**RESEARCH PAPER**

# Trunk perimeter correlates with genetic bottleneck intensity and the level of genetic diversity in populations of *Taxus baccata* L

Sara Stefanowska[1] · Katarzyna Meyza[1] · Grzegorz Iszkuło[2,3] · Igor J. Chybicki[1]

## Abstract

***Key message*** *Taxus baccata* **remnants established recently tend to contribute less to the species' overall genetic variation than historical populations because they are subjected to a greater impact of the founder effect and genetic isolation. As tree trunk perimeter is a rough indicator of genetic variation in a population, this measure should be considered in conservation programs.**

**Context** Genetic variation within *Taxus baccata* (L.) populations is not associated with the current census size but correlates well with the effective size, suggesting that genetic drift intensity reflects variation in demographic histories.

**Aims** We hypothesize that recently established populations are subjected to greater bottleneck than old remnants. Using the mean trunk perimeter as a surrogate of tree age, we test whether the demographic history and genetic variation are associated with the mean tree age.

**Methods** Using 18 microsatellite markers, we analyze the genetic diversity and demographic history of 11 yew populations in Poland to assess the relationship between the mean trunk perimeter and the inferred genetic parameters.

**Results** Populations reveal significant differences in levels of genetic variation and in the intensity and time of genetic bottleneck. After excluding an apparent outlier, the genetic variation is significantly greater while the bottleneck intensity lower in populations with a greater perimeter.

**Conclusion** Due to continuous species decline and increasing fragmentation, the non-uniform contribution of yew remnants to the overall genetic variation tends to decrease together with the mean tree age. Germplasm collections for the species should take into account tree perimeter as a rough indicator of the genetic variation of a population.

**Keywords** Genetic variation · Demographic history · Bottleneck · Tree size · Conservation · *Taxus baccata*

✉ Igor J. Chybicki
igorchy@ukw.edu.pl

Sara Stefanowska
sara.stefanowska@ukw.edu.pl

Katarzyna Meyza
kasiakow@ukw.edu.pl

Grzegorz Iszkuło
iszkulo@man.poznan.pl

# 1 Introduction

Preserving genetic variation is crucial for protecting endangered species and populations (Aravanopoulos 2016; Ralls et al. 2018), especially to mitigate effects of progressing anthropogenic changes, such as increased exploitation and fragmentation of habitats, as well as the on-going climate changes. Even if it is difficult to predict responses of wild

[1] Department of Genetics, Kazimierz Wielki University, Chodkiewicza 30, 85064 Bydgoszcz, Poland

[2] Institute of Dendrology, Polish Academy of Sciences, Parkowa 5, 62035 Kórnik, Poland

[3] Institute of Biological Sciences, University of Zielona Góra, Prof. Z. Szafrana 1, 65516 Zielona Góra, Poland

INRAℇ ⁂ Springer

populations to environmental threats, the standing genetic variation is believed to be a fundamental driver for adaptive changes in the future (Pelletier and Coltman 2018).

In addition to natural selection, gene flow among populations and genetic drift are the main factors shaping the genetic resources of a population (Johnson et al. 2018). Generally, gene exchange between populations helps to maintain the genetic variation within populations, neutralizing the negative effect of genetic erosion due to local genetic drift (Willi et al. 2006). As a result, gene flow can maintain or even increase a population's adaptive potential (Swindell and Bouzat 2006). Furthermore, gene flow can counteract inbreeding depression (Ingvarsson and Whitlock 2000) and stochastic demographic fluctuations (Alleaume-Benharira et al. 2006), which otherwise can push isolated populations towards extinction (Brook et al. 2002).

In a meta-population with extinction and recolonization dynamics (turnover dynamics; Haag et al. 2005), newly established (pioneer) populations are characterized by low genetic diversity due to the founder effect (McCauley et al. 1995). Under a sufficient gene flow rate relative to the extinction rate, a pioneer population gains genetic variation, reaching the migration-drift equilibrium eventually. Stochastic population turnover causes a meta-population to consist of populations of different age and different levels of accumulated genetic diversity. Consequently, at any given snapshot in time, a positive relationship between population age and genetic diversity is expected across a meta-population. However, when gene flow rates are low and/or turnover rates are high, the age vs. genetic diversity relationship may be undetectable because populations go extinct before they start reaching the drift-migration equilibrium (Whitlock 1992; McCauley et al. 1995). Therefore, it is not surprising that empirical studies provide only partial support for the "population age vs. genetic diversity" relationship (e.g., Schneller and Holderegger 1996; Powolny et al. 2016). Natural populations do not conform to the "age vs. genetic diversity" pattern when numerous source populations are located a short distance from a colonized patch (Helsen et al. 2013), or/and when a species has the potential to long-distance dispersal (Raffl et al. 2008; Pluess 2011). In such cases, newly founded populations tend to rapidly reach high genetic diversity and low divergence at the early successional stage and quickly become indistinguishable from source populations (Lefèvre et al. 2004). In this respect, trees seem to be especially "resistant" to the founder effect during colonization (Austerlitz et al. 2000; Born et al. 2008; Elleouet and Aitken 2019; Lander et al. 2020), not only due to high dispersal capabilities but also to specific life-history traits which alleviate genetic drift associated with founder effects.

*Taxus baccata* L. can serve as a good example of a species characterized by a well-fragmented meta-population

subjected to the extinction-recolonization dynamics. *T. baccata* is a dioecious conifer having a wide range of natural distribution across Europe, the Middle East, and North Africa (Thomas and Polwart 2003). Nevertheless, throughout its range, the natural yew's occurrence is limited to small, isolated populations (Thomas and Polwart 2003; Thomas and Garcia-Martínez 2015; Iszkuło et al. 2016). It is believed that such a high fragmentation is mainly the result of adverse climate conditions in the Holocene, intensive browsing, forest management, and the exploitation of this species in the past (e.g., Linares 2013). Studies on the ancient range of the yew showed that until 2500–5000 years ago, it was a fairly common species in Europe (Deforce and Bastiaens 2007), although the abundance was likely relatively variable across the range (Uzquiano et al. 2015; Iszkuło et al. 2016). The progressive specialization of land use, including the forest areas' transformation into agricultural fields and monoculture standards in the recent past, caused that today, *T. baccata* is a rare species in natural sites throughout its distribution. On the other hand, *T. baccata* is very common as an ornamental plant. For this reason, taking into account anthropogenic "populations", it is difficult to consider the yew an endangered species. Nevertheless, the future condition of natural populations is unclear, especially in the context of climate change and the current scale of isolation of individual sites (Thomas and Garcia-Martínez, 2015; Chybicki and Oleksa 2018).

*T. baccata* is characterized by high genetic variation (Myking et al. 2009; Dubreuil et al. 2010; González-Martínez et al. 2010; Litkowiec et al. 2015), but also a strong genetic differentiation among populations (Myking et al. 2009; Dubreuil et al. 2010; González-Martínez et al. 2010; Burgarella et al. 2012; Litkowiec et al. 2018). So far, the studies indicated that the current genetic structure is shaped by a genetic drift that is often driven by genetic bottlenecks (Chybicki et al. 2012; Litkowiec et al. 2018). It seems that this is a direct effect of isolation and limited gene flow both among (Hilfiker et al. 2004; Myking et al. 2009; Dubreuil et al. 2010) and within populations (Lavabre et al. 2016; Chybicki and Oleksa 2018). Limited seed and pollen dispersal create a spatial genetic structure within natural populations (Dubreuil et al. 2010; Chybicki et al. 2016), which also results in mating between relatives (Chybicki et al. 2019). Interestingly, contemporary levels of biparental inbreeding co-vary with distance to the nearest male and tree girth, but does not depend on the census population size (Chybicki et al. 2019). The effect of distance to the nearest male can be associated with spatially restricted mating patterns (Chybicki and Oleksa 2018) and the spatial genetic structure (Dubreuil et al. 2010; Chybicki et al. 2016). On the other hand, the decrease in biparental inbreeding level together with the increase in mean tree girth suggests there is a trend towards decreased opportunities and/or possibilities

for sib-ship mating in older yew remnants. Such a pattern can result from the maintenance of sufficient frequency of unrelated individuals or sufficient frequency of lethal factors (or both) in old populations, suggesting that the level of genetic diversity may co-vary with the mean tree age in a population.

The variation in the ratio of effective to census population size observed previously (Chybicki et al. 2012; Litkowiec et al. 2018) suggested different rates of genetic drift across yew stands, likely due to differences in demographic histories. Here, we studied whether demographic histories and genetic diversity are associated with the mean adult tree age. Because naturally occurring *T. baccata* is under protection in Poland, the trunk perimeter was used as a non-invasive but still highly predictive surrogate of tree age (Chen et al. 2020). Consistently with theoretical predictions for a meta-population with extinction-recolonization dynamics, we hypothesized that genetic diversity increases with the mean adult tree age. In addition, noting that the yew is a very long-lived tree species, we assumed that remnants characterized by older adult trees are representatives of a more abundant historical distribution of the species, while populations with younger adult trees (i.e., presumably established recently) show effects of a strong bottleneck due to a limited number of recent seed sources. Therefore, we hypothesized that molecular signs of genetic bottleneck decrease together with the increase in mean tree age.

# 2 Materials and methods

## 2.1 Plant sampling and DNA analysis

This study examined 11 populations of *Taxus baccata* in Poland. Populations were located in protected areas, except for Kórnik. However, yew trees in Kórnik have been left to regenerate freely, and can therefore be considered an undisturbed, semi-natural (naturalized) population (Iszkuło and Boratyński 2005). Except for two remnants (Cisowa Góra A and B) which are only 1 km apart, populations are well separated, and the mean distance between populations is 238 km (see Fig. 1D). All populations were characterized by the mean trunk perimeter. For this purpose, to avoid the impact of sex ratio differences between samples, only female trees (15 random adult female trees per population) were measured at breast height (1.3 m) (Table 1). In every population, we selected only mature, well seed-bearing trees, representing possibly the oldest individuals within a given location. In this way, we attempted to assess the maximum tree perimeter, reflecting the oldest possible demographic events. Among the study populations, only three were previously characterized for the tree age based on dendrochronological measurements: Cisowa Góra A (173–218 years),

Kórnik (105–148 years), and Rokita (49–51 years) (based on two trees per site; Cedro and Iszkuło 2011). Based on these limited data, tree age and tree perimeter show good agreement with each other, despite substantial differences in habitat characteristics among the three populations. Within each site, 17 to 58 trees were sampled (needles), yielding a total of 384 individuals. The total genomic DNA was extracted using GeneMATRIX Plant & Fungi DNA Purification Kit (EURx, Poland). Eighteen EST (Expressed Sequence Tag) microsatellite sequences (Ueno et al., 2015) were used for genotyping (Chybicki et al. 2020) using the protocol described in Chybicki et al. (2019). All the permissions required to perform the above actions in protected areas were received from dedicated Regional Directorates for Environmental Protection according to the Polish Environmental law.

## 2.2 Data analysis

### 2.2.1 Genetic diversity

The genetic diversity was characterized by the number of alleles ($A$), observed heterozygosity ($H_o$), and expected heterozygosity ($H_e$) using the INEST 2.2 software (Chybicki and Burczyk 2009) and the allelic richness ($AR$) using FSTAT 2.9.4 (Goudet 1995). For each population, the inbreeding coefficients were estimated jointly with null allele frequencies ($p_{null}$) using the Bayesian approach implemented in the INEST 2.2 software (Chybicki and Burczyk 2009). To determine whether inbreeding and/or null alleles were important factors influencing the observed genotypic structure, the Deviation Information Criterion (DIC) was used to select the best model for a given population. To test the homogeneity in genetic variation among populations, we used the Friedman test (the nonparametric equivalent of the repeated measures ANOVA) with AR and $H_e$ across markers used as independent blocking variables and populations used as independent treatments.

### 2.2.2 Genetic structure

To characterize the overall strength of the genetic structure, we ran the classic Analysis of Molecular Variance (AMOVA; Excoffier et al. 1992). The significance of $F_{ST}$ index was determined based on 10,000 permutations. We also ran the $F_{ST}$-outlier test to verify whether microsatellite markers follow the neutral model of evolution. The null distribution of $F_{ST}$ under neutrality was determined based on 20,000 coalescent simulations. Both tests were performed using Arlequin 3.5 (Excoffier and Lischer, 2010).

Subsequently, using STRUCTURE 2.3.4 (Pritchard et al. 2000), we performed clustering of individual genotypes

**Fig. 1** Results of the STRU
CTURE analysis of *Taxus
baccata* populations. A—Mean
LogPr(K) and standard devia-
tion from 10 runs for each K
from 1 to 15; and distribution
of $\Delta K$ for $K = 1$ to 15; and
B—estimators MEDMEDK,
MEDMEAK, MAXMEDK, and
MAXMEAK for each $K$ from
1–15. C—Individual member-
ship coefficients for trees within
*Taxus baccata* populations
according to the STRUCTU
RE results for $K = 9$. D—Geo-
graphic distribution of genetic
clusters in Poland
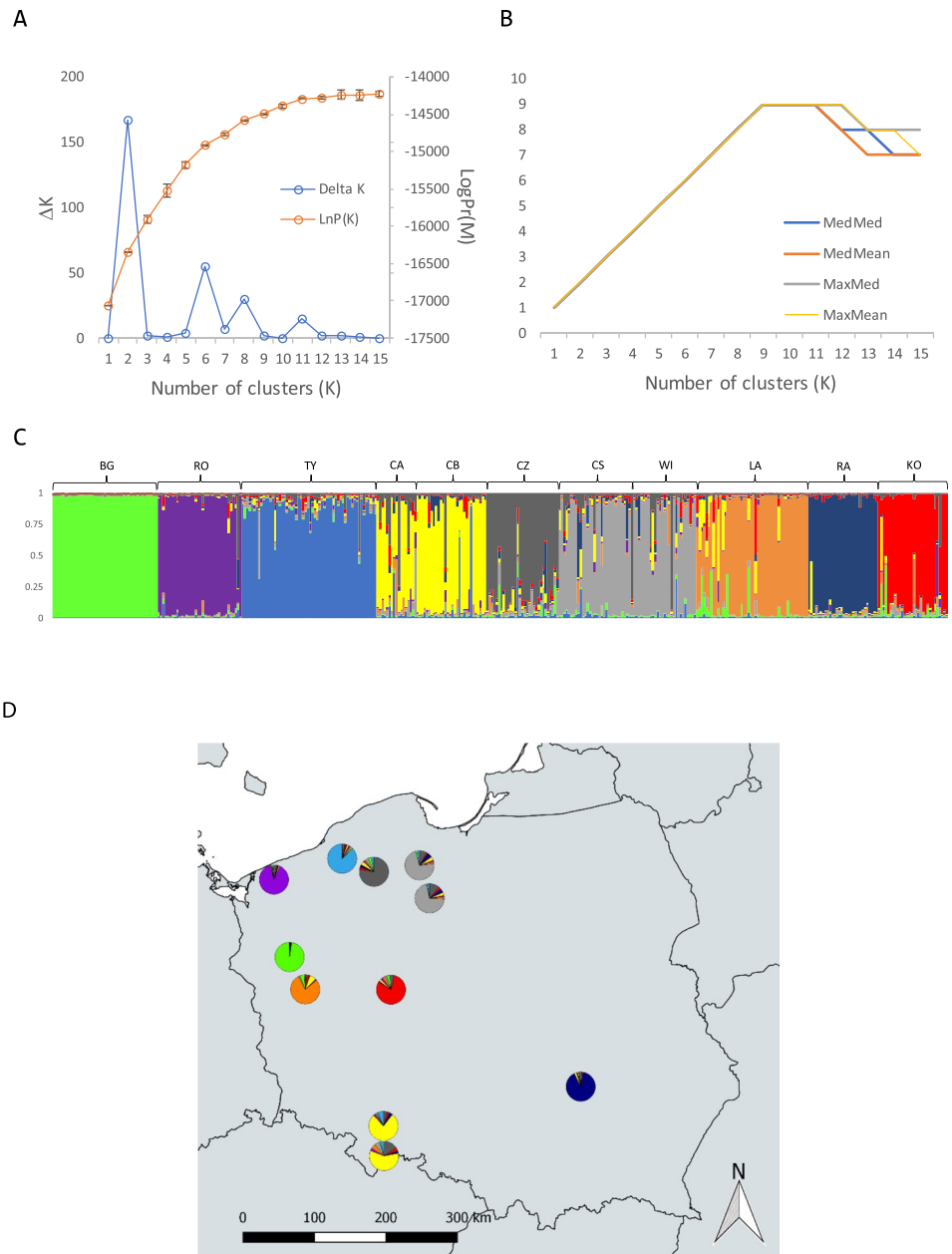


into gene pools, ignoring the information about sampling localities. The number of clusters ($K$) was assumed to vary from 1 to 15. For each $K$, ten runs were performed, with the burn-in length of 250,000 and 250,000 iterations. The best $K$ was determined using the approach described recently in Puechmaille (2016) because the MEDMEDK, MEDMEAK, MAXMEDK, and MAXMEAK estimators tend to be more accurate than the original approach (Pritchard et al. 2000) as well as the $\Delta K$ method (Evanno et al. 2005) for both evenly and unevenly sampled datasets. Puechmaille estima-tors were obtained using the StructureSelector (Li and Liu 2018). The best $K$ was selected as the number of clusters for which Puechmaille's estimators show the maximum value consistently. Finally, the population admixture coefficient

(Admix), an indicator of historical gene flow, was computed under the best $K$ using the Gini-Simpson index.

### 2.2.3 Demographic parameters

The effective population size ($N_e$) was assessed using two esti-mators revealing the best statistical properties out of the avail-able methods (Wang 2016). The first method, implemented in Colony 2.0.6.5 program (Wang 2004), is based on frequencies of full- and half-sib relationships reconstructed using genetic data (SF-$N_e$; Jones and Wang 2010). The second estimator, available in the NeEstimator 2.1 computer program (Do et al. 2014), is based on the level of linkage disequilibrium gener-ated due to genetic drift (LD-$N_e$; Waples 2006).

**Table 1** Characteristics of the study populations of *Taxus baccata*

| Population | Abbreviation | Location | Mean trunk perimeter (cm) (SE) | Sample size | Census number of trees |
|---|---|---|---|---|---|
| Bogdaniec | BG | 52° 41′ N, 15° 04′ E | 48.3 (2.56) | 45 | 2056 |
| Cisowa Góra A | CA | 50° 32′ N, 16° 42′ E | 98.5 (17.50) | 17 | 814 |
| Cisowa Góra B | CB | 50° 32′ N, 16° 42′ E | 97.2 (4.44) | 30 | 167 |
| Czarne | CZ | 53° 44′ N, 16° 58′ E | 87.6 (5.48) | 30 | 439 |
| Czersk | CS | 53° 45′ N, 17° 57′ E | 37.9 (2.56) | 30 | 1200 |
| Kórnik | KO | 52° 14′ N, 17° 05′ E | 95.9 (12.26) | 30 | 1000 |
| Łagów | LA | 52° 18′ N, 15° 16′ E | 85.7 (8.55) | 48 | 157 |
| Radomice | RA | 50° 44′ N, 20° 39′ E | 78.0 (13.38) | 30 | 1005 |
| Rokita | RO | 53° 45′ N, 14° 51′ E | 59.1 (5.09) | 36 | 5500 |
| Tychowo | TY | 53° 56′ N, 16° 20′ E | 95.3 (12.38) | 58 | 70 |
| Wierzchlas | WI | 53° 30′ N, 18° 06′ E | 102.5 (13.30) | 30 | 3559 |

Using the INEST 2.2 software (Chybicki and Burczyk 2009), we tested the study populations for the signs of a genetic bottleneck, using the ratio of the number of alleles ($A$) to the range of allele sizes ($R$) (MR = $A/R$; Garza and Williamson 2001) as an indicator of a genetic bottleneck. The observed MR values were tested against the demographic equilibrium value $MR_{eq}$ using the Wilcoxon signed-rank test. $MR_{eq}$ was computed based on the coalescent simulations (analogously to Cornuet and Luikart 1996), assuming the two-phase mutation model (with the default settings in INEST 2.2). The intensity of the bottleneck was assessed using $\Delta_{MR} = MR_{eq} - MR$. Because of MR < $MR_{eq}$ under bottleneck (Garza and Williamson 2001), the larger $\Delta_{MR}$ value, the higher signal of the bottleneck. The homogeneity in genetic bottleneck signal among populations was evaluated using the Friedman test with $\Delta_{MR}$ across markers used as an independent blocking variable and populations used as independent treatments.

Demographic histories of the studied populations were also characterized using an Approximate Bayesian Computation approach (ABC) implemented in DIYABC 2.1.0 (Cornuet et al. 2014). Due to limited genetic data (as for the robust inference based on the ABC approach), we considered only two simple, a priori equally likely demographic models: either a population evolved without any change in size or a single change in population size was assumed to occur at the time of $t$ generations ago. Each population was analyzed independently, ignoring any gene flow between populations. Based on the estimates of the effective population size obtained independently, the current population size ($N_0$) was assumed to follow the uniform distribution within the range of 4–1000. For the alternative model, we assumed the historical (ancestral) effective size ($N_A$) to follow the uniform distribution within the range of 10–100,000. In addition, for the alternative model, the demographic change was assumed to occur $t$ generations back in the past, with $t$ sampled from the uniform distribution within the range 1–10,000. The default settings were used regarding the mutation model, except for

the average mutation frequency minimum (set to $10^{-5}$) and the individual locus mutation frequency minimum (set to $10^{-6}$). The mean $A$, the mean $H_e$, the mean variance in allele size, and the mean $MR$ were used as summary statistics for the analysis. To approximate the posterior distribution, 2000 best-matching samples out of 1,000,000 random coalescent samples were used. Model probability was estimated using the logistic regression approach. To characterize population demography, we focused on the compound parameters: $N_0\mu$, $N_A\mu$, and $\tau = t\mu$ because their ratios inform about the ratios of the natural parameters, regardless the assumption about the mutation rate $\mu$ (given that the mutation rate $\mu$ is uniform across populations). Parameter estimates were obtained after the log transformation.

#### 2.2.4 Trunk perimeter vs. genetic parameters

Finally, we assessed the relationship between the mean trunk perimeter and the genetic parameters inferred in the above-described analyses. As an indicator of association, we used Spearman's rank correlation because it allows potential non-linear relationships to be taken into account.

## 3 Results

### 3.1 Genetic diversity within populations

The average number of alleles within the study populations ranged from 2.3 (Bogdaniec) to 4.8 (Cisowa Góra B and Czarne), with a mean of 4.1 (Table 2). The allelic richness (based on a minimum of 17 samples) ranged from 2.3 (Bogdaniec) to 4.6 (Cisowa Góra A), with a mean of 3.8. On average, the observed heterozygosity (mean = 0.564, range: 0.508–0.622) was close to the mean expected heterozygosity (mean = 0.564, range: 0.434–0.627) across populations. However, at the population level, we observed homozygosity excess. According to the Bayesian analysis

**INRAE** ✷ Springer

**Table 2** Parameters of the genetic structure of *Taxus baccata* populations, averaged across 18 nuclear microsatellite loci. *A*, number of alleles; *AR*, allelic richness after rarefaction; $H_e$, expected heterozygosity; $H_o$, observed heterozygosity; $p_{null}$, null allele frequency; $F_{IS}$, the multilocus inbreeding coefficient (the posterior median); *95% CI*, 95% highest posterior density interval for the estimate of $F_{IS}$; *Admix*, the admixture coefficient computed based on STRUCTURE results for $K = 9$

| Population | *A* (SE) | AR (SE) | $H_e$ (SE) | $H_o$ (SE) | $p_{null}$ | $F_{IS}$ | 95% CI | Admix |
|---|---|---|---|---|---|---|---|---|
| Bogdaniec | 2.3 (0.958) | 2.3 (0.958) | 0.434 (0.263) | 0.556 (0.356) | 0.000 | 0.000 | NA[2] | 0.053 |
| Cisowa Góra A | 4.6 (2.304) | 4.6 (2.229) | 0.617 (0.141) | 0.554 (0.184) | 0.039 | 0.000 | NA | 0.617 |
| Cisowa Góra B | 4.8 (2.647) | 4.4 (2.235) | 0.565 (0.166) | 0.508 (0.214) | 0.071 | 0.028[1] | 0.000,0.067 | 0.381 |
| Czarne | 4.8 (2.358) | 4.3 (1.690) | 0.591 (0.165) | 0.548 (0.200) | 0.058 | 0.000 | NA | 0.389 |
| Czersk | 4.3 (1.534) | 4.0 (1.404) | 0.597 (0.136) | 0.541 (0.177) | 0.045 | 0.027[1] | 0.000,0.070 | 0.461 |
| Kórnik | 3.9 (1.605) | 3.7 (1.360) | 0.587 (0.137) | 0.622 (0.172) | 0.000 | 0.000 | NA | 0.342 |
| Łagów | 4.2 (1.700) | 3.6 (1.253) | 0.560 (0.152) | 0.592 (0.226) | 0.023 | 0.000 | NA | 0.375 |
| Radomice | 3.9 (1.862) | 3.8 (1.659) | 0.562 (0.164) | 0.551 (0.254) | 0.038 | 0.000 | NA | 0.185 |
| Rokita | 3.2 (1.465) | 3.1 (1.409) | 0.523 (0.186) | 0.572 (0.257) | 0.000 | 0.027[1] | 0.007,0.048 | 0.197 |
| Tychowo | 4.4 (1.580) | 3.7 (1.181) | 0.537 (0.139) | 0.555 (0.183) | 0.028 | 0.000 | NA | 0.244 |
| Wierzchlas | 4.7 (1.742) | 4.3 (1.474) | 0.627 (0.139) | 0.601 (0.190) | 0.030 | 0.000 | NA | 0.485 |
| Overall mean | 4.1 | 4.0 | 0.564 | 0.564 | 0.030 | 0.007 | | 0.340 |

[1]The inbreeding model shows the best fit to the data based on DIC values, [2]estimates not available

results, the excess in homozygosity was mostly due to null alleles (Table 2). Only three populations showed signs of inbreeding, but the estimated inbreeding level corrected for the presence of null alleles was low (between 0.027 and 0.028). Despite the significant contribution of null alleles to the observed genotypic structure, their population-specific mean frequency across markers was rather low, ranging from 0 to 7.2% with a grand mean of 3.0%. With respect to the genetic variation, the study populations deviated significantly from homogeneity in *AR* (*p* value < 0.001) but not in $H_e$ (*p* value = 0.060), as revealed by the Friedman test.

## 3.2 Genetic structure

The AMOVA analysis showed that the genetic variation is significantly structured into among- and within-population components (Table 3), with the differences among populations (sampling locations) explaining c. 11% of the total genetic variation. The $F_{ST}$-outlier test revealed no departures of the microsatellite markers from the null distribution, suggesting that the observed genetic structure was not influenced by selection at any locus (Appendix Fig. 4).

After processing the STRUCTURE analysis results with the Δ*K* approach, *K* = 2 appeared to be the most plausible number of genetic clusters. However, due to the presence of multiple peaks (Fig. 1), the analysis suggested a hierarchical

genetic structure. The MaxMean estimator showed that the most optimal number of clusters *K* = 9. With *K* = 9 as the optimal number of clusters, individuals were grouped into sampling locations, except for Wierzchlas and Czersk (29 km apart). Based on the STRUCTURE results, the admixture coefficient ranged from 0.05 to 0.62, with a mean of 0.34. The level of admixture appeared to vary consistently with the expected heterozygosity.

## 3.3 Demographic parameters

The observed values of *M*-ratio (MR) were below the values expected for the mutation-drift equilibrium ($MR_{eq}$) for all populations (Table 4). The Wilcoxon signed-rank test revealed that the MR deficiency ($MR < MR_{eq}$) was significant in all cases, providing strong support for the presence of bottleneck event(s) in population histories. However, the Friedman test revealed that populations were significantly heterogeneous with respect to bottleneck intensity (*p* value = 0.010).

The SF and the LD estimators appeared to be generally concordant with each other ($r_{Spearman} = 0.836$; *p* value = 0.001), yielding very close harmonic means across populations of 17.1 vs. 17.3. Based on the harmonic mean $N_e$ for the two methods, effective population size ranged from 4.0 (Bogdaniec) to 80.1 (Czersk), with the harmonic mean across populations of 17.2. Interestingly, the two populations (Bogdaniec and Rokita)

**Table 3** Analysis of Molecular Variance (AMOVA) for *Taxus baccata* populations based on sampling locations

| Source of variation | Degrees of freedom | Sum of squares | Variance component | Percentage of explained variation |
|---|---|---|---|---|
| Among populations | 10 | 456.190 | 0.59204 | 11.24 |
| Within populations | 757 | 3539.929 | 4.67626 | 88.76 |
| $F_{ST}$ | 0.112 (*p* value < 0.001) | | | |

**Table 4** Results of the *M*-ratio test for genetic bottleneck and estimates of effective population size of *Taxus baccata* populations. *MR*, the Garza-Williamson statistics (*M*-ratio); $\Delta_{MR}$, the difference between the equilibrium value of MR and the observed value; *p value*, the *p* value for the bottleneck test; *LD-$N_e$*, the linkage disequilibrium-based estimate of effective size; *SF-$N_e$*, the sibling frequency-based effective size; *Mean $N_e$*, the harmonic mean of $N_e$ across two methods; *$N_e/N$*, the ratio of the mean effective to census population size; *95% CI*, the 95% confidence interval

| Population | Genetic bottleneck (the *M*-ratio test) | | | LD-$N_e$ | 95% CI | SF-$N_e$ | 95% CI | Mean $N_e$ | $N_e/N$ |
|---|---|---|---|---|---|---|---|---|---|
| | *MR* | $\Delta_{MR}$ | *p* value | | | | | | |
| Bogdaniec | 0.662 | 0.271 | 0.007 | 23.8 | 13.8,46.0 | 5 | 2.0,20.0 | 8.3 | 0.004 |
| Cisowa Góra A | 0.699 | 0.138 | 0.038 | 37.2 | 21.4,100.0 | 42 | 22.0,100.0 | 39.5 | 0.048 |
| Cisowa Góra B | 0.664 | 0.189 | 0.001 | 31.2 | 14.9,129.3 | 39 | 24.0,70.0 | 34.7 | 0.208 |
| Czarne | 0.674 | 0.179 | 0.001 | 62.7 | 30.6,418.9 | 56 | 34.0,102.0 | 59.2 | 0.135 |
| Czersk | 0.727 | 0.134 | 0.009 | 129.3 | 44.4,+∞ | 58 | 32.0,131.0 | 80.1 | 0.067 |
| Kórnik | 0.683 | 0.196 | 0.001 | 12.2 | 5.2,27.1 | 26 | 15.0,50.0 | 16.6 | 0.017 |
| Łagów | 0.653 | 0.223 | <0.001 | 14.6 | 7.4,28.6 | 27 | 16.0,47.0 | 19.0 | 0.121 |
| Radomice | 0.722 | 0.160 | 0.001 | 37 | 20.3,101.1 | 29 | 17.0,53.0 | 32.5 | 0.032 |
| Rokita | 0.679 | 0.233 | <0.001 | 3.4 | 2.3,7.3 | 5 | 2.0,20.0 | 4.0 | 0.001 |
| Tychowo | 0.669 | 0.204 | <0.001 | 30.1 | 19.8,49.0 | 52 | 34.0,82.0 | 38.1 | 0.545 |
| Wierzchlas | 0.677 | 0.172 | 0.001 | 77.9 | 32.5,+∞ | 42 | 24.0,74.0 | 54.6 | 0.015 |

characterized by the lowest $N_e$ estimates showed also the largest signs of the bottleneck (i.e., the highest values of $\Delta_{MR}$) (Table 4).

Consistently with the MR-based bottleneck test, the coalescent-based ABC approach showed that the scenario with a single change in population size was significantly better than the null model (no change) (Table 5). Although the estimates of demographic parameters obtained under the best model showed wide credible intervals, in all cases, the historical population size was significantly greater than the current population size. The UPGMA clustering of populations based on the Euclidean distance between (log-transformed) posterior modes of coalescent parameters revealed the existence of two main demographic groups (Fig. 2A). The first group (5 populations) was characterized by a very recent

genetic bottleneck (Fig. 2C). The second group (6 populations) embraced populations characterized by a moderate-to-high time to bottleneck. Concerning the time to bottleneck event, populations in the second group were characterized by quite heterogeneous demographic histories, as suggested in Fig. 2C, although the two subgroups had insufficient bootstrap support of 70%. The linear regression analysis for the relationship between the estimate of $N_0\mu$ and the harmonic mean $N_e$ showed that the slope was significantly greater than zero ($b_1 = 0.0048$; SE = 0.0019) while the intercept was non-significant ($b_0 = 0.0778$; SE = 0.0776) (Fig. 2B). Using the regression slope $b_1$ as the estimate of mutation rate $\mu$ to re-scale the estimated demographic history parameters (Table 5), we found that the most recent bottleneck occurred

**Table 5** Demographic histories of *Taxus baccata* populations inferred using the Approximate Bayesian Computing approach. *Pr(const. $N_e$)*, the posterior probability of the scenario of constant population size, *Pr(var. $N_e$)*, the posterior probability of the scenario of varying population size; *$N_0\mu$*, the posterior estimates of the current effective population size multiplied by the mean mutation rate $\mu$, *$N_A\mu$*, the posterior estimates of the effective population size of the ancestral population; $\tau = t\mu$, the mutation-scaled number of generations backward in time to the demographic event; *95% CI*, the 95% credible interval for estimates

| Population | Pr(const. $N_e$) | Pr(var. $N_e$) | Mode | $N_0\mu$ | | | $N_A\mu$ | | | $\tau$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Median | 95% CI | Mode | Median | 95% CI | Mode | Median | 95% CI | |
| Bogdaniec | 0.001 | 0.999 | 0.013 | 0.014 | 0.004,0.035 | 0.70 | 1.47 | 0.249,8.32 | 0.008 | 0.015 | 0.002,0.099 | |
| Cisowa Góra A | 0.263 | 0.737 | 0.480 | 0.480 | 0.099,0.813 | 2.29 | 21.0 | 1.22,71 | 1.17 | 1.81 | 0.042,6.75 | |
| Cisowa Góra B | 0.193 | 0.807 | 0.477 | 0.502 | 0.216,0.807 | 5.48 | 27.7 | 1.82,85.1 | 1.95 | 2.67 | 0.219,8.57 | |
| Czarne | 0.026 | 0.974 | 0.429 | 0.439 | 0.114,0.758 | 2.74 | 21.6 | 1.71,74.5 | 1.44 | 1.72 | 0.082,6.69 | |
| Czersk | 0.05 | 0.951 | 0.275 | 0.279 | 0.101,0.518 | 1.61 | 2.56 | 0.323,6.94 | 0.079 | 0.409 | 0.023,3.06 | |
| Kórnik | 0.255 | 0.746 | 0.070 | 0.086 | 0.017,0.232 | 1.86 | 4.63 | 0.494,22.8 | 0.038 | 0.087 | 0.005,0.62 | |
| Łagów | 0.005 | 0.995 | 0.127 | 0.142 | 0.038,0.351 | 3.10 | 7.8 | 0.938,36.5 | 0.090 | 0.223 | 0.022,1.51 | |
| Radomice | 0.153 | 0.847 | 0.242 | 0.261 | 0.091,0.481 | 1.82 | 14.6 | 0.664,51.6 | 0.656 | 1.06 | 0.054,4.14 | |
| Rokita | 0.353 | 0.647 | 0.037 | 0.046 | 0.012,0.125 | 1.36 | 3.15 | 0.431,16.8 | 0.021 | 0.052 | 0.005,0.380 | |
| Tychowo | 0.312 | 0.688 | 0.266 | 0.263 | 0.089,0.526 | 3.54 | 14.3 | 1.44,59.2 | 0.427 | 0.769 | 0.084,3.81 | |
| Wierzchlas | 0.354 | 0.646 | 0.275 | 0.271 | 0.064,0.543 | 1.71 | 11.6 | 0.747,45.1 | 0.272 | 0.606 | 0.026,3.15 | |
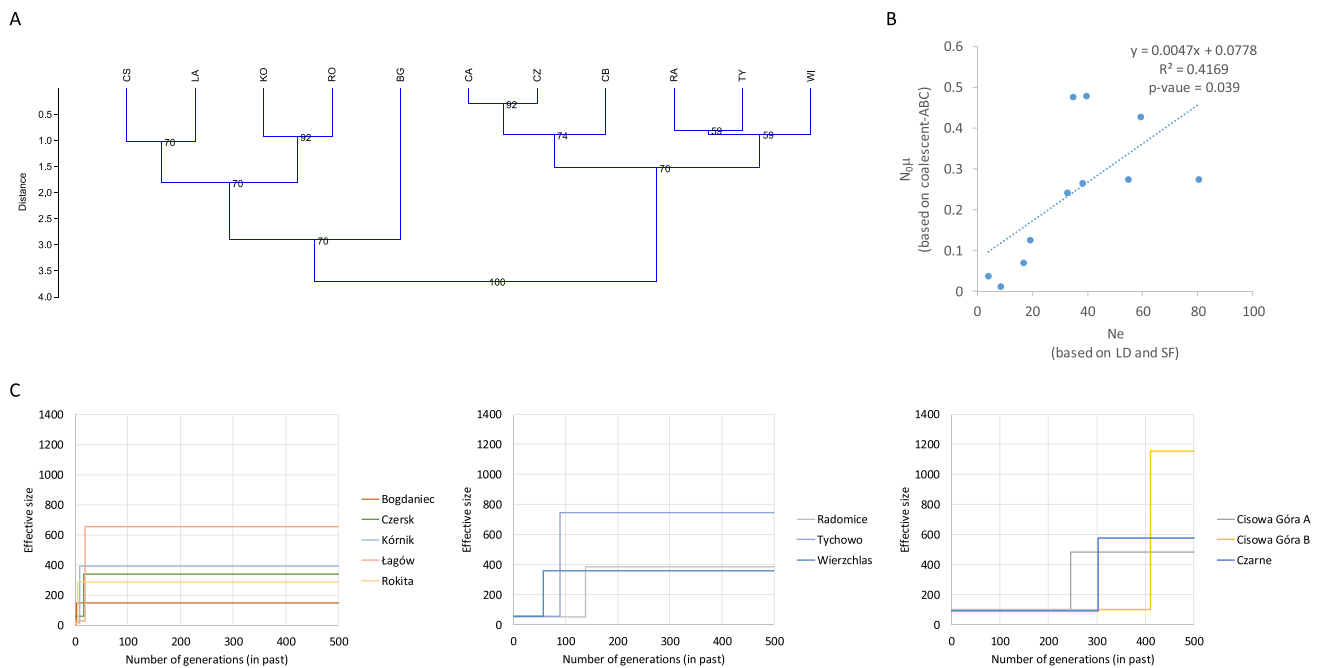
**Fig. 2** Demographic histories of *Taxus baccata* populations inferred with the Approximate Bayesian Computing approach based on coalescent simulations. Panel A—the results of the Euclidean distance-based UPGMA clustering of study populations based on the log-estimates of demographic parameters. Panel B—the relationship between the coalescent-ABC estimate of mutation-scaled current effective size ($N_0\mu$) and the estimate of effective population size obtained as a harmonic mean across linkage disequilibrium (LD) and sibship frequency-based (SF) estimators. Panel C—scale and time of demographic change inferred with the coalescent-ABC approach for the main UPGMA clusters. Scales of axes (effective size and number of generations) obtained after re-scaling the estimated coalescent parameters ($N_0\mu$, $N_A\mu$, and $t\mu$) using the mutation rate estimated as a slope of regression line shown in the panel B

approximately 1.7 generations ago in Bogdaniec, while the most historical bottleneck took place 411 generations ago in Cisowa Góra B (Fig. 2C). The estimated ancestral population size ranged between 147 (Bogdaniec) and 1154 (Cisowa Góra B) with the harmonic mean of 383.

The genetic structure parameters (AR, $H_e$, and Admix) appeared to co-vary quite well with the mean trunk perimeter (Fig. 3). With the remarkable exception of Czersk (CS), study populations with a greater perimeter tended to show a greater allelic richness as well as genetic diversity ($H_e$) and admixture of gene pool (Admix). The CS population appeared to represent an apparent outlier also in correlations between the trunk perimeter and demographic parameters (Fig. 3). After ignoring CS, the trunk perimeter appeared to correlate significantly (positively) with all genetic structure parameters (AR, $H_e$, and Admix) and the following demographic parameters: negatively with the indicator of bottleneck $\Delta_{MR}$ and positively with $N_e$, $N_0\mu$, and $N_0/N_A$.

## 4 Discussion

### 4.1 The main considerations

The study populations showed high genetic differentiation, with gene pools corresponding mostly to sampling locations. Also, they showed a significant but variable signal of a genetic bottleneck. Generally, except for a single population, populations characterized by a greater trunk perimeter were genetically more variable and they tended to have a greater effective population size. Moreover, populations with a greater trunk perimeter tended to show a lower deficiency in the Garza-Williamson index and a higher ratio of current to historical population size, suggesting that younger populations are characterized by stronger genetic bottleneck effects.

In this study, the mean trunk perimeter (girth) was used as a surrogate of tree age. Generally, the tree girth is a measure of radial tree growth that reflects a combined effect of the age (as a primary factor) and the annual growth rate (as a confounding factor), which may differ both among sites and among individuals within a single site (e.g., George et al. 2019). While the impact of individual variation in growth rates can be reduced via averaging, the impact of among-site variation cannot be neglected. Such differences in tree-ring widths were observed among the three populations included in this study (see Cedro and Iszkuło 2011), and attributed to temperature and precipitation. In addition, growth rates may also vary due to light conditions (Devaney et al. 2015), which seem to be different in the study sites due to forest structure differences (pers. obs.). Therefore, tree girth is not a perfect measure of tree age. On the other hand, the perimeter of a tree

**Fig. 3** The relationship between the mean trunk perimeter and genetic parameters in *Taxus baccata* populations. AR, the mean allelic richness; $H_e$, the mean expected heterozygosity; Admix, the admixture rate inferred based on STRUCTURE (for $K=9$); $\Delta_{MR}$, the difference between expected (under demographic stability) and observed value of the Garza-Williamson bottleneck index (MR); $N_e$, the effective population size (the harmonic mean across 2 estimators); $N_e/N$, the ratio of effective to census population size; DIY-ABC parameter estimates: $N_0\mu$, the current (mutation-scaled) effective size; $N_0/N_A$, the ratio of current to ancestral (before demographic event) effective size; $\tau$, the time to demographic event backward in time (in generations × mutation rate). $p$ values show the significance of the Spearman correlation analysis performed with the Czersk population being included and excluded in the analysis
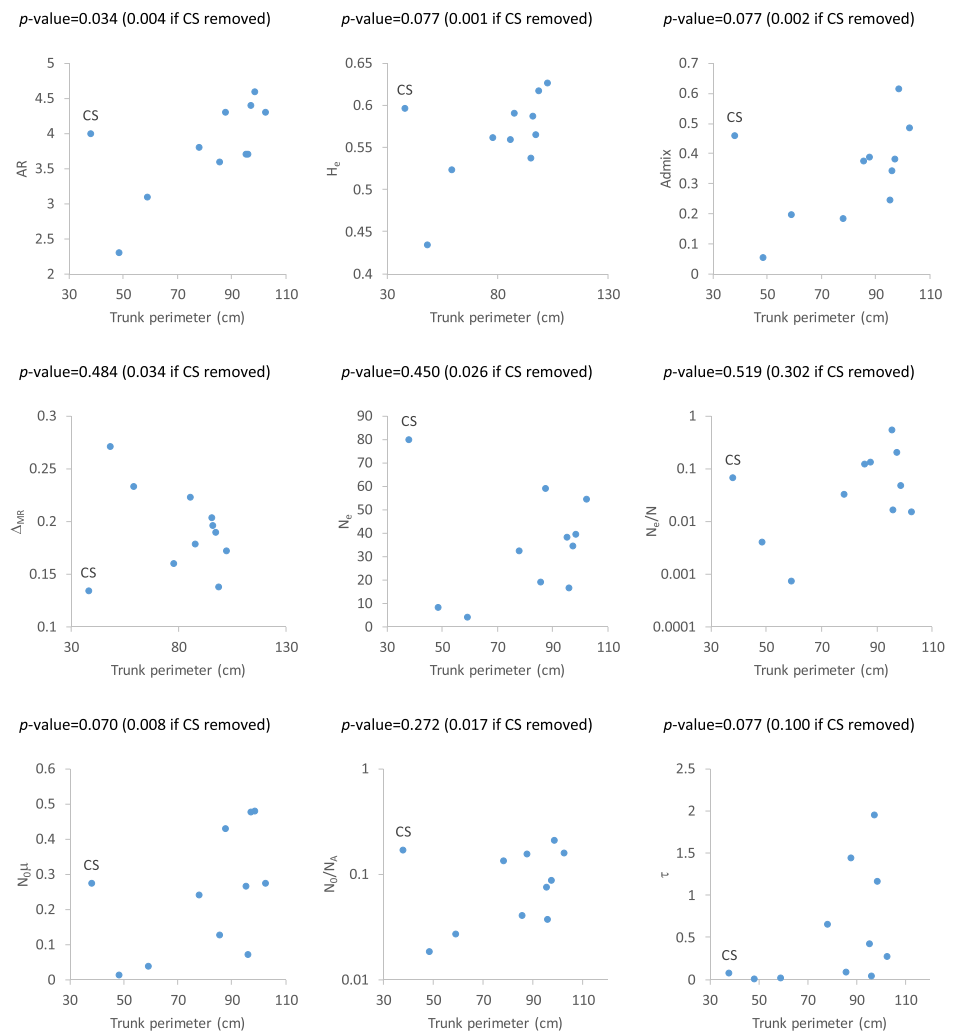


trunk can be measured with high accuracy, while an accurate estimation of tree age poses a real challenge.

Besides ecophysiological factors (De Swaef et al. 2015), a variable response of trees to their environments due to variation in genetically-based preadaptation can shape the trunk perimeter's inter-site variation (Lu et al. 2014). Assuming that a greater trunk perimeter reflects better individual adaptation (more efficient resource acquisition), natural selection may potentially explain the observed relationship between tree girth and genetic parameters. Keeping in mind that the microsatellite markers used in this study were developed based on the transcriptome sequence (Ueno et al. 2015), natural selection could directly impact the pattern of genetic diversity, especially through heterozygosity advantage (Ledig et al. 1983; Bush et al. 1987; Ellis and Burke 2007). However, according to the $F_{ST}$-outlier test, the genetic markers did not show any traces of natural selection, making the selection rather an unreliable factor, especially since it would have to act at most loci simultaneously. Therefore, taking the abovementioned limitations into account, we believe that the relationship between genetic

parameters and the perimeter observed in this study reflects the effect associated with the mean tree age.

The observed patterns of genetic diversity followed quite well the predictions for a meta-population with extinction-recolonization dynamics. Hence, the results imply that the species' genetic structure is shaped by a balanced impact of gene flow and turnover dynamics. The long-term scattered distribution coupled with a limited dispersal potential (relative to spatial isolation; Chybicki and Oleksa 2018) suggests that gene flow among yew populations could be low. Such a conclusion is in line with the findings of earlier studies (Myking et al. 2009; Dubreuil et al. 2010; González-Martínez et al. 2010; Chybicki et al. 2012). According to the STRUCTURE results, older populations showed higher admixture levels suggesting that populations have exchanged genes since their establishment. On the other hand, young populations (except for the outlier) showed homogeneous gene pools, with little or no sign of gene flow. Assuming constantly low gene exchange over generations, the observed "population age vs. genetic diversity" relationship could evolve under low extinction rates to

enable older populations to approach the migration-drift equilibrium. It seems that exceptional yew longevity and ability to regenerate from branches (Thomas and Polwart 2003) may facilitate slow turnover dynamics. Alternatively, relatively high genetic diversity, admixture, and effective size observed in old populations could directly result from greater species abundance in the Subboreal. In such conditions, suitable habitats could be colonized with seeds from numerous, potentially large source populations, leading to the pattern of high initial genetic diversity and weak founder effect (Raffl et al. 2008; Pluess 2011; Elleouet and Aitken 2019). Consequently, it cannot be excluded that the observed genetic diversity of old populations is mostly reminiscent of the ancestral gene pool dated back to the optimum of species abundance, and represents a quasi-stationary maximum point due to limited gene flow.

The long-lasting species decline, proceeding since the Subboreal, is generally well reflected in demographic history results. All populations showed signs of a genetic bottleneck, as revealed by a simple test with the Garza-Williamson index. The post hoc clustering of DIYABC results showed that differences in the inferred demographic histories between populations were mostly due to bottleneck time and not the ancestral population's size. However, as the time to demographic event was not related to trunk perimeter, older populations did not necessarily show a longer time to bottleneck. On the other hand, the positive association between trunk perimeter and the ratio of current to ancestral population size suggests that older populations tended to reveal lower bottleneck intensity. It should be mentioned that the DIYABC results were obtained based on a very simple demographic model (single demographic event, no gene flow), so any conclusions regarding demographic histories should be drawn with caution. Nonetheless, when coupled with the results on genetic diversity and effective population size, they support the scenario that older populations are better representatives of the ancestral gene pool than recently established ones.

Our results showed that, in the current meta-population of *T. baccata*, old populations seem to preserve enough genetic variation to be sources of gene diversity for newly colonized sites. This is important especially for the sites revealing strong founder effect such as Rokita and Bogdaniec. These two populations revealed not only the lowest genetic diversity and effective size but also the apparent excess in heterozygosity, which can be attributed to a small number of breeders (Balloux 2004). Interestingly, despite the low genetic variation, high census numbers in Rokita and Bogdaniec demonstrate that these populations do not show signs of population decline. It is not clear, however, how the populations will respond in the long-term. According to our earlier study, Rokita showed an excess of biparental inbreeding at the seed stage, beyond the predictions based on simple ecological determinants such as population density and trunk perimeter (Chybicki et al. 2019). Thus, without a recurrent gene flow, they are potentially at risk of inbreeding depression. However, parentage analysis at the seedling stage demonstrated that the dispersal potential of yew may limit the increase in genetic diversity, especially in the context of the species' strong spatial isolation (Chybicki and Oleksa 2018). Therefore, comparative assessment of the genetic variation of successive generations is a pending goal of our future study.

### 4.2 The outlier population

Czersk was characterized by high genetic variation and large effective population size despite being relatively young, which makes it an outlier population in terms of the relationship between population age and genetic diversity. According to the STRUCTURE analysis, Czersk clustered with Wierzchlas, which is well known as one of the oldest and the largest (if trees of age > 100 years are counted) yew remnants in Europe. Such a genetic homogeneity suggests either intensive gene flow between these populations or an artificial origin of Czersk. According to predictions on pollen and seed dispersal (Chybicki and Oleksa 2018), it is very unlikely that Czersk and Wierzchlas form a panmictic unit, given that pollen needs to overcome both the distance of 29 km and the potential barrier of the forest canopy. On the other hand, the forest in Czersk is characterized by a high tree species richness, and especially the presence of (non-native) Douglas fir, suggesting that the stand was planted. Our genetic data, collected in this and the previous study (Chybicki et al. 2012), show that there is no spatial genetic structure in Czersk (Appendix Fig. 5), supporting an artificial origin. If the yew population was established by humans, the plant material was almost surely taken from Wierzchlas, which is the closest yew remnant and a protected nature reserve since the 1820s (Tobolski 2002). The high genetic variation in Czersk suggests that the plant material was collected, representatively, as if the goal was to protect gene resources of the Wierzchlas reserve. According to the reserve documentation, the yew decline in Wierzchlas has been noted for the last 100 years (from 5500 in 1910 to < 3000 living trees at present). Hence, it might be reasonable to consider there has been the need for ex-situ conservation for a long time, but we cannot go beyond speculations whether such efforts were actually made.

## 5 Conclusions

Previous studies showed that yew populations characterized by the large census number unnecessarily represent the richest source of genetic variation (Chybicki et al. 2012; Litkowiec et al. 2018). Our study demonstrated that the trunk
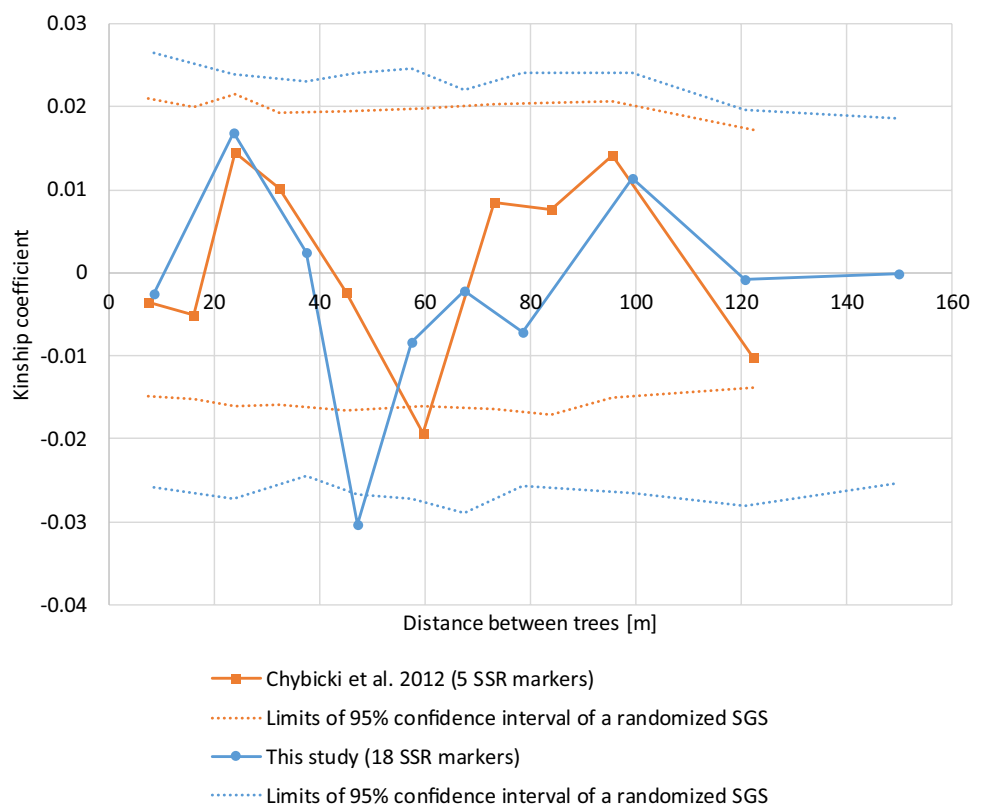
perimeter, which is a readily available tree measure, can be used as a proxy for genetic variation in yew populations. We also showed that the perimeter could be used as a simple indicator of demographic history in the study species. Revealing connection between population history and genetic parameters, the study showed that genetic structure in *T. baccata* is in agreement with predictions for a metapopulation with extinction-recolonization dynamics, at least in areas of historically scattered species distribution (cf. Gargiulo et al. 2019). Keeping in mind that additional studies are required to fully understand the observed pattern, we believe that our results can help optimize conservation efforts and manage the species' genetic resources. Considering the significant genetic differentiation, gene bank collections should optimally include as many populations as possible (Hoban 2019). However, because yew populations do not contribute equally to the overall genetic pool, as revealed in this study, in order to achieve the maximum genetic variation, the sampling intensity across conservation units may be proportional to the mean trunk perimeter. In this way, the ancestral gene pool of the species would likely be captured more representatively. Future studies can focus on the applicability of the observed trend to the other species subjected to historical fragmentation (e.g., *Eucalyptus caesia*; Bezemer et al. 2019).

# Appendix
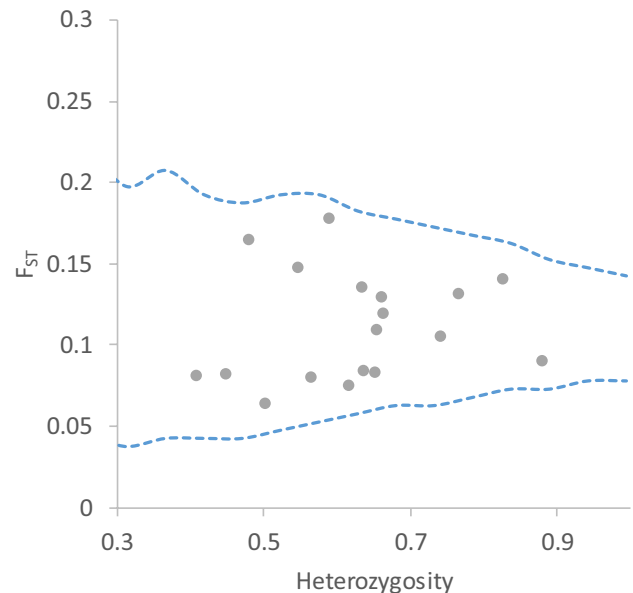


**Fig. 4** The results of the $F_{ST}$-outlier test for neutrality. Dots represent the observed values and dashed lines show the 95% quantiles of the null distribution obtained based on 20,000 coalescent simulations

**Fig. 5** The spatial genetic structure within the Czersk population assessed based on two different data sets. The dataset from Chybicki et al. (2012) included 66 trees genotyped at 5 microsatellite markers. The dataset from this study included 30 trees genotyped at 18 microsatellite markers. As a measure of kinship, the Nason kinship coefficient was used (Loiselle et al., 1995). Limits of the 95% confidence interval for a random spatial genetic structure were computed based on 10,000 random permutations of tree positions using INEST 2.2 software (Chybicki and Burczyk 2009)



Chybicki et al. 2012 (5 SSR markers)

Limits of 95% confidence interval of a randomized SGS

This study (18 SSR markers)

Limits of 95% confidence interval of a randomized SGS

INRAE   ⌂ Springer

**Data availability** Microsatellite data generated in this study are available publicly at https://doi.org/10.5281/zenodo.4153442

## Declarations

**Ethics approval** All the permissions required to perform the actions in protected areas were received from dedicated Regional Directorates for Environmental Protection according to the Polish Environmental law.

**Consent to participate** Not applicable.

**Consent for publication** All authors gave their informed consent to this publication and its content.

**Competing interests** The authors declare no competing interests.

## References

Alleaume-Benharira M, Pen IR, Ronce O (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. J Evol Biol 19:203–215. https://doi.org/10.1111/j.1420-9101.2005.00976.x

Aravanopoulos FA (2016) Conservation and monitoring of tree genetic resources in temperate forests. Curr Forestry Rep 2:119–129. https://doi.org/10.1007/s40725-016-0038-8

Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B (2000) Effects of colonization processes on genetic diversity: differences between annual plants and tree species. Genetics 154(3):1309–1321. https://doi.org/10.1046/j.1523-1739.2003.02365.x

Balloux F (2004) Heterozygote excess in small populations and the heterozygote-excess effective population size. Evolution 58(9):1891–1900. https://doi.org/10.1111/j.0014-3820.2004.tb00477.x

Bezemer N, Krauss SL, Roberts DG, Hopper SD (2019) Conservation of old individual trees and small populations is integral to maintain species' genetic diversity of a historically fragmented woody perennial. Mol Ecol 28:3339–3357. https://doi.org/10.1111/mec.15164

Born C, Kjellberg F, Chevallier MH, Vignes H, Dikangadissi JT, Sanguié J, Wickings EJ, Hossaert-Mckey M (2008) Colonization processes and the maintenance of genetic diversity: insights from a pioneer rainforest tree *Aucoumea Klaineana*. Proc Royal Soc B 275(1647):2171–2179. https://doi.org/10.1098/rspb.2008.0446

Brook BW, Tonkyn DW, O'Grady JJ, Frankham R (2002) Contribution of inbreeding to extinction risk in threatened species. Conserv Ecol 6(1):16. http://www.consecol.org/vol6/iss1/art16/

Burgarella C, Navascués M, Zabal-Aguirre M, Berganzo E, Riba M, Mayol M, Vendramin GG, González-Martínez SC (2012) Recent population decline and selection shape diversity of taxol-related genes. Mol Ecol 21:3006–3021. https://doi.org/10.1111/j.1365-294X.2012.05532.x

Bush RM, Smouse PE, Ledig FT (1987) the fitness consequences of multiple-locus heterozygosity: the relationship between heterozygosity and growth rate in pitch pine (Pinus rigida Mill.). Evolution 41(4):787–798. https://doi.org/10.1111/j.1558-5646.1987.tb05853.x

Chen J, Yang H, Man R et al (2020) Using machine learning to synthesize spatiotemporal data for modelling DBH-height and DBH-height-age relationships in boreal forests. Forest Ecology and Management 466:118104. https://doi.org/10.1016/j.foreco.2020.118104

Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. J Hered 100:106–113. https://doi.org/10.1093/jhered/esn088

Chybicki IJ, Dering M, Iszkuło G, Meyza K, Suszka J (2016) Relative strength of fine-scale spatial genetic structure in paternally vs. biparentally inherited DNA in a dioecious plant depends on both sex proportions and pollen-to-seed dispersal ratio. Heredity 117:449–459

Chybicki IJ, Oleksa A, Kowalkowska K (2012) Variable rates of random genetic drift in protected populations of English yew: implications for gene pool conservation. Conserv Genet 13:899–911

Chybicki IJ, Oleksa A (2018) Seed and pollen gene dispersal in *Taxus baccata*, a dioecious conifer in the face of strong population fragmentation. Ann of Bot 122:409–421. https://doi.org/10.1093/aob/mcy081

Chybicki IJ, Iszkuło G, Suszka J (2019) Bayesian quantification of ecological determinants of outcrossing in natural plant populations: computer simulations and the case study of biparental inbreeding in English yew. Mol Ecol 28:4077–4096. https://doi.org/10.1111/mec.15195

Chybicki IJ, Stefanowska S, Meyza K, & Iszkuło G (2020) Taxus genotype data for "Trunk perimeter correlates with genetic bottleneck intensity and the level of genetic diversity in populations of Taxus baccata L." . Zenodo. V1. https://doi.org/10.5281/zenodo.4153442

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Jean-Michel M, Estoup A (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism DNA Sequence and Microsatellite Data. Bioinformatics 30:1187–1189. https://doi.org/10.1093/bioinformatics/btt763

Cedro A, Iszkuło G (2011) Do females differ from males of European yew (*Taxus baccata L.*) in dendrochronological analysis? Tree Ring Res 67:3–11. https://doi.org/10.3959/2009-9.1

Deforce K, Bastiaens J (2007) The Holocene history of *Taxus baccata* (yew) in Belgium and neighbouring regions. Belg J Bot 140:222–237

De Swaef T, De Schepper V, Vandegehuchte MW, Steppe K (2015) Stem diameter variations as a versatile research tool in ecophysiology. Tree Physiology 35(10):1047–1061. https://doi.org/10.1093/treephys/tpv080

Devaney JL, Whelan PM, Jansen MA (2015) Light responses of yew (*Taxus baccata L.*); does size matter? Trees 29:109–118

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Res 14(1):209–214. https://doi.org/10.1111/1755-0998.12157

Dubreuil M, Riba M, González-Martínez SC, Vendramin GG, Sebastiani F, Mayol M (2010) Genetic effects of chronic habitat fragmentation revisited: strong genetic structure in a temperate tree, *Taxus baccata* (*Taxaceae*), with great dispersal capability. Am J Bot 97:303–310. https://doi.org/10.3732/ajb.0900148

Elleouet JS, Aitken SN (2019) Long-distance pollen dispersal during recent colonization favors a rapid but partial recovery of genetic diversity in *Picea sitchensis*. New Phytol 222(2):1088–1100. https://doi.org/10.1111/nph.15615

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity 92(2):125–132. https://doi.org/10.1038/sj.hdy.6801001

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

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

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

Gargiulo R, Saubin M, Rizzuto G, West B, Fay MF, Kallow S, Trivedi C (2019) Genetic diversity in British populations of *Taxus baccata L.*: is the seedbank collection representative of the genetic variation in the wild? Biol Conserv 233:289–297. https://doi.org/10.1016/j.biocon.2019.01.014

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. Mol Ecol 10:305–318. https://doi.org/10.1046/j.1365-294X.2001.01190.x

George JP, Grabner M, Campelo F, Karanitsch-Ackerl S, Mayer K, Klumpp RT, Schüler S (2019) Intra-specific variation in growth and wood density traits under water-limited conditions: long-term-, short-term-, and sudden responses of four conifer tree species. Sci Total Environ 660:631–643. https://doi.org/10.1016/j.scitotenv.2018.12.478

González-Martínez SC, Dubreuil M, Riba M, Vendramin GG, Sebastiani F, Mayol M (2010) Spatial genetic structure of *Taxus baccata L.* in the western Mediterranean Basin: past and present limits to gene movement over a broad geographic scale. Mol Phylogenet Evol 55:805–815. https://doi.org/10.1016/j.ympev.2010.03.001

Goudet J (1995) Fstat version 1.2: a computer program to calculate F statistics (version 2.9.4). J Hered 86:485–486

Haag CR, Riek M, Hottinger JW, Pajunen VI, Ebert D (2005) Genetic diversity and genetic differentiation in Daphnia metapopulations with subpopulations of known age. Genetics 170(4):1809–1820. https://doi.org/10.1534/genetics.104.036814

Helsen K, Jacquemyn H, Hermy M, Vandepitte K, Honnay O (2013) Rapid buildup of genetic diversity in founder populations of the gynodioecious plant species Origanum vulgare after semi-natural grassland restoration. PLoS ONE 8(6):e67255. https://doi.org/10.1371/journal.pone.0067255

Hilfiker K, Gugerli F, Schütz JP, Rotach P, Holderegger R (2004) Low RAPD variation and female-biased sex ratio indicate genetic drift in small populations of the dioecious conifer *Taxus baccata* in Switzerland. Conserv Genet 5:357–365

Hoban S (2019) New guidance for ex situ gene conservation: sampling realistic population systems and accounting for collection attrition. Biol Conserv 235:199–208. https://doi.org/10.1016/j.biocon.2019.04.013

Ingvarsson PK, Whitlock MC (2000) Heterosis increases the effective migration rate. Proc Royal Soc B 267:1321–1326. https://doi.org/10.1098/rspb.2000.1145

Iszkuło G, Boratyński A (2005) Different age and spatial structure of two spontaneous subpopulations of *Taxus baccata* as a result of various intensity of colonization process. Flora - Morphology, Distribution, Funct Ecol Plants 200(2):195–206. https://doi.org/10.1016/j.flora.2004.03.001

Iszkuło G, Pers-Kamczyc E, Nalepka D, Rabska M, Walas L, Dering M (2016) Postglacial migration dynamics helps to explain current scattered distribution of *Taxus baccata*. Dendrobiology 76:81–89. https://doi.org/10.12657/denbio.076.008

Johnson MT, Prashad CM, Lavoignat M, Saini HS (2018) Contrasting the effects of natural selection, genetic drift and gene flow on urban evolution in white clover (*Trifolium repens*). Proc Royal Soc B 285:20181019. https://doi.org/10.1098/rspb.2018.1019

Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. Mol Ecol Res 10:551–555. https://doi.org/10.1111/j.1755-0998.2009.02787.x

Lander TA, Klein EK, Roig A, Oddou-Muratorio S (2020) Weak founder effects but significant spatial genetic imprint of recent contraction and expansion of European beech populations. Heredity (in Press). https://doi.org/10.1038/s41437-020-00387-5

Lavabre JE, Gilarranz LJ, Fortuna MA, Bascompte J (2016) How does the functional diversity of frugivorous birds shape the spatial pattern of seed dispersal? A case study in a relict plant species. Philos Trans R Soc Lond B Biol Sci 371:20150280. https://doi.org/10.1098/rspb.2015.0280

Ledig FT, Guries RP, Bonefeld B (1983) the relation of growth to heterozygosity in pitch pine. Evolution 37(6):1227–1238

Lefèvre F, Fady B, Fallour-Rubio D, Ghosn D, Bariteau M (2004) Impact of founder population, drift and selection on the genetic diversity of a recently translocated tree population. Heredity 93(6):542–550. https://doi.org/10.1038/sj.hdy.6800549

Li YL, Liu JX (2018) StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. Mol Ecol Res 18:176–177. https://doi.org/10.1111/1755-0998.12719

Linares JC (2013) Shifting limiting factors for population dynamics and conservation status of the endangered English yew (*Taxus baccata* L., Taxaceae). Forest Ecol Manage 291:119–127. https://doi.org/10.1016/j.foreco.2012.11.009

Litkowiec M, Lewandowski A, Wachowiak W (2018) Genetic variation in *Taxus baccata L.*: a case study supporting Poland's protection and restoration program. Forest Ecol Manage 409:148–160. https://doi.org/10.1016/j.foreco.2017.11.026

Litkowiec M, Plitta-Michalak BP, Lewandowski A, Iszkuło G (2015) Homogenous genetic structure in populations of *Taxus baccata* with varied proportions of male and female individuals. Silva Fenn 49:1236

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (*Rubiaceae*). Am J Bot 82:1420–1425

Lu P, Parker WH, Cherry M, Colombo S, Parker WC, Man R, Roubal N (2014) Survival and growth patterns of white spruce (*Picea glauca* [Moench] Voss) rangewide provenances and their implications for climate change adaptation. Ecolo Evol 4(12):2360–2374. https://doi.org/10.1002/ece3.1100

McCauley DE, Raveill J, Antonovics J (1995) Local founding events as determinants of genetic structure in a plant metapopulation. Heredity 75(6):630–636. https://doi.org/10.1038/hdy.1995.182

Myking T, Vakkari P, Skrøppa T (2009) Genetic variation in northern marginal *Taxus baccata L.* populations. Implications for

Conservation Forestry 82:529–539. https://doi.org/10.1093/forestry/cpp022

Pelletier F, Coltman DW (2018) Will human influences on evolutionary dynamics in the wild pervade the Anthropocene? BMC Biol 16:7

Pluess AR (2011) Pursuing glacier retreat: genetic structure of a rapidly expanding *Larix decidua* population. Mol Ecol 20(3):473–485. https://doi.org/10.1111/j.1365-294X.2010.04972.x

Powolny M, Poschlod P, Reisch C (2016) Genetic variation in *Silene acaulis* increases with population age. Botany 94(4):241–247. https://doi.org/10.1139/cjb-2015-0195

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Puechmaille SJ (2016) The program structure does not reliably recover the correct population structure when sampling is uneven: sub-sampling and new estimators alleviate the problem. Mol Ecol Res 16:608–627. https://doi.org/10.1111/1755-0998.12512

Raffl C, Holderegger R, Parson W, Erschbamer B (2008) Patterns in genetic diversity of *Trifolium pallescens* populations do not reflect chronosequence on alpine glacier forelands. Heredity 100:526–532. https://doi.org/10.1038/hdy.2008.8

Ralls K, Ballou JD, Dudash MR, Eldridge MD, Fenster CB, Lacy RC, Sunnucks P, Frankham R (2018) Call for a paradigm shift in the genetic management of fragmented populations. Conserv Lett 11:e12412. https://doi.org/10.1111/conl.12412

Schneller JJ, Holderegger R (1996) Genetic variation in small, isolated fern populations. J Veg Sci 7(1):113–120. https://doi.org/10.2307/3236423

Swindell WR, Bouzat JL (2006) Gene flow and adaptive potential in *Drosophila melanogaster*. Conserv Genet 7:79–89

Thomas PA, Garcia-Martínez X (2015) Response of European yews to climate change: a review. Forest Systems 24:1–11. https://doi.org/10.5424/fs/2015243-07465

Thomas PA, Polwart A (2003) Taxus baccata L. J Ecol 91:489–524

Tobolski K (2002) Pomijana tematyka badawcza rezerwatu „Cisy Staropolskie im. Leona Wyczółkowskiego" koło Wierzchlasu: Park Narodowy „Bory Tucholskie" na tle projektowanego rezerwatu biosfery (ed. by J Banaszczak, K Tobolski) Park Narodowy Bory Tucholskie, Charzykowy, Poland, pp. 165–194.

Ueno S, Wen Y, Tsumura Y (2015) Development of EST-SSR markers for *Taxus cuspidata* from publicly available transcriptome sequences. Biochem Syst Ecol 63:20–26. https://doi.org/10.1016/j.bse.2015.09.016

Uzquiano P, Allué E, Antolín F et al (2015) All about yew: on the trail of *Taxus baccata* in southwest Europe by means of integrated palaeobotanical and archaeobotanical studies. Veg Hist Archaeobot 24:229–247. https://doi.org/10.1007/s00334-014-0475-x

Wang J (2004) Sibship reconstruction from genetic data with typing errors. Genetics 166:1963–1979

Wang J (2016) A comparison of single-sample estimators of effective population sizes from genetic marker data. Mol Ecol 25:4692–4711. https://doi.org/10.1111/mec.13725

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conserv Genet 7:167–184

Whitlock MC (1992) Temporal fluctuations in demographic parameters and the genetic variance among populations. Evolution 46(3):608–615. https://doi.org/10.1111/j.1558-5646.1992.tb02069.x

Willi Y, Van Buskirk J, Hoffmann AA (2006) Limits to the adaptive potential of small populations. Annu Rev Ecol Evol Syst 37:433–458