-3513 doi:10.1111/j.1365-294X.2006.03040.x

Duminil J, Di Michele M (2009) Plant species delimitation: A comparison of morphological and molecular markers. Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology 143:528-542 doi:10.1080/11263500902722964

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361 doi:10.1007/s12686-011-9548-7

Eckert C (2001) The loss of sex in clonal plants. Evolutionary Ecology 15:501-520 doi:10.1023/A:1016005519651

Elwes HJ, Henry A (1913) The Trees of Great Britain and Ireland. Privately printed, Edinburgh

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology 14:2611-2620 doi:10.1111/j.1365-294X.2005.02553.x

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564-567 doi:10.1111/j.1755-0998.2010.02847.x

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

Fineschi S, Salvini D, Taurchini D, Carnevale S, Vendramin GG (2003) Chloroplast DNA variation of *Tilia cordata* (Tiliaceae). Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 33:2503-2508 doi:10.1139/x03-179

Fromm M, Hattemer HH (2003) Inheritance of allozymes and hybridization in two European *Tilia* species. Heredity 91:337-344 doi:10.1038/sj.hdy.6800325

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. J Hered 86:485-486

Hamrick JL (2004) Response of forest trees to global environmental changes. Forest Ecology and Management 197:323-335 doi:http://dx.doi.org/10.1016/j.foreco.2004.05.023

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 351:1291-1298 doi:10.1098/rstb.1996.0112

Hermy M, Honnay O, Firbank L, Grashof-Bokdam C, Lawesson JE (1999) An ecological comparison between ancient and other forest plant species of Europe, and the implications for forest conservation. Biological Conservation 91:9-22 doi:http://dx.doi.org/10.1016/S0006-3207(99)00045-2

Hosseinzadeh Colagar A, Yusefi M, Zarei M, Yousefzadeh H (2013) Assessment of genetic diversity of *Tilia rubra* DC. by RAPD analysis in the Hyrcanian forests, north of Iran. Polish Journal of Ecology 61:341 - 348

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322-1332 doi:10.1111/j.1755-0998.2009.02591.x

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801-1806 doi:10.1093/bioinformatics/btm233

Jost L (2008) GST and its relatives do not measure differentiation. Molecular Ecology 17:4015-4026 doi:10.1111/j.1365-294X.2008.03887.x

Kelleher CT, Hodkinson TR, Douglas GC, Kelly DL (2005) Species distinction in Irish populations of *Quercus petraea* and *Q. robur*: morphological versus molecular analyses. Annals of Botany 96:1237-1246 doi:10.1093/aob/mci275

Larcombe MJ, Vaillancourt R, xe, E, Jones RC, Potts BM (2014) Assessing a Bayesian approach for detecting exotic hybrids between plantation and native Eucalypts. International Journal of Forestry Research 2014:13 doi:10.1155/2014/650202

Liesebach H, Sinkó Z (2008) A contribution to the systematics of the genus *Tilia* with respect to some hybrids by RAPD analysis. Dendrobiology 59:13-22

Maurer WD, Tabel U (1995) A methodological study to improve Isozyme analysis for identification of clones of *Tilia* (Linden syn. Lime tree). Silvae Genetica 44:351 - 356

Meimberg H et al. (2010) Evidence for species differentiation within the *Ancistrocladus tectorius* complex (Ancistrocladaceae) in Southeast Asia: a molecular approach. Plant Systematics and Evolution 284:77-98 doi:10.1007/s00606-009-0241-1

Morgan-Richards M, Wolff K (1999) Genetic structure and differentiation of *Plantago major* reveals a pair of sympatric sister species. Molecular Ecology 8:1027-1036 doi:10.1046/j.1365-294x.1999.00665.x

Natural_England (2014). www.gov.uk/government/organisations/natural-england. Accessed May 2014
Neophytou C (2014) Bayesian clustering analyses for genetic assignment and study of hybridization in oaks: effects of asymmetric phylogenies and asymmetric sampling schemes. Tree Genetics & Genomes 10:273-285 doi:10.1007/s11295-013-0680-2
Opdam P, Wascher D (2004) Climate change meets habitat fragmentation: linking landscape and biogeographical scale levels in research and conservation. Biological Conservation 117:285-297 doi:http://dx.doi.org/10.1016/j.biocon.2003.12.008
Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics (Oxford, England) 28:2537-2539
Peterken GF (1974) A method for assessing woodland flora for conservation using indicator species. Biological Conservation 6:239-245 doi:http://dx.doi.org/10.1016/0006-3207(74)90001-9
Peterken GF (1977) Habitat conservation priorities in British and European woodlands. Biological Conservation 11:223-236 doi:http://dx.doi.org/10.1016/0006-3207(77)90006-4
Peterken GF, Welch RC (1975) Bedford Purlieus: its history, ecology and management.
Petit S, Griffiths L, Smart SS, Smith GM, Stuart RC, Wright SM (2004) Effects of area and isolation of woodland patches on herbaceous plant species richness across Great Britain. Landscape Ecology 19:463-471 doi:10.1023/b:land.0000036141.30359.53
Pfeiffer A, Olivieri AM, Morgante M (1997) Identification and characterization of microsatellites in Norway spruce (Picea abies K.). Genome 40:411-419 doi:10.1139/g97-055
Phuekvilai P (2014) Relicts, refugia, and reticulation: a study of population history, hybrids and phylogeny in the long-lived flowering tree genus *Tilia.* Ph.D. Thesis, Newcastle University
Phuekvilai P, Wolff K (2013) Characterization of microsatellite loci in *Tilia platyphyllos* (Malvaceae) and cross-amplification in related species. Applications in Plant Sciences 1:1200386 doi:10.3732/apps.1200386
Pigott CD (1969) The status of *Tilia Cordata* and *T. Platyphyllos* on the Derbyshire limestone. Journal of Ecology 57:491-504
Pigott CD (1975) Natural regeneration of *Tilia cordata* in relation to forest structure in the Forest of Białowieża, Poland. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 270:151-179 doi:10.2307/2417401
Pigott CD (1981) The status, ecology and conservation of *Tilia platyphyllos* in Britain. In: The biological aspects of rare plant conservation. Chichester etc., Wiley. pp 305-317
Pigott CD (1991) *Tilia cordata* MILLER. Journal of Ecology 79:1147-1207 doi:10.2307/2261105
Pigott CD (2000) A probably native and regenerating population of *Tilia platyphyllos* Scop. Bedfordshire Watsonia 23:344-345
Pigott D (2012) Lime-trees and Basswoods. Cambridge University Press,
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959
Rackham O (2003) Ancient Woodland: it's history, vegetation and uses in England., 2nd edn. Castlepoint Press, Dalbeattie, UK
Rackham O (2008) Ancient woodlands: modern threats. New Phytologist 180:571-586 doi:10.1111/j.1469-8137.2008.02579.x
Radoglou K, Dobrowolska D, Spyroglou G, Nicolescu V (2009) A review on the ecology and silviculture of limes (*Tilia cordata* Mill., *Tilia platyphyllos* Scop. and *Tilia tomentosa* Moench.) in Europe. Die Bodenkultur 60:9-19
Raymond M, Rousset F (1995) Genepop (version 1.2) - population genetics software for exact tests and ecumenicism. J Hered 86:248-249
Riutta T, Slade EM, Morecroft MD, Bebber DP, Malhi Y (2014) Living on the edge: quantifying the structure of a fragmented forest landscape in England. Landscape Ecology 29:949-961 doi:10.1007/s10980-014-0025-z

Rosenberg NA (2004) distruct: a program for the graphical display of population structure. Mol Ecol Notes 4:137-138 doi:10.1046/j.1471-8286.2003.00566.x

SilviFuture (2015). http://www.silvifuture.org.uk/about.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? Molecular Ecology 19:3496-3514 doi:10.1111/j.136-294X.2010.04691.x

Streiff R, Labbe T, Bacilieri R, Steinkellner H, Glössl J, Kremer A (1998) Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. Molecular Ecology 7:317-328 doi:10.1046/j.1365-294X.1998.00360.x

Sun GL, Díaz O, Salomon B, von Bothmer R (2001) Genetic diversity and structure in a natural *Elymus caninus* population from Denmark based on microsatellite and isozyme analyses. Plant Systematics and Evolution 227:235-244 doi:10.1007/s006060170051

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology 15:63-72 doi:10.1111/j.1365-294X.2005.02773.x

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535-538 doi:10.1111/j.1471-8286.2004.00684.x

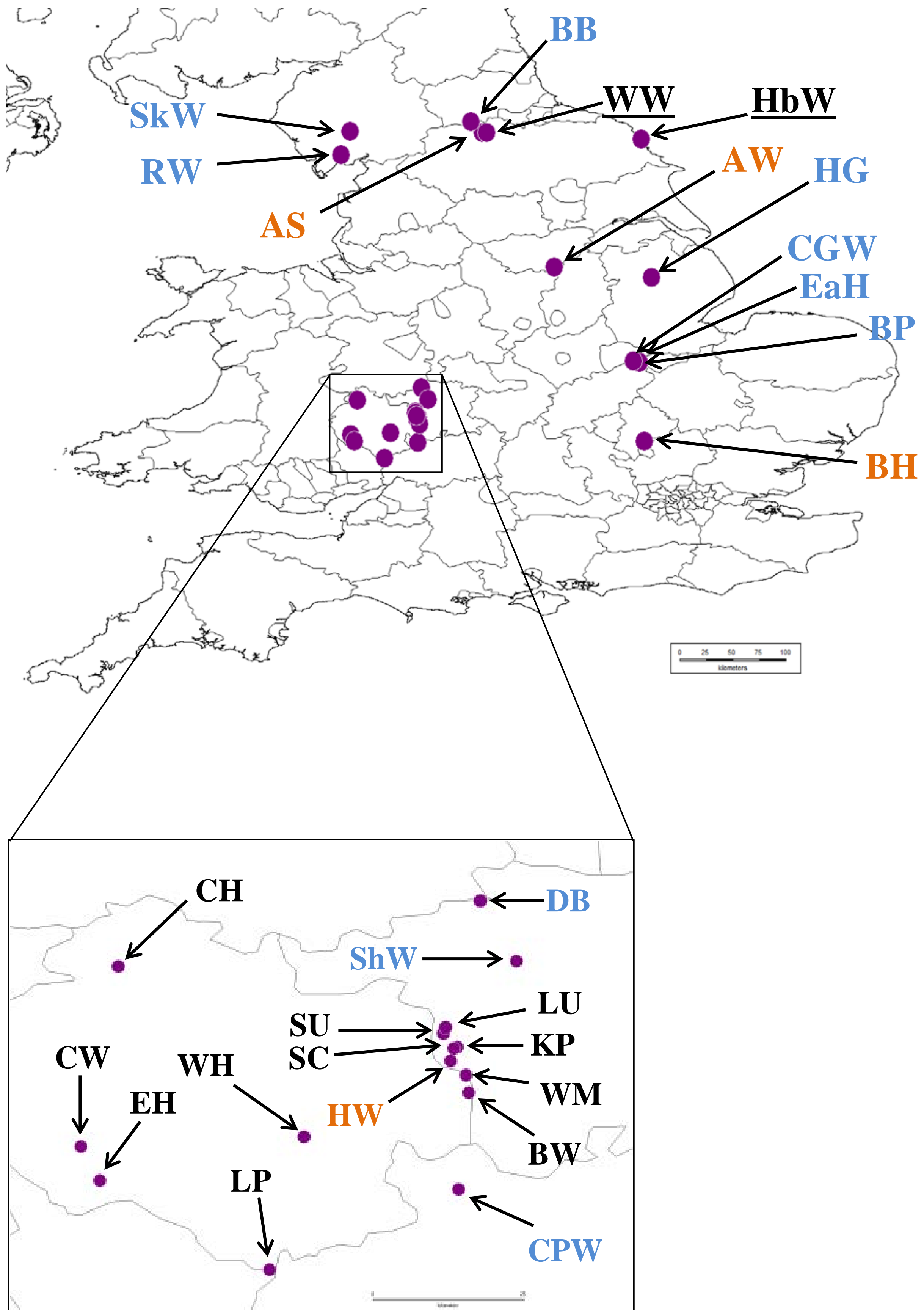
Weir BS, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370 doi:citeulike-article-id:143979

Wesolowski T, Rowinski P (2006) Timing of bud burst and tree-leaf development in a multispecies temperate forest. Forest Ecology and Management 237:387-393 doi:10.1016/j.foreco.2006.09.061

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. Trends in Ecology & Evolution 11:413-418 doi:http://dx.doi.org/10.1016/0169-5347(96)10045-8

Yousefzadeh H, Hosseinzadeh Colagar A, Tabari M, Sattarian A, Assadi M (2012) Utility of ITS region sequence and structure for molecular identification of *Tilia* species from Hyrcanian forests, Iran. Plant Systematics and Evolution 298:947-961 doi:10.1007/s00606-012-0604-x

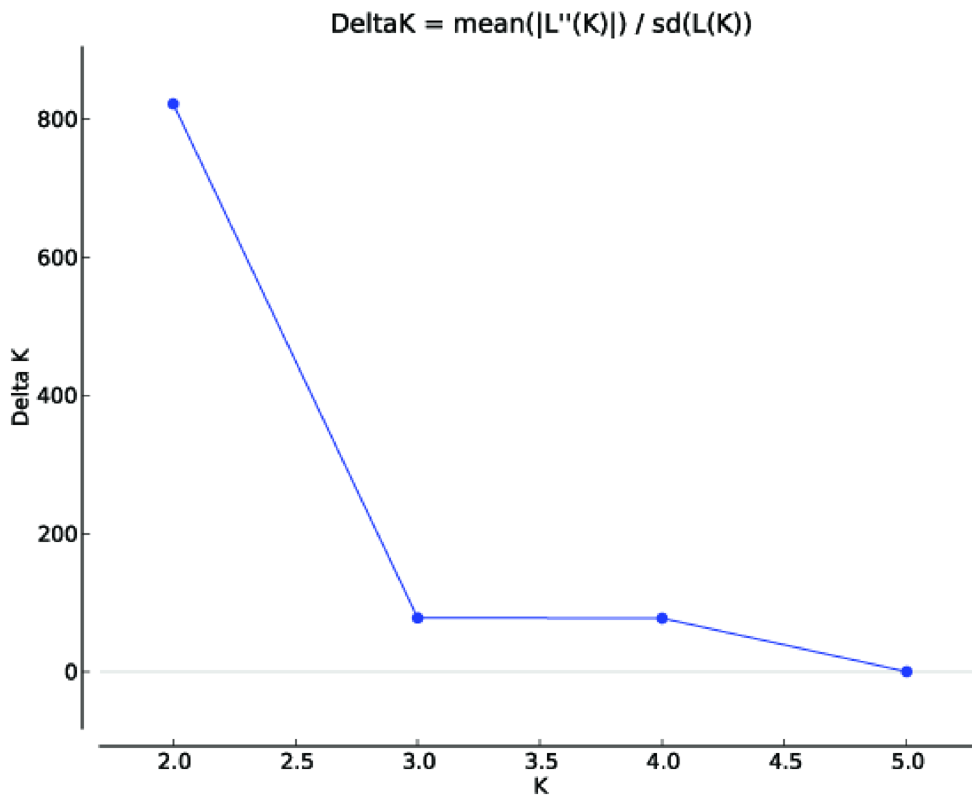
Online Resources



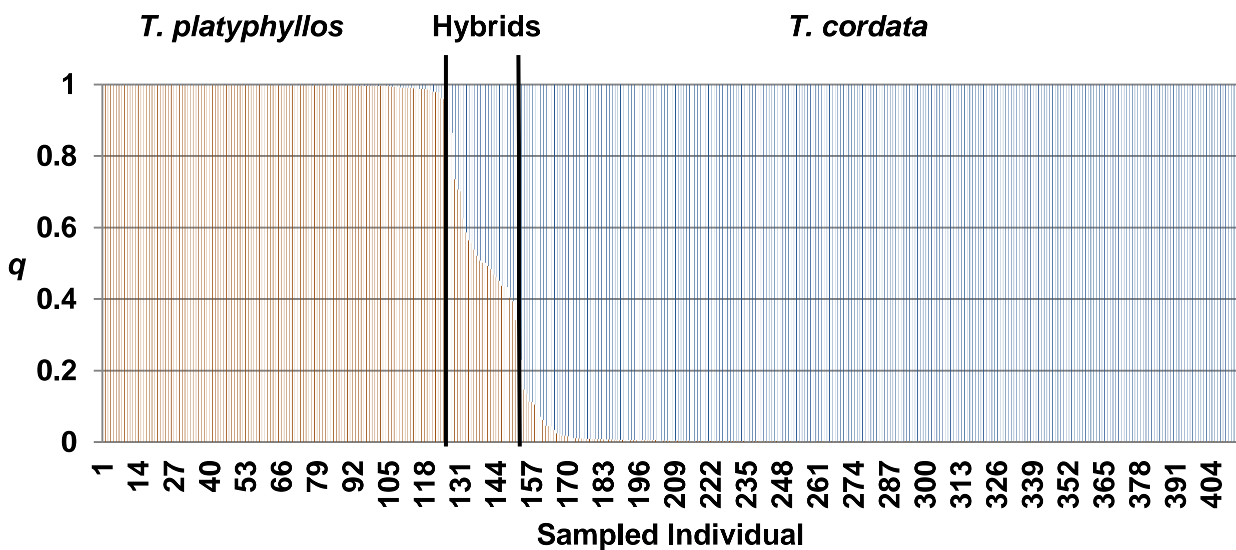
Online Resource 1 Locations of sampled *Tilia* populations, post-genetic analysis. Blue are *T. cordata*, orange are *T. platyphyllos*, black are mixed, and underlined were thought to be *T. platyphyllos* but were *T. x europaea*

Online Resource 2 Results from Micro-Checker showing possible null alleles at each locus in each species.

	Tc6	Tc937	Tc920	Tc8	Tc943	Tc31	Tc4	Tc927	Tc915	Tc963	Tc5	Tc951	Tc7
Pop													
TcLP	-	-	-	-	-	-	null	-	-	-	-	-	-
TcBB	-	-	-	-	-	null	-	-	-	-	-	-	-
TcKP	null	-	-	-	-	-	-	-	-	null	-	-	-
TcCW	-	-	-	-	-	-	-	-	-	-	-	-	-
TcDB	-	-	-	-	-	-	-	-	-	-	-	-	-
TcSkW	-	-	-	-	-	-	-	-	-	-	-	-	-
TcEaH	-	-	-	-	-	-	-	-	-	null	-	-	-
TcBP	-	-	-	-	-	-	-	-	-	-	-	-	-
TcCPW	-	-	-	-	-	-	-	-	-	-	-	-	-
TcHG	-	-	-	-	-	-	-	-	-	null	-	-	-
TcCGW	-	-	-	-	-	-	-	-	-	-	-	-	null
TcShW	-	-	-	-	-	-	-	-	-	-	-	-	-
TcRW	-	-	-	-	-	-	-	-	-	null	-	-	-
TcLu	-	-	-	-	-	-	-	-	-	-	-	-	-
TcSu	-	-	-	-	-	-	-	-	-	-	-	-	-
TcEH	-	-	-	-	-	-	-	-	-	-	-	-	-
TpAS	-	-	-	-	-	-	-	-	-	-	-	-	-
TpBH	-	-	-	-	-	-	-	-	-	-	-	-	-
TpAW	-	-	-	-	-	-	-	-	-	-	-	-	-
TpLP	-	-	-	-	-	-	-	-	-	-	-	-	-
TpHW	-	-	-	-	-	-	-	-	-	null	-	null	-
TpKP	-	-	-	-	-	-	-	-	-	null	-	null	-
TpCW	-	-	-	-	-	-	-	-	-	-	-	-	-
TpSu	-	-	-	-	-	-	-	-	-	null	-	null	-



Online Resource 3 Evanno's ΔK revealing K = 2, implemented in Structure harvester



Online Resource 4 Membership coefficient (*q*) values of all individuals. Individuals (with *q*-value of between 0.20 and 0.80), shown between the two black lines, are considered hybrids, edited in Microsoft Excel 2010

Online Resource 5 Allele fragment size at locus *Tc8*, and *q* – values of 32 (including shaded individuals) and 25 putative hybrid individuals following threshold 0.10 and 0.20, respectively. Cluster 1 is *T. platyphyllos* and cluster 2 is *T. cordata*

Allele size		*q*- value	
Locus *Tc8*	Tree ID	Cluster 1	Cluster 2
156,158	WM03	0.866	0.134
156,158	WM01	0.865	0.135
141,158	CW06	0.735	0.265
141,158	EaH04	0.706	0.294
158,164	HbW02	0.701	0.299
141,158	WW01	0.625	0.375
141,158	LP27	0.586	0.414
141,158	CW_sd2	0.564	0.436
141,158	KP18	0.556	0.444
141,158	CW25	0.538	0.462
141,160	AW11	0.521	0.479
141,158	KP06	0.508	0.492
141,160	AW12	0.501	0.499
141,158	AW13	0.501	0.499
141,156	BB03	0.492	0.508
141,156	BB05	0.483	0.517
141,166	WW03	0.468	0.532
141,158	LP28	0.461	0.539
141,156	BW002	0.451	0.549
141,158	CW13	0.437	0.563
141,160	CW11	0.434	0.566
141,158	CW28	0.433	0.567

Online Resource 5 continued

Allele size		*q*- value	
Locus *Tc8*	Tree ID	Cluster 1	Cluster 2
141,164	HbW04	0.404	0.596
141,162	LP10	0.394	0.606
141,158	AS16	0.341	0.659
141,141	BW001	0.257	0.743
141,141	RW26	0.230	0.770
141,141	BP04	0.147	0.853
141,141	HG01	0.134	0.866
141,141	ShW10	0.113	0.887
141,141	RW37	0.112	0.888
141,141	RW05	0.105	0.895

Online Resource 6 Number of alleles per locus and populations with private alleles at each locus

Locus	***T. cordata***		***T. platyphyllos***	
	No. of alleles	**Pop. Private alleles**	**No. of alleles**	**Pop. Private alleles**
Tc6	8	–	11	TpCW, TpLP
Tc937	6	–	9	TpHW, TpCW
Tc920	11	TcSU	11	TpLP
Tc8	1	–	8	TpBH
Tc943	8	TcEaH	5	–
Tc31	6	–	10	TpHW, TpCW
Tc4	13	TcCPW	13	TpBH, TpKP
Tc927	4	TcEH	15	TpBH, TpHW, TpLP
Tc915	16	TcCPW	18	–
Tc963	26	TcEH	15	TpCW, TpHW
Tc5	13	TcCGW	12	TpBH, TpCW
Tc951	7	TcEaH	9	TpHW
Tc7	8	TcCGW	13	TpAW, TpBH, TpSU
Total	127	9	149	21

Online Resource 7 Diversity measures of 16 T. cordata populations. N – number of samples; N_A – Average number of alleles; A_E – Effective number of alleles; H_O – Observed heterozygosity; H_E – Expected heterozygosity.

Pop	***N***	***N_A***	***A_E***	***H_O***	***H_E***
TcBB	*8*	*3.69*	*2.76*	*0.58*	*0.51*
TcBP	*20*	*5.15*	*3.06*	*0.54*	*0.53*
TcCPW	*8*	*4.23*	*3.01*	*0.53*	*0.49*
TcCW	*13*	*3.69*	*2.67*	*0.50*	*0.49*
TcCGW	*31*	*4.62*	*2.91*	*0.58*	*0.52*
TcDB	*17*	*4.92*	*3.09*	*0.53*	*0.52*
TcEaH	*18*	*5.38*	*3.54*	*0.53*	*0.56*
TcEH	*6*	*3.31*	*2.61*	*0.51*	*0.50*
TcHG	*20*	*5.46*	*3.20*	*0.49*	*0.51*
TcKP	*15*	*4.23*	*2.57*	*0.46*	*0.51*
TcLP	*11*	*4.15*	*2.98*	*0.51*	*0.53*
TcLU	*5*	*2.46*	*1.93*	*0.34*	*0.35*
TcRW	*30*	*5.38*	*3.24*	*0.56*	*0.57*
TcShW	*20*	*5.46*	*3.47*	*0.59*	*0.56*
TcSkW	*14*	*4.31*	*2.79*	*0.52*	*0.48*
TcSU	*10*	*3.38*	*2.73*	*0.52*	*0.45*
Mean	*15.38*	*4.36*	*2.90*	*0.52*	*0.51*

Online Resource 8 Diversity measures of eight T. platyphyllos populations. N – number of samples; N_A – Average number of alleles; A_E – Effective number of alleles; H_O – Observed heterozygosity; H_E – Expected heterozygosity.

Pop	***N***	N_A	A_E	H_O	H_E
TpAS	*8*	*3.69*	*2.68*	*0.73*	*0.60*
TpAW	*14*	*5.08*	*3.36*	*0.69*	*0.66*
TpBH	*14*	*6.31*	*3.80*	*0.78*	*0.71*
TpCW	*15*	*5.62*	*3.27*	*0.75*	*0.66*
TpHW	*26*	*6.69*	*3.91*	*0.72*	*0.71*
TpKP	*11*	*4.85*	*3.38*	*0.67*	*0.67*
TpLP	*15*	*5.62*	*3.24*	*0.70*	*0.66*
TpSU	*13*	*5.77*	*4.00*	*0.66*	*0.73*
Mean	*14.5*	*5.45*	*3.46*	*0.71*	*0.68*

Online Resource 9 Among species population pairwise F_{ST} and their significance.
* - 0.05; ** - 0.01; NS – not significant

pop	TcBB	TcBP	TcCPW	TcCW	TcCGW	TcDB	TcEaH	TcEH	TcHG
TcBB	–	NS	NS	**	**	*	NS	NS	NS
TcBP	0.053	–	NS	**	**	**	*	**	**
TcCPW	0.044	0.018	–	NS	*	NS	NS	NS	NS
TcCW	0.112	0.065	0.046	–	**	**	**	*	**
TcCGW	0.091	0.039	0.054	0.093	–	**	**	**	**
TcDB	0.083	0.070	0.021	0.070	0.101	–	**	**	**
TcEaH	0.026	0.029	0.012	0.048	0.051	0.047	–	**	NS
TcEH	0.173	0.186	0.161	0.176	0.204	0.148	0.156	–	**
TcHG	0.053	0.024	0.026	0.072	0.054	0.095	0.021	0.171	–
TcKP	0.093	0.102	0.072	0.123	0.113	0.045	0.059	0.171	0.104
TcLP	0.038	0.045	-0.009	0.025	0.074	0.031	0.009	0.140	0.039
TcLU	0.146	0.127	0.052	0.077	0.163	0.048	0.073	0.185	0.123
TcRW	0.046	0.083	0.040	0.096	0.113	0.081	0.054	0.167	0.090
TcShW	0.069	0.047	0.024	0.040	0.059	0.035	0.025	0.151	0.061
TcSkW	0.115	0.079	0.067	0.114	0.098	0.108	0.083	0.240	0.083
TcSU	0.106	0.094	0.035	0.054	0.130	0.092	0.066	0.189	0.079
TpAS	**0.387**	**0.374**	**0.396**	**0.403**	**0.397**	**0.399**	**0.364**	**0.378**	**0.397**
TpAW	**0.361**	**0.360**	**0.368**	**0.382**	**0.386**	**0.373**	**0.348**	**0.354**	**0.383**
TpBH	**0.318**	**0.329**	**0.332**	**0.350**	**0.348**	**0.331**	**0.306**	**0.307**	**0.355**
TpCW	**0.351**	**0.351**	**0.359**	**0.374**	**0.375**	**0.352**	**0.335**	**0.339**	**0.373**
TpHW	**0.324**	**0.329**	**0.335**	**0.346**	**0.358**	**0.343**	**0.319**	**0.318**	**0.353**
TpKP	**0.354**	**0.352**	**0.361**	**0.387**	**0.386**	**0.372**	**0.347**	**0.347**	**0.379**
TpLP	**0.340**	**0.347**	**0.358**	**0.375**	**0.380**	**0.362**	**0.340**	**0.338**	**0.375**
TpSU	**0.320**	**0.335**	**0.329**	**0.348**	**0.360**	**0.346**	**0.314**	**0.307**	**0.352**

Online Resource 9 Among species population pairwise F_{ST}. continued

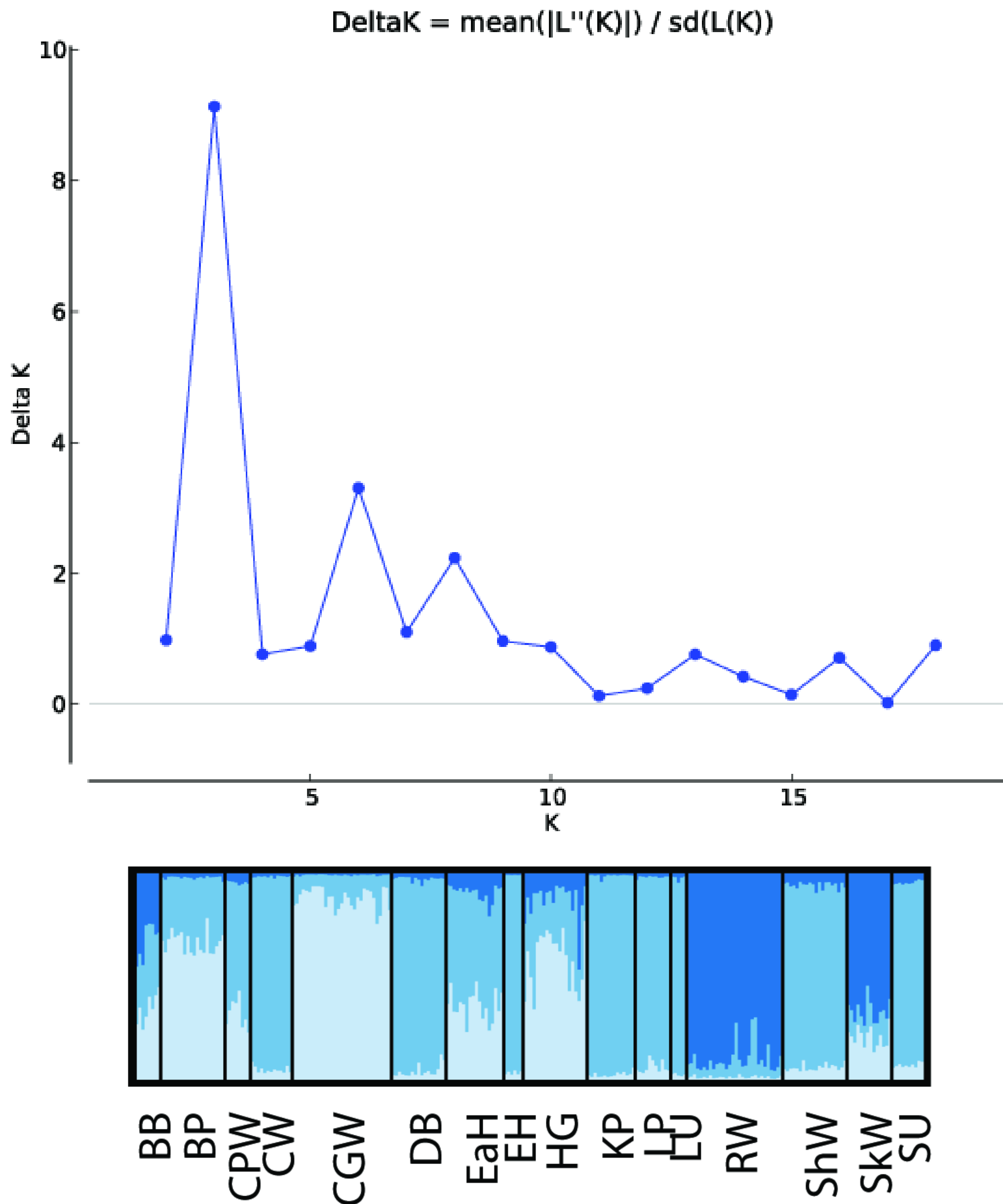
pop	TcKP	TcLP	TcLU	TcRW	TcShW	TcSkW	TcSU
TcBB	**	NS	NS	**	**	**	*
TcBP	**	**	**	**	**	**	**
TcCPW	NS	NS	NS	NS	NS	**	NS
TcCW	**	NS	NS	**	**	**	*
TcCGW	**	**	**	**	**	**	**
TcDB	*	NS	NS	**	**	**	**
TcEaH	**	NS	NS	**	*	**	**
TcEH	**	*	NS	**	**	*	NS
TcHG	**	*	*	**	**	**	**
TcKP	–	NS	NS	**	**	**	**
TcLP	0.048	–	NS	**	NS	**	NS
TcLU	0.099	0.034	–	*	NS	*	NS
TcRW	0.095	0.029	0.101	–	**	**	**
TcShW	0.062	0.030	0.091	0.067	–	**	**
TcSkW	0.131	0.083	0.155	0.090	0.061	–	**
TcSU	0.142	0.051	0.134	0.100	0.074	0.149	–
TpAS	**0.398**	**0.380**	**0.454**	**0.366**	**0.358**	**0.387**	**0.421**
TpAW	**0.374**	**0.356**	**0.414**	**0.347**	**0.342**	**0.380**	**0.403**
TpBH	**0.316**	**0.314**	**0.371**	**0.314**	**0.305**	**0.355**	**0.365**
TpCW	**0.347**	**0.344**	**0.395**	**0.348**	**0.336**	**0.375**	**0.396**
TpHW	**0.338**	**0.322**	**0.375**	**0.309**	**0.316**	**0.344**	**0.357**
TpKP	**0.372**	**0.356**	**0.411**	**0.345**	**0.342**	**0.376**	**0.394**
TpLP	**0.360**	**0.343**	**0.401**	**0.333**	**0.337**	**0.370**	**0.388**
TpSU	**0.340**	**0.318**	**0.371**	**0.313**	**0.316**	**0.350**	**0.364**

Online Resource 9 Among species population pairwise F_{ST}. continued

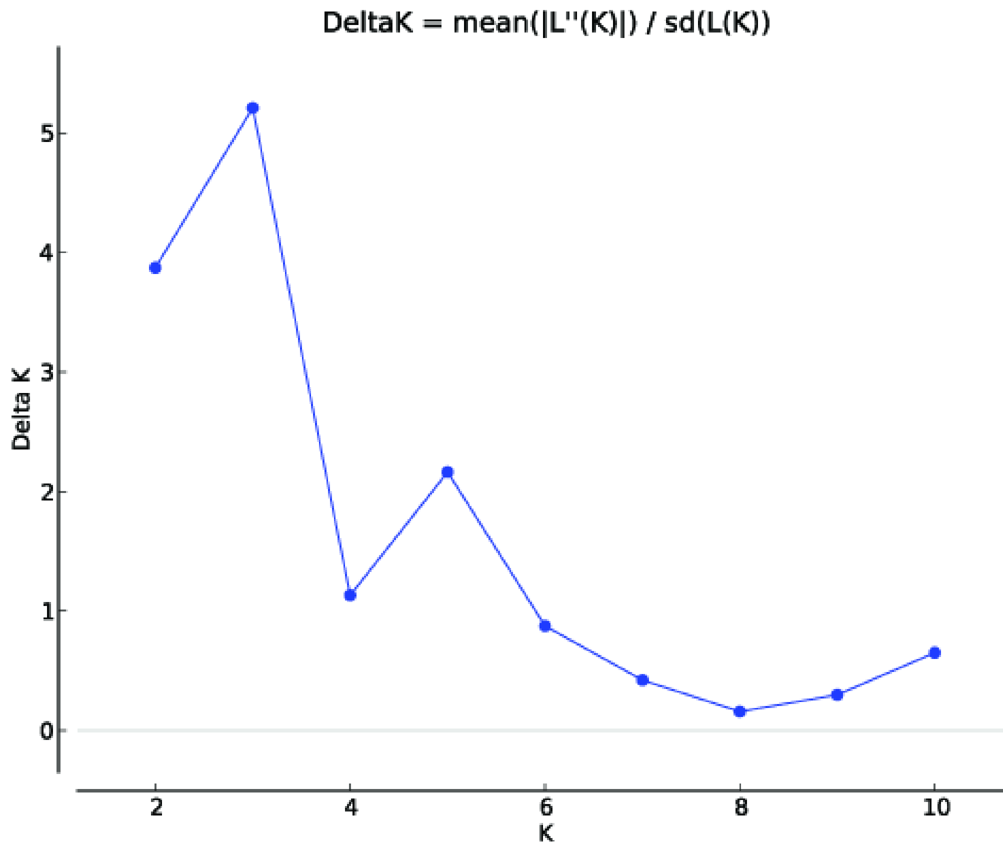
pop	TpAS	TpAW	TpBH	TpCW	TpHW	TpKP	TpLP	TpSU
TcBB	*	**	**	**	**	*	**	*
TcBP	**	**	**	**	**	**	**	**
TcCPW	*	**	**	**	**	*	**	*
TcCW	**	**	**	**	**	**	**	**
TcCGW	**	**	**	**	**	**	**	**
TcDB	**	**	**	**	**	**	**	**
TcEaH	**	**	**	**	**	**	**	**
TcEH	**NS**	**	**	*	**	**NS**	**	**NS**
TcHG	**	**	**	**	**	**	**	**
TcKP	**	**	**	**	**	**	**	**
TcLP	**	**	**	**	**	**	**	**
TcLU	**NS**	*	*	*	**	**NS**	*	**NS**
TcRW	**	**	**	**	**	**	**	**
TcShW	**	**	**	**	**	**	**	**
TcSkW	**	**	**	**	**	**	**	**
TcSU	*	**	**	**	**	**	**	*
TpAS	–	**	**	**	**	*	**	*
TpAW	0.122	–	**	**	**	**	**	**
TpBH	0.196	0.164	–	**	**	**	**	**
TpCW	0.139	0.101	0.129	–	**	**	**	**
TpHW	0.139	0.086	0.118	0.105	–	**	**	**
TpKP	0.106	0.059	0.149	0.079	0.079	–	**	*
TpLP	0.108	0.077	0.136	0.084	0.083	0.046	–	**
TpSU	0.108	0.095	0.119	0.090	0.074	0.075	0.073	–

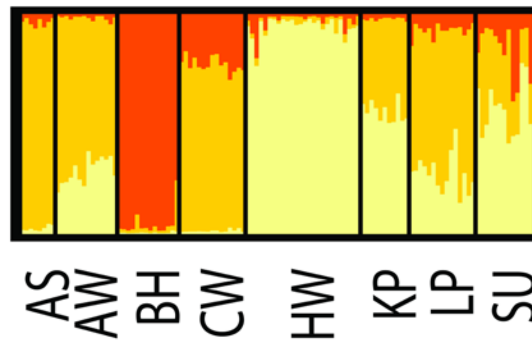
Online Resource 10 Analysis of Molecular Variance (AMOVA) showing the partitioning of genetic variation among species, populations and individuals

Source of variation	d.f.	Sum of Squares	Variance Components	Percentage of Variation	*p* value
Among species	1	862.15	1.77	28.54%	<0.001
Among populations within species	22	450.75	0.43	7.05%	<0.001
Within populations	922	3646.55	3.96	64.41%	<0.001
Total	945	4959.45	6.15		



Online Resource 11 Evanno's ΔK revealing K = 3 in *T. cordata*, implemented in Structure harvester and assignment of 246 individuals from 16 *T. cordata* populations with $K = 3$ inferred by Bayesian clustering analysis implemented in STRUCTURE, averaged in CLUMPP and visualized in DISTRUCT





Online Resource 12 Evanno's ΔK revealing K = 3 in *T. platyphyllos*, implemented in Structure harvester and assignment of 116 individuals from eight *T. platyphyllos* populations when $K = 3$ inferred by Bayesian clustering analysis implemented in STRUCTURE, averaged in CLUMPP and visualized in DISTRUCT