Reproductive and population genetic characteristics of leading-edge and central populations of two temperate forest tree species and implications for range expansion

Running Title: Genetics of leading-edge vs central tree populations

Authors:
Samuel A. Logan¹, Prattana Phuekvilai¹,², Roy Sanderson¹, Kirsten Wolff¹

Address:
¹ Newcastle University, School of Natural and Environmental Sciences, Newcastle NE1 7RU, UK
² Current address: Department of Biology, Faculty of Science, Silpakorn University, Nakhon Pathom 73000, Thailand

Author for correspondence: Dr Kirsten Wolff
Kirsten.wolff@ncl.ac.uk

Newcastle University
School of Natural and Environmental Sciences,
Ridley Building
Newcastle NE1 7RU
UK

Declaration of interest: none
Abstract

It is still a matter of debate how reproductive and genetic characteristics of range-edge populations differ from those central to its range, yet this is important for future changes in species’ ranges. Here we use microsatellite markers to assess the genetic diversity, relatedness and clonal reproduction of two lime tree species, *Tilia cordata* and *T. platyphyllos*, from leading range-edge and central locations.

Clonal reproduction was limited in the populations studied, but leading range-edge populations contained more clones than populations sampled from lower latitudes. Although no inbreeding or lower genetic diversity was detected, leading range-edge populations have a higher average relatedness of individuals and a lower effective population size than those populations closer to the centre of the species’ distribution. Trees further apart than 26m are unlikely to be clones and those further apart than 61m are not likely to be closely related.

The implications for forest managers are that although leading range-edge populations have a lower effective population size, they are also likely to be better adapted to northern climes. Because locally sourced trees appear to have sufficient genetic diversity and predominantly result from sexual reproduction they are likely to adapt to climate change and be suited for natural migration and a good source for assisted migration. This is a promising prospect for the potential of future natural or managed expansion and increased species abundance.

Keywords: Assisted migration; Climate change; Clonal reproduction; Leading range-edge; Microsatellites; Relatedness; Range shifts; *Tilia*
1. Introduction

Range-edge populations exist in less than optimum conditions, and are often smaller and more isolated than central populations. The majority of the studies indeed show that range-edge populations are vulnerable to loss of genetic diversity and increased population differentiation through random genetic drift (e.g. Arnaud-Haond et al., 2006; Eckert et al., 2008; Gapare & Aitken, 2005). However, a few phylogeographic studies have suggested that current genetic diversity has been influenced by historical range expansions and contractions leading to boundary populations in some species that contain a larger proportion of the species diversity compared to the central populations (Hewitt, 2004; Petit et al., 2003). In addition, the demography and mode of reproduction in range-edge populations may also differ from central populations (Dorken & Eckert, 2001; Vucetich & Waite, 2003), potentially also affecting genetic diversity, population differentiation and the ability to adapt.

Asexual reproduction is a survival mechanism common in plants and is achieved through agamospermy or clonal (vegetative) reproduction e.g. root suckering, root collar growth, epicormic shoots, (Bond & Midgley, 2001; Ellstrand & Roose, 1987; Evans & Morris, 2016; Widén et al.,1994). Clonality can be seen as an adaptation to persist in adverse conditions (Macaya-Sanz et al., 2016; Meloni et al., 2013) and offers both advantages and disadvantages. For example, being clonal prolongs the lifespan of an individual which may lead to lower genetic diversity among adult stems (De Witte & Stöcklin, 2010). Also, clonal organisms suspend the exchange and recombination of genetic material, and thus fewer novel genotypes are generated. If sexual reproduction is limited to only a few individuals then asexual reproduction is advantageous as otherwise inbreeding and random allele loss may occur (Ennos, 2003), and genetic diversity is maintained (Balloux et al., 2003). In addition, large clones may accumulate somatic mutations, potentially generating diversity and elevating heterozygosity (Balloux et al., 2003; Halkett et al., 2005). It is also possible selection plays a role, i.e. higher survival of more heterozygous genotypes. However,
adaptation, for example to a changing climate, is limited in largely clonally reproducing populations. On the contrary, sexual reproduction would produce an excess of seedlings for natural selection to act on. Therefore, the differences in relative amounts of sexual and asexual reproduction can change the genetic composition, diversity and adaptability of organisms.

In several plant species, strong clonal reproduction, low levels of genetic diversity and high population differentiation in range-edge populations, as compared to those at the centre of the species' range, have been reported. These studies concentrated often on shrubs and weeds (e.g. Beatty et al., 2008; Dorken & Eckert, 2001), but some significant work has also been carried out on long-lived organisms, such as forest trees (e.g. Jankowska-Wroblewska et al., 2017; Santos-del-Blanco et al., 2013; Stoeckel et al., 2006; Vaughan et al., 2007). Although clonal spread certainly occurs in forest trees, these are generally not considered to be species where clonal propagation is the predominant form of reproduction.

Trees from the genus *Tilia* were important components of woodlands in Europe until a few centuries ago, and the woodlands where they currently exist are largely remnants of their original distribution (Pigott, 2012). As a survival strategy, in sub-optimum conditions (e.g. cooler temperatures resulting in shorter flowering period and/or fewer pollinating insects), members of the genus may reproduce asexually (Evans & Morris, 2016; Logan et al., 2018). *Tilia*, may be considered a partially clonal reproducing species, but the extent of asexual reproduction and intra-population relatedness across their range is not known. Therefore, *Tilia* serves as an excellent forest tree model to investigate leading range-edge effects, adding valuable information to the existing literature (e.g. Myking, 2002; Siefert et al., 2015).

Species within the genus *Tilia* tend to cope well with cyclic management regimes, such as coppicing (Buckley & Mills, 2015; Pigott, 2012). Coppicing encourages sprouting, and therewith prolongs the lifespan and (clonal) spread of a genotype. Clonal spread is also aided by land erosion, particularly on steep banks and cliffs and individual trees can spread through low hanging branches and fallen trunks (S. A. Logan pers. obs.). The ability to
regenerate through vegetative growth and the low rate of regeneration by seed at some
locations, e.g. range-edge (attributed to the present climate), may explain current
distributions (Pigott & Huntley, 1978). Earlier work has shown that populations in the UK
clonally reproduce at some sites (Logan et al., 2015; Mylett, 2015; Phuekvilai, 2014).
Likewise, Novák et al. (2014) suggest that *T. sibirica* (Siberian lime) shows evidence of clonal
growth, which has been confirmed by microsatellite genotyping (Logan et al., 2018). Also,
remnant stands of *T. americana* (American basswood) on islands off the coast of Georgia
(USA) largely depend on vegetative growth (Evans & Morris, 2016).

Radoglou et al. (2009), and references cited therein, suggest that *T. cordata* in south-western Russia has a greater incidence of clonal reproduction than sexual reproduction,
claiming that nearly all *Tilia* in the north-east of its European distribution originated from
vegetative reproduction. They also claim that 90% of *Tilia* trees in the Strict Reserve of
the Białowieża National Park (BNP) sprout from root collars, and that this phenomenon
extends to the range-edge populations in the UK, Finland, and Siberia. These statements,
based on morphology, need genetic confirmation. In contrast, it has been noted that *T.
cordata* regenerates freely from seedlings in the BNP (Pigott, 1975) and in other areas of
Poland (Jaworski et al., 2005). Clonal reproduction was certainly not as high as 90% in
sampled UK populations (Logan et al., 2015), and while *T. cordata* trees from western
Siberia did show some clonal occurrence, lime trees from BNP in Poland appear not to be
clonal but largely a result of sexual reproduction (Logan et al., 2018).

Climate change is altering both the potential and actual distribution, of many animal
and plant species (Gibson et al., 2009; Rehm et al., 2015; Walther et al., 2002). Species may
need to track their climate induced range shifts (Essl et al., 2015) to higher latitudes, where it
is colder - the leading edge, or to lower latitudes, where it is warmer - the trailing edge
(Hampe & Petit, 2005; Hewitt, 2004). Here we concentrate on a comparison between leading
distribution. Due to the rapidity of climate change, the shift of ranges suitable for habitation
and sexual reproduction of the species may not be synchronized with the actual migration of organisms (Hoegh-Guldberg et al., 2008; McLachlan, et al., 2007; Seddon, 2010; Walther et al., 2002). The migration rates of trees, and in particular of *Tilia*, in the past were much slower than migration rates needed now for some organisms to track potential ranges (McLachlan et al., 2005; Myking, 2002), also due to human habitat use and geography. Therefore, in addition to natural migration, assisted migration, *i.e.* transfer of trees or planting of seeds, may be needed (Aitken & Whitlock, 2013; Seddon, 2010; Whittet et al., 2016; Winder et al. 2011).

Trees from populations nearer to their central ecological/geographical area are unlikely to have the adaptation for optimal reproduction and survival at higher latitudes further north, *e.g.* differences in season and day length or winter and late frost resistance (Kreyling et al., 2014; Rehm et al. 2015). Therefore, genetically diverse populations with a diversity of microhabitats outside of the main central range might be key to the survival of some tree species as they are more likely to harbor pre-adaptation to environmental change (Hampe & Petit, 2005; Lobo et al., 2018b). This is particularly important in species such as *Tilia*, where season length is crucial for sexual reproduction. Although there is a view that range-edge populations may not be optimally adapted to their habitat due to persistent gene flow from central populations (*e.g.* Mimura & Aitken, 2007), it is likely that if there is sexual reproduction in a population there will be ample seedlings upon which selection can act. Seed and pollen dispersal in *Tilia* species has not yet been assessed, but is expected to be limited compared to other tree species (Myking, 2002) because *Tilia* is insect pollinated and seeds are heavy and wind-dispersed. Therefore, persistent gene flow from central to range-edge population is unlikely to be the case in *Tilia*, in particular with (currently) fragmented populations. Moreover, there is evidence for adaptation to latitude as spring phenology adaptation has been recorded in *T. cordata* (Lobo et al., 2018a).

Microsatellite markers indicated high levels of polymorphism in *T. cordata* and *T. platyphyllos* (Logan et al., 2015; Phuekvilai & Wolff, 2013) and allowed identification of
clones. Here we investigate clonal architecture, relatedness and genetic diversity at leading range-edge populations in comparison and central populations, using two putative partially clonal *Tilia* tree species. In particular, this study investigates; [1] the level of clonality, [2] the within-population relatedness of individuals and *F*<sub>IS</sub>, [3] genetic diversity and structure, and [4] the effective population size of populations. We expect that in regions with hampered sexual reproduction (northern range-edge), asexual reproduction must be more pronounced and that genetic diversity and effective population size will be lower, while population differentiation and relatedness will be higher than in central populations. The study will indicate whether range-edge populations of *Tilia* are potentially suitable for natural and assisted migration, aiding future management of these woodland species.

### 2. Materials and Methods

#### 2.1 Study sites and sample collection

Two species of temperate, long-lived forest tree, *Tilia cordata* Mill. (small-leaved lime) and *T. platyphyllos* Scop. (large-leaved lime) were studied. These two species are mostly found in European parts of the northern hemisphere and their distribution is limited because of poor seed set in the short and cool summers at the northern range-edge (Pigott & Huntley, 1978). The distribution of *T. platyphyllos* extends north to Sweden and south to Greece and Turkey, while its east-west range extends from Wales to Poland, Slovakia and Romania (Fig 1). *Tilia cordata* has a wider distribution, from Finland and Norway in the north to Greece in the south, while its east-west range extends from Wales and Spain to Siberia, Russia. Both species now mainly exist in native or ‘ancient’ woodlands.

We define leading (northern) range-edge populations as those occurring at >52°N for *T. platyphyllos* and >54°N for *T. cordata* (Table 1, Fig 1). Populations sampled from locations between latitudes 43° - 52°N for *T. platyphyllos* and 43° - 54°N for *T. cordata* are considered central populations. UK samples were from Sites of Special Scientific Interest
and National Nature Reserves. Other collection sites were remnants of ancient woods and not planted, as there is no history of planting *Tilia* in these woodlands. No further details were available about the stands. Only adult trees were included in this study, i.e. trees with a circumference of 10 cm or more. Coordinates were recorded for as many individuals as possible using a Garmin hand-held GPS unit. We calculated geographic distances between trees in GenAlEx v6.502 (Peakall & Smouse, 2006; 2012). Locations could not be systematically sampled with the same distances between trees, due to the terrain and the number and density of trees. Average distances between sampled trees are presented (Table 1). In two populations (SWWC, RUVA) only one coordinate for the population was available and hence mean distances between sampled trees could not be calculated. Also, in POBF locations of nine trees were not available.

### 2.2 DNA extraction and microsatellite genotyping

Leaf samples were dried and stored at -20°C until required for genomic DNA extraction, using the CTAB method (Morgan-Richards & Wolff, 1999). A multiplex Polymerase Chain Reaction (PCR) procedure amplified 12 nuclear microsatellite regions (*Tc6, Tc937, Tc920, Tc8, Tc943, Tc4, Tc927, Tc915, Tc963, Tc5, Tc951 and Tc7*) following Phuekvilai and Wolff (2013). Microsatellites were genotyped using an ABI 3130XL Genetic Analyser, scored using Genemapper (Applied Biosystems) and binned manually.

### 2.3 Clonal structure

Deviations from Hardy-Weinberg equilibrium (HWE) were tested in GENEPOP on the web v4.2 (Raymond & Rousset, 1995; Rousset, 2008) and the presence of null alleles was tested in MICRO-CHECKER v2.2.3 (Van Oosterhout *et al.*, 2004). The combination of all alleles across all 12 markers together is called a Multi Locus Genotype (MLG). Individuals sharing MLG are likely to be clones. However, mutation rates, in particular for microsatellites,
can be high enough for mutations to occur in old clones, making genotypes of members of a clone (ramets) differ from each other for one or more alleles. Although such cases were checked and corrected for genotyping errors this could have contributed to ramets of a single clone being similar but not identical. In such a case ramets might have a MLG that differs by one or few alleles from other ramets of the same clone. In the literature this is sometimes called a Multi Locus Lineage (MLL). Hereafter, when we use the term 'clone' instead of MLL.

It is essential to determine the threshold at which MLGs should be considered to be part of a clone (Aizawa et al., 2017; Meloni et al., 2013). Microsatellite mutation rates and an objective criterion are not available. We used GenoDive 2.0b (Meirmans & Van Tienderen, 2004) to determine the optimal threshold for assigning samples to clones, by determining the allele difference represented by the smallest number of pairs, using the infinite allele model. For this a matrix of genetic distance between all pairs of trees was calculated. Two trees identical for all alleles are clearly clones. Those differing for a small number of alleles could have arisen due to somatic mutation or due to genotyping errors. It is important to distinguish those that have clearly a different MLG (i.e. at multiple loci) from those that could erroneously be considered different. The threshold indicates the maximum distance (alleles different) allowed between two trees for them to still be considered clone mates. The limit to be set is reached by constructing a histogram of the frequency of the allelic differences. Those to the left of the valley (defined as the number of alleles different with the lowest frequency) will be considered clone members and those to the right of the valley will be considered to be different. This analysis had strong power and showed that trees with a difference of maximally three alleles would be assigned to the same clone (Fig A1). Clonal spread was quantified by averaging the distance between trees belonging to the same clone.

Once the clones within each population were determined we used GenClone (Arnaud-Haond & Belkhir, 2007) to provide an unbiased (i.e. not affected by genet size) estimate of genotypic richness \( R = (G-1)/(N-1) \), where \( N \) is the number of samples and \( G \) the number of genotypes. \( R \) is a modified measure of proportion distinguishable genotypes \( Pd \),
Ellstrand & Roose, 1987). The value will be ‘0’ when stands consist of a single clone and ‘1’ when all sampled trees are separate genets (Dorken & Eckert, 2001). To describe clonal heterogeneity, an adapted estimate of the Simpson’s complement index (i.e. $D^*$), independent of sample size (Pielou, 1969), was calculated. The Simpson index ($D$) represents the probability that two randomly sampled plants belong to the same species (Simpson, 1949). The Simpson’s complement index (1-$D$) of diversity is commonly reported in clonal studies as $D^*$ (Arnaud-Haond et al., 2007) and ranges from ‘0’ when all trees within a population are a single clone to ‘1’ when they are all unique. Additionally, the statistics $P_{gen}$, the probability of trees having the same genotype (i.e. are part of same clone) by chance following the methods of Parks and Werth (1993) and $P_{sex}$, the probability that a clonal genotype originates from sexual reproduction at the first reencounter, were estimated. $P_{gen}$ assumes populations are in HWE. An adjusted measure $P_{gen}(f)$, taking into account HWE departure can be estimated, providing a more conservative estimate of $P_{sex}$ (Arnaud-Haond et al., 2007). These last two estimates indicate the power of the markers to detect clones.

2.4 Genetic diversity and differentiation

In the population genetic analyses we used single representatives of each clone, i.e. at genet level, reflecting genetic characteristics comparable between populations. Genetic diversity, calculated as the expected heterozygosity ($H_e$), and the fixation index $F_{IS}$ were calculated in GenAlEx 6.502. When testing the difference in $F_{IS}$ between non-clonal and clonal genotypes the fixation index $F_{IS}$ was calculated by hand as $F = (H_e - H_o)/H_e$, with $H_o$ being the expected heterozygosity over all clonal genotypes in that population. AMOVA analyses were performed to quantify population differentiation of the leading range-edge and of central populations of each species in GenoDive 2.0b using the Weir and Cockerham (1984) method. In addition, GenoDive supplied $G$-statistics $G_{ST}$ (Nei, 1987), adjusted $G'_{ST}$ (correcting for the bias due to small numbers of populations) as well as Jost’s $D_{est}$ (Jost, 2008), which is independent of $H_e$. It also supplied standard deviations through jack-knifing
over loci and 95% confidence intervals (CI) through bootstrapping over loci. These statistics were tested in GenoDive (Compare Groups) for significance of difference between leading range-edge and central populations, using 9999 permutations. We tested Isolation by Distance (IBD) to determine if genetic distance between populations was correlated to geographic distance by performing a Mantel Test in GenAlEx.

2.5 Relatedness and effective population size

Relatedness and effective population size were estimated at both genet (with single representative per clone) and ramet level (including all samples), where ramets reflect the actual standing crop. ML-Relate (Kalinowski et al., 2006) uses simulations to determine the most likely relationship (unrelated, half sib, full sib or parent-offspring) and compares this with the likelihood of other relationships for the same pair of trees. We used 100,000 randomisations and counted for each population the number of pairs that have the highest probability of being full sib (50% related), parent-offspring (50% related) and half sibs (25% related). We calculated the number of first degree related pairs as well as the sum of first and second degree related pairs (50% relatedness and 50%+25% relatedness, respectively) as a proportion of the maximum number of pairs within a population ($n \times (n-1)/2$, with $n$ being the number of genotypes or ramets). To obtain a quantitative measure of relatedness we averaged the maximum likelihood estimates of relatedness ($r$, ML-Relate) for all pairs of trees in a population. The distance for each pair of genets within a population with a 50% relatedness was averaged to indicate the distance at which two trees can be closely related.

To detect small scale population structure, i.e. whether trees that were growing near to one another within a population were genetically more closely related than those at larger distances, the autocorrelation coefficient was calculated in GenAlEx. We tested whether the autocorrelation $r$ as a function of the pairwise distance between trees is significantly different from zero and determined the 95% confidence intervals and error bars using 9999 permutations and 9999 bootstraps. In each case we used eight even pairwise distance
classes, with sufficient pairwise comparisons per distance class (maximum sample distance in a population divided by eight).

We estimated the contemporary (or recent) effective population size \((N_e)\) using the molecular co-ancestry method of Nomura (2008) and the Linkage Disequilibrium (LD) method as implemented in NeEstimator V2.01 (Do et al., 2014). The LD method is based on Waples and Do (2008) and includes the Waples (2006) bias correction, which corrects the bias that could be introduced when sample size is lower than the actual effective population size. The 95% CIs are based on Jack-knifing over loci, and in the LD method only alleles with a frequency > 0.05 were used.

2.6 Contrasting groups of populations

For testing the fixed-effect categorical variables of region, species and their interaction on clonal reproduction, genetic diversity and relatedness estimates we used averages per population. For most tests we used the average distance between trees (Av Dist) as a covariate to ensure that a difference in a measure was not biased by differences of within population sampling distances between the contrasting groups of populations. For the two populations without individual coordinates we replaced the missing Av Dist with the average for the group. The sample size (number of populations) was 28 for all tests. Three types of statistical analyses were carried out in R (R Core Development Team, 2016). For variables bounded between 0 and 1 \((H_e, r, \text{ML-Relate}, 1^{\text{st}} \text{ and } 2^{\text{nd}} \text{ degree relatedness, } D^* \text{ values})\) beta regression was performed (Ferrari & Cribari-Neto, 2004; betareg R package: Cribari-Neto & Zeileis, 2010). For variables from 0 to infinity \((N_e \text{ values})\) \(\log(n+1)\) transformation of the values was used, with infinity set to 999. An infinite \(N_e\) estimate is likely caused by an actually large number, but the data do not allow an exact estimate (19 out of 60 estimates).

The \(N_e\) data were analysed with a GLM model with Gaussian (normal) errors. For variables bounded by -1 and +1 \((F_{is})\) we used a logistic quantile regression (Bottai, et al., 2010; lqr R package Galarza, et al., 2016), and we could not use Av Dist as a covariate in this test.
Genetic diversity ($H_e$) at the genet level were compared using a Welch’s (unequal variance) $t$-test in R.

In addition, differences between range-edge and central populations for the degree of clonal reproduction ($D^*$), $H_e$, $F_{IS}$ and population differentiation, were tested using the procedure ‘compare groups’ in GenoDive 2.0b using 9999 permutations with randomisation of populations over the groups. The difference in $H_e$ and $F_{IS}$ between non-clonal and clonal genotypes in populations was tested using a paired $t$-test in R.

3. Results

Genetic diversity was medium to high in both species for all populations. In *T. platyphyllos* the average number of alleles for the 12 microsatellite loci across all samples was 19.7, with a range of 8-27 across loci. Within *T. platyphyllos* populations this was an average of 7.6 alleles per locus (range 3.1-9.8). For *T. cordata* the average number of alleles for the 12 microsatellite loci across all samples was 13.7, with a range of 2-35 across loci. Within *T. cordata* populations this was an average of 5.8 alleles per locus (range 1.1-14.3). Deviations from HWE were observed at some loci, but were not consistent across populations. No null alleles were detected in the loci included in the analyses.

3.1 Clone identification

In total 229 *T. platyphyllos* and 376 *T. cordata* trees were genotyped using microsatellite markers and clones were identified. Using a difference of three alleles as a maximum threshold value, ten trees with a unique MLG were deemed to be part of a clone. In all cases it produced clone members located closely together (average 5m for *T. platyphyllos* and 22m for *T. cordata*). The genotypes of members of a clone that did not have the exact same genotype, but differed for fewer alleles than the threshold value, were compared. Locus Tc963 was the most frequently deviating locus between clone-members, and therefore
seemed to accumulate mutations for members of a single clone (10 out of 17 presumed
mutations). A total of 116 trees were not unique genotypes, i.e. belonged to a clone (19% of
samples). Out of the total of 531 genotypes analysed, 489 were unique and 42 were
represented by more than one tree (7.9%). Eight of the 12 populations of *T. platyphyllos* and
six of the 16 *T. cordata* populations showed multiple adult trees belonging to a clone. No
clones were found between locations. None of the clones were dominating a population, i.e.
most clones (86%) were presented by two or three trees each (Table A2).

The number of loci screened was ample to identify MLG according to the resampling
procedure in GenClone (data not presented, Arnaud-Haond et al., 2006). The probability of
genotypes being identical by chance in a randomly mating population (*P*<sub>gen</sub>) as well as the
probability that identical genotypes arose through sexual reproduction (*P*<sub>sex</sub>) were <0.001 at
all locations (Table S1), confirming that repeated genotypes were not due to chance and are
in fact clones. Genotypic richness (*R*) was high and rather variable across populations,
ranging from 0.278 to 1.00 (Table A1). Clonal heterogeneity or Simpson's index (*D*<sup>*</sup>) was
high in most populations, ranging from 0.743 to 1 (Table A1), suggesting a high probability
that two randomly chosen trees were genetically different and a low rate of clonal
reproduction.

The heterogeneity in clonal reproduction across populations was largely explained by
‘region’, with range-edge populations having higher clonal reproduction (for *R*: *P* <0.0005 and
for *D*<sup>*</sup>: *P* <0.0005) (Table 2). The average distance between trees, used as covariate in the
analysis, was not significant for both measurements (*P* = 0.053 and 0.099, respectively).
Species and region x species interaction effects were not significant (*P* >0.05). Distances
between trees belonging to a single clone was variable, on average 16.8m and top quintile
26m, albeit smaller in *T. platyphyllos* than in *T. cordata* (average 5.7 and 19.3m, respectively,
Fig. 2a and Table A3).

3.2 Genetic diversity and fixation index
Genetic diversity is represented by the expected heterozygosity ($H_e$) and calculated at the genet level (Table A4). $H_e$ was significantly higher in *T. platyphyllos* than in *T. cordata* ($P < 0.0005$, Table 3). There was no effect of region, Av Dist or interaction on genetic diversity. Other measures of genetic diversity, such as effective number of alleles per locus, led to the same conclusions and are not presented here. Values of $F_{IS}$ (fixation index) were variable across populations and reflect the outcrossing breeding system, ranging from -0.284 to 0.105. However, there were no significant effects of species or region (species $P = 0.729$, region $P = 0.545$, Table 3, Table A4).

To test whether large clones have accumulated mutations and have the potential to contribute to the maintenance of genetic diversity we contrasted genotypes that consist of single trees with those consisting of more than one tree (*i.e.* non-clonal vs clonal). This was done in the six *T. cordata* populations that had >2 genotypes of both types. In all populations the observed heterozygosity ($H_o$) was higher and the fixation index ($F_{IS}$) lower in clonal than in non-clonal genotypes (Table 4). The difference in $H_o$ is significant, both using a paired t-test ($t = -3.73, P = 0.014$) and using permutations in comparisons of groups in GenoDive ($P = 0.013$). The difference in $F_{IS}$ between clonal and non-clonal genotypes is significant in a paired t-test ($t = 3.60, P = 0.015$), but is not significant using group comparison in GenoDive ($P = 0.113$).

### 3.3 Population differentiation: contrast range-edge to central populations

$F_{ST}$ values were slightly higher in *T. platyphyllos* as a whole than in *T. cordata*, with 9% and 7% of the variation between populations ($F_{ST}$ 0.090 and 0.070 and $F'_{ST}$ 0.369 and 0.180, respectively). Within each species the differentiation between populations was higher in the range-edge group than in the central group (Table 5 and 6). $G_{ST}$, $G'_{ST}$ and $D_{est}$ were significantly higher for range-edge than central populations in *T. platyphyllos* ($P$-values 0.014, 0.013 and 0.013, respectively), while this was significant for $G'_{ST}$ and $D_{est}$ in *T. cordata* ($P$-values 0.050, 0.040 and 0.003, respectively. The 95% confidence interval (CI) of $F_{ST}$ for
range-edge and for central was not overlapping for *T. platyphyllos*, with 11.2% and 5.5% of variation between range-edge and central populations, respectively, while in *T. cordata* this is 9.1% and 5.6%, respectively, with a small overlap of the 95% CI. A Mantel Test revealed slight positive correlation between genetic and geographic distances within all four groups: *T. platyphyllos* range-edge ($R^2 = 0.208 \ P < 0.000$) and central ($R^2 = 0.0811 \ P < 0.000$); *T. cordata* range-edge ($R^2 = 0.0571 \ P < 0.000$) and central ($R^2 = 0.0101 \ P = 0.001$).

### 3.4 Relatedness

The range-edge populations had a higher relatedness than populations from central regions, expressed as $r$ (ML-Relate), but only at the ramet level ($P < 0.0005$), and the Av Dist effect was also significant at the ramet level ($P = 0.040$) (Table 7 and 8). $R$ (ML-Relate) was not significantly different between the species ($P = 0.079$ genotype level and $P = 0.279$ ramet level).

To test the reliability of indicating the most likely relationships in ML-Relate we combined trees from different locations, expecting no relationships. We found some half-sib relationships across locations, albeit in much smaller numbers than within locations, and no parent-offspring or full sib relationships between locations. The fraction of first degree related pairs ranged from 0 to 0.056 at genet level and 0 to 0.038 at ramet level, while the fraction of first plus second degree relatedness ranged from 0 to 0.097 at genet level and 0.018 to 0.086 at ramet level across populations (Table A4 and A5). Contrasting range-edge and central populations showed more related pairs in range-edge populations, only for the proportion of first degree relationships ($P < 0.003$, $P < 0.045$, at genotype and ramet, respectively). There was no significant difference between species in fraction of first or first plus second degree relatedness. The co-variate Av Dist and the range x species interaction effects were not significant either. The average distance of parent-offspring and full sib pairs
was 52m across populations, with top quintile 61m (average 79m for *T. platyphyllos*, 34.4m for *T. cordata*, Figure 2b, Table A3).

35 Small scale population structure

The correlation of the autocorrelation coefficient \( r \) and geographic distance was tested at the genotype level so that the results were not biased due to ramets of the same clone being close together and having a relatedness of 1. Sixty percent of the populations (12 out of 20) for which we had individual tree locations and enough clones showed a positive correlation of \( r \) with geographic distance in the first distance class (average 40m). Those populations that were collected with large distances between trees did not show autocorrelation in their first distance class.

36 Effective population size

Although the estimates for effective population size \( N_e \) differ between the estimation methods, coalescent and LD, and 95% CI are large, the values are likely to be comparable within this study. There was an effect of species for three of the four estimates \( (P = 0.048 \text{ and } P = 0.042 \text{ for the ramet level and } P = 0.009 \text{ at the genet level, for } N_e) \). In all cases the \( N_e \) of *T. platyphyllos* was smaller than of *T. cordata*. Also, range-edge populations had lower \( N_e \) values than central populations \( (P = 0.005 \text{ and } P < 0.001 \text{ at the ramet level and } P = 0.049 \text{ at the genotype level for } N_e \text{ coalescent and LD, respectively, Tables A2 and A4}). The Av Dist or region x species interaction were not significant for any \( N_e \) estimate.

4. Discussion

Our study shows that leading range-edge populations of two *Tilia* species have more clonal reproduction than populations from central areas of their distribution. From *F*-statistics we
can conclude that range-edge populations are not likely to be affected by gene flow from other populations. This is largely concordant with several previous studies showing that range-edge populations, as compared to central range populations, have higher population differentiation ($F_{ST}$) because of fragmentation and lower long distance gene flow, potentially enabling adaptation to current local conditions (Arnaud-Haond et al., 2006; Eckert et al., 2008) or due to repeated extinction-founder effects. However, we did not find a significantly lower genetic diversity in range-edge populations, as some studies have indicated for other species (Arnaud-Haond et al., 2006). Instead, *Tilia* range-edge populations have medium to high neutral genetic diversity (comparable to other trees species). However, provenance trials would be required to confirm that these range-edge populations also contain sufficient adaptive variation. With a warming climate, gene flow among populations may promote adaptation to novel conditions at northern latitudes.

### 4.1 Clonality in *Tilia*

Clonal occurrence varies greatly across plant species (Ellstrand & Roose, 1987) and populations (Dorken & Eckert, 2001). Genotypic richness ($R$) quantifies sexual versus asexual reproduction (Silvertown, 2008) and can be broadly compared across taxa. Clonality in *Tilia* at both the range-edge and central populations was generally low compared to other forest trees *e.g.* *Populus* species, *Prunus avium*, *P. sspiori*, *Ulmus minor*, and *Sorbus torminalis* (Fuentes-Utrilla et al., 2014; Mock et al., 2008; Nagamitsu et al., 2004; Rasmussen & Kollmann, 2008; Santos-del-Blanco et al, 2013; Stoeckel et al., 2006; Vaughan et al., 2007). This is significant in that even in less optimial conditions, *Tilia* may still mostly reproduce sexually. Therefore, clonality was lower than expected from existing literature, which was based on observations and morphology (*e.g.* Radoglou et al., 2009).

As expected, we found more clonal reproduction in range-edge than in central populations. High clonality was expected at Scandinavian, Russian and northerly UK sites due to range-edge effects and limited sexual reproduction (Pigott, 1981; Pigott & Huntley, 2008).
1981; Radoglou et al., 2009), while clonality at central range sites, was expected to be lower
due to greater sexual regeneration and a greater viability of seed (Pigott, 1981). Various
additional factors can be involved in this. For example, *T. platyphyllos* trees at UK sites were
large and old and with age comes more opportunity for clonal reproduction. In some
populations there is also signs of past coppicing or self-coppicing due to steep terrain (e.g.
UKBB). Other mechanisms, such as freezing/thawing, especially at northerly locations, or
grazing can also have played a role in the tendency for sprouting (Morris et al., 2014; Sjölund & Jump, 2014; Wilmking et al., 2017).

Production of fertile seeds and regeneration seems to be a relatively recent
phenomenon in UK *T. platyphyllos* (Pigott, 2000; 2012). Although references in Pigott (1975)
state that fertile fruit are produced in Finland during very warm summers, our study suggests
that asexual reproduction is important for maintaining *Tilia* in Finland where it competes with
*Picea abies* (Pigott, 2012). Irregular sexual reproduction (of randomly mating individuals)
intermittent to clonal reproduction permits the preservation of diversity and that diversity
reflects the last time sexual reproduction occurred. Ennos (2003), reported that aspen
showed high clonal occurrence and limited flowering but had similar genetic diversity to other
outcrossing woody plants and attributes this survival to asexual reproduction of a previously
random mating population which flowered regularly when conditions were more suitable. So
essentially, genetic diversity was ‘frozen in time’ (Ennos, 2003).

We expected that other northern-most populations may have a similar ecology and
demography to *Tilia* in Finland. However, we found no clones in the Norwegian population,
but this is likely because samples were collected with large distances between trees (Table
1) and so additional sampling from a smaller area is needed for a better understanding of the
reproductive strategy in Norway. Climatic influences, topography and past management may
also play a role in the clonal development of *Tilia* in Russia. Similar to *T. cordata* in our study,
clonal occurrence has also been observed in *T. sibirica* (Siberian lime), further east in
southern Siberia (Logan et al., 2018; Novák et al., 2014). Although, unlike many of the *Tilia*
populations sampled in this present study, *T. sibirica* showed low levels of genetic diversity (Mean $H_e = 0.318$) and high levels of clonality (Mean $R = 0.601$). Logan *et al.* (2018) attributed this to population fragmentation and difficulties for regeneration of seedlings due to competition from tall forbs, except in recent canopy gaps where *Tilia* seedlings may emerge (Novák *et al.*, 2014).

The BNP in Poland has had little human management (Miścicki, 2012). A recent study found no clones in the adult trees sampled from three plots within the Special Reserve (Logan *et al.*, 2018), and repeating the analyses here no clones were found. This contradicts Radoglou *et al.* (2009) who reported that *T. cordata* populations in Poland are highly clonal. In contrast rather, there were young trees present in the BNP (S. A. Logan, pers. obs.), aggregated in tight clusters, probably due to establishment of sexually produced seedlings in response to the creation of a recent gap in the canopy (Bobiec, 2007). The young trees from the BNP were not clones of any of the adult trees (Logan, 2016), confirming observations by Pigott (1975) and in other forests of southern Poland by Jaworski *et al.* (2005) that sexual reproduction is taking place there.

The concept that somatic mutations (or genotyping errors) contribute to a clone being composed of similar MLGs is accepted in other studies (*e.g.* James & McDougall, 2014). In an extreme case in *Robinia pseudoacacia* a ‘hypervariable’ locus was detected (Lian *et al.*, 2004). In *Tilia*, the locus that seemed most mutable (Tc963) also has the largest number of alleles, namely 41. In addition, our finding that genets that are clonal had a higher observed heterozygosity than non-clonal genets (Table 4) follows expectations that large and long lived clones accumulate somatic mutations (Balloux *et al.*, 2003) and therewith potentially contribute to maintenance and generation of genetic diversity. It would be worth repeating this analysis with populations that have larger sample sizes. However, it may also be that trees with higher heterozygosity live longer and/or are more clonal than trees with lower heterozygosity (*e.g.* Vrankx *et al.*, 2014).
While clonal reproduction is an important survival tool for *Tilia* in several parts of its European range, sexual reproduction is still the most important mode of reproduction. This is particularly advantageous for range-edge populations: as the summers become warmer more viable seed will be produced and the exchange and recombination of genetic material will produce genotypes to allow for adaptation to warmer climates.

4.2 Genetic diversity and differentiation

Genetic diversity (*H_e*) in both *Tilia* species was similar to other partially clonal trees (Santos-del-Blanco *et al.*, 2013; Stoeckel *et al.*, 2006; Vaughan *et al.*, 2007). Range-edge populations do not have a significantly lower genetic diversity than those at central locations. This may be because range-edge populations are relicts of a diverse migration front that, through clonal reproduction, have maintained diversity or even increased genetic diversity through somatic mutations (Balloux *et al.*, 2003; Silvertown, 2008). *H_e* was higher in *T. platyphyllos* than in *T. cordata*, which could potentially be caused by ascertainment bias since the markers used were derived from *T. platyphyllos*. However, preliminary evidence using markers derived from *T. cordata* (Mylett, 2015) also revealed lower genetic diversity in *T. cordata* than in *T. platyphyllos* (Stephenson, Logan & Wolff unpubl.).

The near zero fixation index (*F_{IS}*), reflects the outcrossing breeding system of the two species. There was no significant difference in *F_{IS}* values between range-edge and central populations, similar to *Acer* (Chybicki *et al.*, 2014). Our low *F_{IS}* values are in contrast to Ennos (2003), where he reported high *F_{IS}* values in clonal Aspen, likely due to a shortage of flowering mates. While we have no knowledge of the flowering status of the populations in our study, the *F_{IS}* values we report suggest that this is not the case in *Tilia*. The overall slightly negative *F_{IS}* values in the species and the slightly (but not significantly) lower *F_{IS}* in clonal range-edge populations may well be explained by somatic mutations (Balloux *et al.*, 2003; Meloni *et al.*, 2013).
Assuming ideal populations the overall $F_{ST}$ values seem low compared to some other studies and indicate migration rates (Nm values) that cannot be ignored. However, taking into account the long generation time and rare sexual reproduction, migration will have a small effect in these *Tilia* species. Similar to other studies (*e.g.* Arnaud-Haond *et al.*, 2006; Chybicki *et al.*, 2014) we confirmed here that range-edge populations show higher divergence from each other than central populations. It is likely that random genetic drift due to smaller effective population size and higher fragmentation is causing this. This may promote adaptation to current local conditions making them important for range expansion (Arnaud-Haond *et al.*, 2006). However, the difference in $F_{ST}$ between range-edge and central populations could be an artefact, namely because range-edge populations are geographically further apart from each other than central populations. The weak but positive relationship between genetic and geographic distance (see section 3.3) hints at this.

4.3 Relatedness within populations and effective population size

Range-edge populations have a higher relatedness and a smaller effective population size than those from central areas at ramet, and partially, at genet level. Of course, these parameters are not independent and all contribute to low genetic connectivity between edge populations and indicate the fact that trees in range-edge populations are more closely related and fewer genotypes contribute to the next generation in range-edge population than in central populations. This could be because there is less turnover in range-edge populations; genets may be very old and contributed proportionally more to the standing crop than younger trees, especially if there is a lack of sexual reproduction. In central populations there may be more turnover and more even aged genets, with more equal contribution to the next generation. Also, population fragmentation is likely to be higher at the range-edge, also lowering $N_e$ (Chybicki *et al.*, 2014).

The average distance between sampled trees was not identical in all populations and could potentially bias outcomes if, for example, pairwise distances between trees from range-
edge populations were smaller than those from central populations. However, Av Dist only
significantly affected relatedness (r ML-Relate), and only at the ramet level: this effect is as
expected as the average distance between first degree related trees is 52m. Since average
distance of trees is not significant as a covariate in other measures of clonal reproduction,
diversity and relatedness, we are confident that the main-effect differences between range-
edge and central populations are not biased due to sampling scenarios differing between
sites. Moreover, the number of ramet pairs within 26m (excluding the top-quintile for clonal
pairs) across all populations was similar in range-edge and in central populations in both
species (182 vs 165 in T. platyphyllos and 594 vs 395 in T. cordata). The number of genet
pairs within 61m (excluding the top-quintile for first degree related pairs) was also similar in
range edge and in central populations (272 vs 310 in T. platyphyllos and 652 vs 674 in T.
cordata). This indicates that there were ample potential candidates for clonal or closely
related pairs, i.e. within ‘critical distance’, in both regions. The conclusion that relatedness
between Tilia trees is unlikely over distances of more than 61m (top quintile), has important
consequences for woodland managers.

4.5 Population substructure

Several studies in forest trees have shown that clones or related trees can often only be
expected at relatively short distances, e.g. only up to 2m in Quercus crispula (Aizawa et al.,
2017). In Tilia, the average distance of full-sibs and parent-offspring related pairs is 52m
(median 21m) and the average distance between clones is 16.8m (median 15m), which is
similar to the spatial scales revealed in Fagus grandifolia and Prunus avium (Kitamura et al.,
2003; Vaughan et al., 2007) and slightly more than in tanoak (Dodd et al., 2014). This means
that close relatedness is only expected to be observed at a rather small scale and in
populations with high density of adult trees (Duminil et al., 2016). Indeed this is reflected in
some populations with large distances between samples, such as NOSO, not showing any
small scale population structure (data not presented).
We can now attempt to answer the questions whether leading range-edge populations are at risk and whether they are suitable for range expansion. Leading range-edge populations of *Tilia* do not have a lower genetic diversity than central populations, making it likely that range-edge populations are remnants of a genetically diverse moving front, with no prolonged bottleneck effects. Although range-edge populations are more clonal than central populations, they contain medium to high genetic diversity (perhaps from an earlier period in time, e.g. Ennos, 2003) and show no inbreeding ($F_{IS}$) so should be able to regenerate, naturally or assisted (Macaya-Sanz *et al.*, 2016). With the climate warming at range-edge locations sexual reproduction may become more prevalent and therefore should promote *in situ* adaptation to novel conditions. In addition, current populations may spread: with more continuous populations natural gene flow from more southern populations may occur. However, the effective population size is smaller at range-edge, meaning that if the fragmentation and demographics stay the same, there is a risk that genetic diversity may be lost as few individuals contribute to the next generation, while in central range populations this risk is limited as there is a more even spread of sexual reproduction amongst adult trees. Range-edge populations generally have $N_e < 50$, which is considered low and makes the species vulnerable due to low ability to adapt to changes (Myking, 2002), and this is exacerbated by clonal reproduction, lowering $N_e$ to below 10. For example, the *T. platyphylllos* population UKAS, has small $N_e$ and a high rate of clonal reproduction. Those types of populations seem particularly at risk, and maybe at a tipping point, due to the combined effects of strong fragmentation, human activity, and the small number of, sometimes very old, trees, belonging to few clones. Regeneration here may require good management to allow adaptation, *e.g.* to climate change.

Woodland managers and foresters wishing to sample trees for planting further north (assisted migration) can use populations that are currently at the northern range-edge of the
species because in those populations diversity is not lower than in central regions.

Furthermore, northern populations are likely to be better adapted to conditions in the north than those from central regions, e.g. to climatic, season length and day light differences (Lobo et al., 2018a). However, further work is required to test whether Tilia set seed at these northern locations. An alternative option for replication would be through clonal propagation, natural or assisted. The additional use of other sources, e.g. from central ranges with diverse microclimates, may generate the optimal genetic composition and diversity for adaptation. In the wake of a warming climate and a shifting natural range, successful sexual regeneration will greatly contribute to adaptation to local and novel environmental conditions (Thomas et al., 2014).

Forest managers need to take into account sufficient distance between donor trees (be it seed or clonal material), i.e. 26m between sampling to avoid two identical genotypes, and 61m to avoid two trees that are related to the first degree (avoiding all but top quintile). Adding diverse tree species to forests has become popular, and Tilia in particular has shown to increase the herb layer (Normann et al., 2016) and the decomposition of leaf litter (Muys et al., 1992). Furthermore, Tilia is known to be a strong tree with specific associated communities performing important ecosystem services (Hommel et al., 2007).

Here we have compared northern range-edge with central populations, but we recognise that a similar study should be performed on ‘trailing edge’ or ‘stable rear edge’ Tilia populations to ensure conservation of the diversity in these species (Hampe & Petit, 2005). Pinpointing the existence of such populations and obtaining samples is particularly difficult. Nonetheless, our findings may also be of importance for trailing edge populations because Tilia used to be one of the dominant tree species in Europe and current populations are remnants of these woodlands (e.g. UK and Denmark), and are, similarly to leading edge populations, now fragmented and limited due to human interference and climate change.

Acknowledgements:
Comment from anonymous reviewers are gratefully acknowledged as they helped to improve the manuscript tremendously. We are also grateful to everyone who contributed to sample collection: Hanne Hegre Grundt, Tor Myking, Arthur Leewis, Bruno Fady, Bruno Chopart, Honor Prentice, Leena Yrjänä, Ole Hansen, Seb Mankelow, Tim Laurie, Jon Tomkinson and Jade Lauren Gunnell. Thanks to Heino Konrad for indicating Austrian populations and to Milan Chytrý for providing *Tilia cordata* samples from Siberia. Thanks to Natural England, the National Trust and the private landowners who granted permission for us to collect leaf samples from their woodlands and Renata Krzyściak-Kosińska for her assistance in gaining permission to collect leaf samples from the Białowieża National Park, Poland. Thanks to Helen Martin for laboratory support and Patrick Meirmans for support using GenoDive. This study was funded by a RB Cooke Studentship (SL), the Thai government (PP) and a Royal Society of Biology Travel Grant (SL collection of Polish samples).
References:


doi:10.1093/aob/mcu025


**Data Accessibility Statement:** Sampling locations and microsatellite genotypes have been made available through Mendeley DOI: doi:10.17632/t7j48wg28z.1

**Author contributions:**

SL, PP and KW collected the genotype data, SL and KW designed the experiment, collected most of the samples, performed the genetic data analyses and wrote the manuscript, RS performed the GLM statistical analyses. All authors read and commented on the ms.
Table 1  *Tilia* populations, ordered by location codes within groups, country, average distance (Av Dist) between trees in m, number (N) and coordinates of trees: non-shaded blocks are leading range-edge populations, while shaded blocks are central range populations (na not assessed).

<table>
<thead>
<tr>
<th>Species/Code</th>
<th>Location</th>
<th>Country</th>
<th>Av Dist</th>
<th>N</th>
<th>Latitude (°N)</th>
<th>Longitude (°E)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. platyphyllos</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEBOp</td>
<td>Bolderslev</td>
<td>Denmark</td>
<td>38</td>
<td>11</td>
<td>54.9978</td>
<td>9.3512</td>
</tr>
<tr>
<td>SWWC</td>
<td>West Coast</td>
<td>Sweden</td>
<td>na</td>
<td>15</td>
<td>58.8967</td>
<td>11.0419</td>
</tr>
<tr>
<td>UKAS</td>
<td>Aplegarth</td>
<td>UK</td>
<td>67</td>
<td>13</td>
<td>54.4093</td>
<td>-1.8165</td>
</tr>
<tr>
<td>UKAW</td>
<td>Anston Wd</td>
<td>UK</td>
<td>54</td>
<td>15</td>
<td>53.3402</td>
<td>-1.1984</td>
</tr>
<tr>
<td>UKHD</td>
<td>Hudswell</td>
<td>UK</td>
<td>269</td>
<td>25</td>
<td>54.4012</td>
<td>-1.7573</td>
</tr>
<tr>
<td>UKHW</td>
<td>Halesend Wd</td>
<td>UK</td>
<td>52</td>
<td>24</td>
<td>52.1407</td>
<td>-2.3815</td>
</tr>
<tr>
<td>AUDOp</td>
<td>Dobra</td>
<td>Austria</td>
<td>27</td>
<td>14</td>
<td>48.5907</td>
<td>15.3974</td>
</tr>
<tr>
<td>AULE</td>
<td>Leopoldsberg</td>
<td>Austria</td>
<td>91</td>
<td>25</td>
<td>48.2774</td>
<td>16.3522</td>
</tr>
<tr>
<td>AUSOp</td>
<td>Sommerein</td>
<td>Austria</td>
<td>58</td>
<td>14</td>
<td>47.9848</td>
<td>16.6947</td>
</tr>
<tr>
<td>FRML</td>
<td>Mont Lure</td>
<td>France</td>
<td>1250</td>
<td>20</td>
<td>44.1252</td>
<td>5.8705</td>
</tr>
<tr>
<td>FRVE</td>
<td>Ventoux</td>
<td>France</td>
<td>374</td>
<td>20</td>
<td>44.1869</td>
<td>5.2382</td>
</tr>
<tr>
<td>GELC</td>
<td>Lichtenstein</td>
<td>Germany</td>
<td>362</td>
<td>31</td>
<td>48.4071</td>
<td>9.2627</td>
</tr>
<tr>
<td><em>T. cordata</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEBOc</td>
<td>Bolderslev</td>
<td>Denmark</td>
<td>55</td>
<td>30</td>
<td>54.9978</td>
<td>9.3516</td>
</tr>
<tr>
<td>FIIS</td>
<td>Iso Kirvessaari</td>
<td>Finland</td>
<td>49</td>
<td>20</td>
<td>62.5548</td>
<td>27.7626</td>
</tr>
<tr>
<td>FINI</td>
<td>Niinisaaar</td>
<td>Finland</td>
<td>36</td>
<td>20</td>
<td>61.8171</td>
<td>29.3895</td>
</tr>
<tr>
<td>NOSO</td>
<td>Sogn Fjordane</td>
<td>Norway</td>
<td>854</td>
<td>16</td>
<td>61.8505</td>
<td>6.1375</td>
</tr>
<tr>
<td>RUVA</td>
<td>Vagay area</td>
<td>Russia</td>
<td>na</td>
<td>20</td>
<td>57.5096</td>
<td>69.1954</td>
</tr>
<tr>
<td>SWOL</td>
<td>Oland</td>
<td>Sweden</td>
<td>71</td>
<td>27</td>
<td>56.6070</td>
<td>16.4503</td>
</tr>
<tr>
<td>UKBB</td>
<td>Brignal Banks</td>
<td>UK</td>
<td>36</td>
<td>19</td>
<td>54.4968</td>
<td>-1.9114</td>
</tr>
<tr>
<td>UKRO</td>
<td>Roudsea</td>
<td>UK</td>
<td>47</td>
<td>40</td>
<td>54.2337</td>
<td>-3.0260</td>
</tr>
<tr>
<td>AUDOc</td>
<td>Dobra</td>
<td>Austria</td>
<td>50</td>
<td>21</td>
<td>48.5909</td>
<td>15.3975</td>
</tr>
<tr>
<td>AUSOc</td>
<td>Sommerein</td>
<td>Austria</td>
<td>83</td>
<td>23</td>
<td>47.9848</td>
<td>16.6947</td>
</tr>
<tr>
<td>AUST</td>
<td>Stams</td>
<td>Austria</td>
<td>106</td>
<td>21</td>
<td>47.2757</td>
<td>10.9772</td>
</tr>
<tr>
<td>AUTH</td>
<td>Thayatal</td>
<td>Austria</td>
<td>725</td>
<td>17</td>
<td>48.8448</td>
<td>15.8878</td>
</tr>
<tr>
<td>CZVO</td>
<td>Velky Osek</td>
<td>Czech Rep</td>
<td>67</td>
<td>20</td>
<td>50.1012</td>
<td>15.1788</td>
</tr>
<tr>
<td>FRMO</td>
<td>Mouthiers</td>
<td>France</td>
<td>69</td>
<td>20</td>
<td>48.9120</td>
<td>4.9138</td>
</tr>
<tr>
<td>GECO</td>
<td>Colbitz</td>
<td>Germany</td>
<td>270</td>
<td>20</td>
<td>52.3310</td>
<td>11.5559</td>
</tr>
<tr>
<td>POBF</td>
<td>Białowieża</td>
<td>Poland</td>
<td>930</td>
<td>40</td>
<td>52.7272</td>
<td>23.8351</td>
</tr>
</tbody>
</table>
**Table 2** Means of estimates of clonal occurrence in *Tilia* and significance of fixed and covariate effects.

<table>
<thead>
<tr>
<th>Species/region/P-values</th>
<th>N</th>
<th>G</th>
<th>R</th>
<th>D*</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. platyphyllos</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>17.2</td>
<td>14.7</td>
<td>0.804</td>
<td>0.947</td>
</tr>
<tr>
<td>Central</td>
<td>21.0</td>
<td>20.7</td>
<td>0.981</td>
<td>0.998</td>
</tr>
<tr>
<td><em>T. cordata</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>24.0</td>
<td>17.6</td>
<td>0.685</td>
<td>0.937</td>
</tr>
<tr>
<td>Central</td>
<td>23.0</td>
<td>23.0</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**P-values**

<table>
<thead>
<tr>
<th></th>
<th>&lt; 0.0005</th>
<th>&lt; 0.0005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>0.143</td>
<td>0.186</td>
</tr>
<tr>
<td>Av Dist</td>
<td>0.053</td>
<td>0.099</td>
</tr>
<tr>
<td>Region x species interaction</td>
<td>0.291</td>
<td>0.425</td>
</tr>
</tbody>
</table>

**Footnote:** Number of samples (N), number of genotypes (G) and clonal occurrence presented as genotypic richness $R = (G-1)/(N-1)$ and Simpson’s complement index for genotypic diversity $D^*$, from range-edge and central locations, $P$ values for significance of fixed-effect of region and species, co-variate effect of the average distance between samples (Av Dist), interaction of species and region effect, significant effects are indicated in bold.
Table 3 Genetic characteristics of *Tilia* as means per group of populations at the genet level and outcomes of statistical tests (*P* values)

<table>
<thead>
<tr>
<th>Species/region</th>
<th>$H_e$</th>
<th>$F_{IS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. platyphyllos</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>0.690</td>
<td>-0.061</td>
</tr>
<tr>
<td>Central</td>
<td>0.745</td>
<td>-0.024</td>
</tr>
<tr>
<td><em>T. cordata</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>0.553</td>
<td>-0.037</td>
</tr>
<tr>
<td>Central</td>
<td>0.573</td>
<td>0.013</td>
</tr>
<tr>
<td><strong>P-values</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect region</td>
<td>0.319</td>
<td>0.545</td>
</tr>
<tr>
<td>Effect species</td>
<td>&lt; 0.0005</td>
<td>0.729</td>
</tr>
<tr>
<td>Av Dist</td>
<td>0.373</td>
<td>NA</td>
</tr>
<tr>
<td>Region x species interaction</td>
<td>0.161</td>
<td>0.874</td>
</tr>
</tbody>
</table>

Footnote: Expected heterozygosity ($H_e$) and fixation index ($F_{IS}$) and *P*-values for fixed effects of region, species and interaction effects, as well as significance of the average distance between samples (Av Dist) within a population, used as a covariate, significant effects are indicated in bold.
Table 4. Expected heterozygosity $H_e$ in all genotypes of a population, observed heterozygosity $H_o$ and fixation index $F_{IS}$ of non-clonal and clonal MLL in six populations of *T. cordata*, with number of samples in a group indicated as n.

<table>
<thead>
<tr>
<th>Population</th>
<th>$H_e$ over all (n)</th>
<th>$H_o$ non-clonal (n)</th>
<th>$H_o$ clonal (n)</th>
<th>$F_{IS}$ non-clonal genets</th>
<th>$F_{IS}$ clonal genets</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEBOc</td>
<td>0.575 (18)</td>
<td>0.545 (13)</td>
<td>0.550 (5)</td>
<td>0.052</td>
<td>0.043</td>
</tr>
<tr>
<td>FIIS</td>
<td>0.505 (12)</td>
<td>0.552 (8)</td>
<td>0.583 (4)</td>
<td>-0.093</td>
<td>-0.155</td>
</tr>
<tr>
<td>FINI</td>
<td>0.499 (13)</td>
<td>0.481 (9)</td>
<td>0.583 (4)</td>
<td>0.035</td>
<td>-0.169</td>
</tr>
<tr>
<td>RUVA</td>
<td>0.576 (14)</td>
<td>0.483 (10)</td>
<td>0.563 (4)</td>
<td>0.161</td>
<td>0.024</td>
</tr>
<tr>
<td>UKBB</td>
<td>0.520 (6)</td>
<td>0.556 (3)</td>
<td>0.639 (3)</td>
<td>-0.069</td>
<td>-0.229</td>
</tr>
<tr>
<td>UKRO</td>
<td>0.566 (29)</td>
<td>0.561 (19)</td>
<td>0.600 (10)</td>
<td>0.008</td>
<td>-0.060</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.540</strong></td>
<td><strong>0.530</strong></td>
<td><strong>0.586</strong></td>
<td><strong>0.016</strong></td>
<td><strong>-0.091</strong></td>
</tr>
</tbody>
</table>
Table 5. Diversity and G-statistics of leading range-edge and central populations of *Tilia* species and *P*-values for comparisons between the regions (GenoDive 2.0b) with 95% confidence intervals (bootstrapping) between brackets at the genet level, significant effects are indicated in bold.

<table>
<thead>
<tr>
<th>Species/region</th>
<th><em>H</em>&lt;sub&gt;o&lt;/sub&gt;</th>
<th><em>H</em>&lt;sub&gt;s&lt;/sub&gt;</th>
<th><em>G</em>&lt;sub&gt;is&lt;/sub&gt;</th>
<th><em>G</em>&lt;sub&gt;st&lt;/sub&gt;</th>
<th><em>G</em>'&lt;sub&gt;st&lt;/sub&gt;</th>
<th><em>D</em>&lt;sub&gt;est&lt;/sub&gt;</th>
<th>Clonal div</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. platyphyllos</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>0.732</td>
<td>0.72</td>
<td>-0.017</td>
<td>0.108</td>
<td>0.442</td>
<td>0.375</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td>(0.651-0.805)</td>
<td>(0.652-0.772)</td>
<td>(-0.064-0.032)</td>
<td>(0.092-0.124)</td>
<td>(0.325-0.564)</td>
<td>(0.252-0.507)</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>0.764</td>
<td>0.760</td>
<td>-0.005</td>
<td>0.049</td>
<td>0.234</td>
<td>0.195</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>(0.649-0.861)</td>
<td>(0.648-0.850)</td>
<td>(-0.042-0.034)</td>
<td>(0.041-0.060)</td>
<td>(0.166-0.350)</td>
<td>(0.123-0.320)</td>
<td></td>
</tr>
<tr>
<td><em>P</em>-value</td>
<td>0.054</td>
<td><strong>0.035</strong></td>
<td>0.657</td>
<td><strong>0.014</strong></td>
<td><strong>0.013</strong></td>
<td><strong>0.013</strong></td>
<td>0.068</td>
</tr>
<tr>
<td><em>T. cordata</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>0.570</td>
<td>0.575</td>
<td>0.008</td>
<td>0.084</td>
<td>0.214</td>
<td>0.142</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>(0.420-0.699)</td>
<td>(0.424-0.705)</td>
<td>(-0.050-0.069)</td>
<td>(0.064-0.104)</td>
<td>(0.135-0.331)</td>
<td>(0.070-0.261)</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>0.558</td>
<td>0.587</td>
<td>0.050</td>
<td>0.048</td>
<td>0.126</td>
<td>0.082</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(0.406-0.694)</td>
<td>(0.425-0.731)</td>
<td>(-0.015-0.126)</td>
<td>(0.030-0.070)</td>
<td>(0.069-0.206)</td>
<td>(0.036-0.158)</td>
<td></td>
</tr>
<tr>
<td><em>P</em>-value</td>
<td>0.644</td>
<td>0.525</td>
<td>0.189</td>
<td>0.064</td>
<td><strong>0.050</strong></td>
<td><strong>0.040</strong></td>
<td><strong>0.003</strong></td>
</tr>
</tbody>
</table>

*H*<sub>o</sub> observed heterozygosity and *H*<sub>s</sub> expected heterozygosity within populations, *G*<sub>is</sub> inbreeding coefficient and *G*<sub>st</sub> fixation index, *G*'<sub>st</sub> is the fixation index corrected for bias due to limited number of populations, and Jost’s *D*<sub>est</sub> population differentiation independent from *H*<sub>s</sub>, and clonal diversity as measure of amount of sexual reproduction, analogous to Simpsons D*.
Table 6. Distribution of genetic diversity (AMOVA) within and among groups of populations in *Tilia* species at the genet level (GenoDive 2.0b). Standard deviations (St dev) of F-statistics were obtained through jack-knifing over loci and 95% confidence intervals of F-statistics were obtained through bootstrapping over loci. P values indicate significance of F-value > 0.

<table>
<thead>
<tr>
<th></th>
<th>%var</th>
<th>F-statistic</th>
<th>F-value</th>
<th>St dev</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. platyphyllos</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Range edge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within individuals</td>
<td>0.886</td>
<td>$F_{it}$</td>
<td>0.114 (0.063-0.167)</td>
<td>0.028</td>
<td>-</td>
</tr>
<tr>
<td>Among individuals</td>
<td>0.002</td>
<td>$F_{is}$</td>
<td>0.002 (-0.050-0.059)</td>
<td>0.029</td>
<td>0.486</td>
</tr>
<tr>
<td>Among populations</td>
<td>0.112</td>
<td>$F_{ST}$</td>
<td>0.112 (0.098-0.127)</td>
<td>0.008</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>Central</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within individuals</td>
<td>0.940</td>
<td>$F_{it}$</td>
<td>0.053 (0.016-0.096)</td>
<td>0.210</td>
<td>-</td>
</tr>
<tr>
<td>Among individuals</td>
<td>0.003</td>
<td>$F_{is}$</td>
<td>0.003 (-0.038-0.034)</td>
<td>0.019</td>
<td>0.631</td>
</tr>
<tr>
<td>Among populations</td>
<td>0.055</td>
<td>$F_{ST}$</td>
<td>0.055 (0.046-0.068)</td>
<td>0.006</td>
<td>0.001</td>
</tr>
<tr>
<td><em>T. cordata</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Range edge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within individuals</td>
<td>0.900</td>
<td>$F_{it}$</td>
<td>0.099 (0.044-0.148)</td>
<td>0.027</td>
<td>-</td>
</tr>
<tr>
<td>Among individuals</td>
<td>0.008</td>
<td>$F_{is}$</td>
<td>0.099 (-0.048-0.065)</td>
<td>0.031</td>
<td>0.306</td>
</tr>
<tr>
<td>Among population</td>
<td>0.091</td>
<td>$F_{ST}$</td>
<td>0.091 (0.066-0.115)</td>
<td>0.013</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>Central</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within individuals</td>
<td>0.903</td>
<td>$F_{it}$</td>
<td>0.097 (0.027-0.176)</td>
<td>0.041</td>
<td>-</td>
</tr>
<tr>
<td>Among individuals</td>
<td>0.041</td>
<td>$F_{is}$</td>
<td>0.043 (-0.022-0.117)</td>
<td>0.038</td>
<td>0.001</td>
</tr>
<tr>
<td>Among population</td>
<td>0.056</td>
<td>$F_{ST}$</td>
<td>0.056 (0.038-0.079)</td>
<td>0.011</td>
<td>0.001</td>
</tr>
</tbody>
</table>
Table 7. Means or medians of relatedness characteristics of *Tilia*, at the genet level and outcomes of statistical tests for region, species and interaction effects (*P*-values, significant effects are indicated in bold).

<table>
<thead>
<tr>
<th></th>
<th><em>T. platyphyllos</em></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>r</em> ML-Relate mean</td>
<td>1st+2nd degr rel mean</td>
<td>1st degr rel mean</td>
<td><em>N_e</em> Co-an median</td>
<td><em>N_e</em> LD median</td>
</tr>
<tr>
<td>Range-edge</td>
<td>0.060</td>
<td>0.060</td>
<td>0.027</td>
<td>8.7</td>
<td>16.2</td>
</tr>
<tr>
<td>Central</td>
<td>0.038</td>
<td>0.042</td>
<td>0.013</td>
<td>48.1</td>
<td>44.5</td>
</tr>
<tr>
<td></td>
<td><em>T. cordata</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>0.054</td>
<td>0.060</td>
<td>0.023</td>
<td>60.9</td>
<td>38.2</td>
</tr>
<tr>
<td>Central</td>
<td>0.046</td>
<td>0.048</td>
<td>0.009</td>
<td>∞</td>
<td>494.5</td>
</tr>
</tbody>
</table>

*P*-values

<table>
<thead>
<tr>
<th></th>
<th><em>r</em> ML-Relate mean</th>
<th>1st+2nd degr rel mean</th>
<th>1st degr rel mean</th>
<th><em>N_e</em> Co-an median</th>
<th><em>N_e</em> LD median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect region</td>
<td>0.236</td>
<td>0.120</td>
<td>0.003</td>
<td>0.049</td>
<td>0.031</td>
</tr>
<tr>
<td>Effect species</td>
<td>0.079</td>
<td>0.269</td>
<td>0.674</td>
<td>0.009</td>
<td>0.078</td>
</tr>
<tr>
<td>Av.Dist.</td>
<td>0.381</td>
<td>0.697</td>
<td>0.518</td>
<td>0.578</td>
<td>0.889</td>
</tr>
<tr>
<td>Region x species interaction</td>
<td>0.134</td>
<td>0.215</td>
<td>0.180</td>
<td>0.885</td>
<td>0.292</td>
</tr>
</tbody>
</table>

**Footnote:** Shown are the average relatedness of individuals (*r*, ML-Relate), proportion of first and second degree related individuals, proportion of first degree related individuals and effective population size (*N_e*) calculated based on Co-ancestry and on Linkage Disequilibrium (LD).
Table 8 Means or medians of relatedness characteristics of *Tilia*, at the ramet (sample) level and outcomes of statistical tests for region, species and interaction effects (*P* values, significant effects are indicated in bold).

<table>
<thead>
<tr>
<th></th>
<th>$r_{ML}$-Relate mean</th>
<th>1st+2nd degr rel mean</th>
<th>1st degr rel mean</th>
<th>$N_e$ Co-an median</th>
<th>$N_e$ LD median</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. platyphyllos</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>0.109</td>
<td>0.060</td>
<td>0.027</td>
<td>9.4</td>
<td>5.85</td>
</tr>
<tr>
<td>Central</td>
<td>0.040</td>
<td>0.042</td>
<td>0.013</td>
<td>48.2</td>
<td>37.1</td>
</tr>
<tr>
<td><em>T. cordata</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range-edge</td>
<td>0.108</td>
<td>0.060</td>
<td>0.023</td>
<td>8.8</td>
<td>3.0</td>
</tr>
<tr>
<td>Central</td>
<td>0.046</td>
<td>0.048</td>
<td>0.009</td>
<td>∞</td>
<td>494.5</td>
</tr>
</tbody>
</table>

$P$-values

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect region</td>
<td>&lt;0.0005</td>
<td>0.519</td>
<td><strong>0.045</strong></td>
<td>0.005</td>
<td>&lt; 0.0005</td>
</tr>
<tr>
<td>Effect species</td>
<td>0.279</td>
<td>0.154</td>
<td>0.453</td>
<td>0.048</td>
<td>0.042</td>
</tr>
<tr>
<td>Av.Dist.</td>
<td><strong>0.040</strong></td>
<td>0.942</td>
<td>0.474</td>
<td>0.335</td>
<td>0.352</td>
</tr>
<tr>
<td>Region x species interaction</td>
<td>0.688</td>
<td>0.349</td>
<td>0.132</td>
<td>0.425</td>
<td>0.079</td>
</tr>
</tbody>
</table>

Footnote: Shown are the average relatedness of individuals ($r$, ML-Relate), proportion of first and second degree related individuals, proportion of first degree related individuals and effective population size ($N_e$) calculated based on Co-ancestry and on Linkage Disequilibrium (LD).
Figure legends:

**Fig 1.** Distribution of *Tilia cordata* (light green area and dots) and *T. platyphyllos* (dark green area and dots). Distribution data from EUFORGEN (www.euforgen.org). Map constructed using QGIS v2.16.2 (www.qgis.org). Triangles are *T. cordata* populations sampled while circles are *T. platyphyllos*. Orange icons show leading range-edge populations, for *T. cordata (>54°N)* with yellow background and for *T. platyphyllos (>52°N)* with white background. Blue icons are populations from central range of the two species (>43°N), *T. cordata* with orange background and *T. platyphyllos* with grey background.

**Figure 2** Counts of numbers of a) clone pairs, in *T. platyphyllos* (green) and *T. cordata* (blue) across distance classes b) first degree related pairs of individuals (full sib and parent-offspring), three pairs with distance of 483, 684 and 809m are combined into >300m class.
Fig 1. Distribution of *Tilia cordata* (light green area and dots) and *T. platyphyllos* (dark green area and dots). Distribution data from EUFORGEN (www.euforgen.org). Map constructed using QGIS v2.16.2 (www.qgis.org). Triangles are *T. cordata* populations sampled while circles are *T. platyphyllos*. Orange icons show leading range-edge populations, for *T. cordata* (>54°N) with yellow background and for *T. platyphyllos* (>52°N) with white background. Blue icons are populations from central range of the two species (>43°N), *T. cordata* with orange background and *T. platyphyllos* with grey background.
Figure 2 Counts of numbers of a) clone pairs b) first degree related pairs of individuals (full sib and parent-offspring), in *T. platyphyllos* (green) and *T. cordata* (blue) across distance classes. Three related pairs with distances of 483, 684 and 809m are combined into the >300m class.