

RESEARCH ARTICLE

Open Access



Recent demographic histories of temperate deciduous trees inferred from microsatellite markers

Yu Cao¹, Da-Yong Zhang¹, Yan-Fei Zeng² and Wei-Ning Bai^{1*}

Abstract

Background: Accurate inference of demographic histories for temperate tree species can aid our understanding of current climate change as a driver of evolution. Microsatellites are more suitable for inferring recent historical events due to their high mutation rates. However, most programs analyzing microsatellite data assume a strict stepwise mutation model (SMM), which could cause false detection of population shrinkage when microsatellite mutation does not follow SMM.

Results: This study aims to reconstruct the recent demographic histories of five cool-temperate tree species in Eastern Asia, *Quercus mongolica*, *Q. liaotungensis*, *Juglans cathayensis*, *J. mandshurica* and *J. ailantifolia*, by using 19 microsatellite markers with two methods considering generalized stepwise mutation model (GSM) (MIGRAINE and VarEff). Both programs revealed that all the five species experienced expansions after the Last Glacial Maximum (LGM). Within butternuts, *J. cathayensis* experienced a more serious bottleneck than the other species, and within oaks, *Q. mongolica* showed a moderate increase in population size and remained stable after the expansion. In addition, the point estimates of the multistep mutation proportion in the GSM model (p_{GSM}) for all five species were between 0.50 and 0.65, indicating that when inferring population demographic history of the cool-temperate forest species using microsatellite markers, it is better to assume a GSM rather than a SMM.

Conclusions: This study provides the first direct evidence that five cool-temperate tree species in East Asia have experienced expansions after the LGM with microsatellite data. Considering the mutation model of microsatellite has a vital influence on demographic inference, combining multiple programs such as MIGRAINE and VarEff can effectively reduce errors caused by inappropriate model selection and prior setting.

Keywords: Asian butternuts, Demographic inference, Effective population size, MIGRAINE, Oaks, VarEff

Background

Understanding the demographic history of species is a central issue in evolutionary ecology. At present, it is widely believed that repeated Quaternary glacial cycles

since the Cenozoic, especially the Last Glacial Maximum (LGM, 20–18 kya [1]) [2–4], have strongly influenced the distribution and population dynamics of temperate tree species in the northern hemisphere [5–8]. In recent years, whole-genome sequencing coupled with new methods, such as the Sequentially Markovian Coalescent (SMC) based methods, e.g., the Pairwise Sequentially Markovian Coalescent (PSMC) model [9] and the Multiple Sequentially Markovian Coalescent (MSMC) model [10], and the site frequency spectrum (SFS) based method, e.g., stairway plot [11], contributes to illustrate

*Correspondence: baiwn@bnu.edu.cn

¹ State Key Laboratory of Earth Surface Process and Resource Ecology and Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing 100875, China

Full list of author information is available at the end of the article



the patterns of population demography. However, the SMC methods cannot efficiently reveal the continuous changes of population size within the last ten thousand years [12] and the SFS methods cannot provide an accurate estimation of population demography if the sample size is limited. However, sample size may be necessarily small for the projects of whole-genome sequencing in the non-model organisms [13]. Thus, to accurately predict recent population dynamics is still a major challenge for evolutionary ecologists, which requires molecular markers with much more polymorphisms. Microsatellites may be an option because of their high mutation rate [14, 15] of 10^{-6} – 10^{-2} per generation per locus [16, 17] and relatively selective neutrality [18, 19]. Moreover, considering there are thousands of studies that have employed microsatellite markers to study species evolution histories in the last two decades, it provides us opportunity to study population demography from these existing data.

Microsatellites have been widely used in population dynamics since the beginning of this century. Before the likelihood-based methods have been developed, most moment-based programs relied on summary statistics of genetic data and tested departure from theoretical distributions under a given demographic and mutational model, such as BOTTLENECK [20] and M-RATIO [21], which can only detect population declines. However, the moment-based methods suffer from a limited statistical power because they do not give any estimate of the severity and duration of the bottleneck. Likelihood-based methods such as MSVAR [22, 23] and MIGRATE [24] coupled with Monte Carlo sampling offer a powerful alternative to the moment-based methods [25–27], but they only take into account the strict stepwise mutation models (SMMs) [28], although many models have been developed to describe microsatellite mutation mechanisms [29]. It has been widely recognized that deviations from an SMM can cause false detection of population shrinkage [22, 30]. For these reasons, Leblois et al. [31] developed a program (MIGRAINE) which used a maximum-likelihood method to infer the past changes of population size with a generalized stepwise mutation model (GSM, [32]). Additionally, Nikolic and Chevalet [33] developed another program, VarEff, which used a composite-likelihood method to infer transient changes in population size in the past with multiple mutation models.

Ecologically dominant forest trees with wide distribution provide an excellent opportunity to explore the influences of climate and geography on their demographic histories [34]. In this study, we chose five temperate tree species distributed in East Asia, two oak species, *Quercus mongolica* and *Q. liaotungensis*, and three Asian butternut species, *Juglans cathayensis*, *J. mandshurica* and *J. ailantifolia*, to detect their recent population dynamics

with microsatellite data. All five species are wind-pollinated, and have low to moderate genetic differentiation among populations [35–37]. For oaks, *Quercus mongolica* is mainly distributed in northeastern China [38], whereas *Q. liaotungensis* is divided into two clades, a northeast group scattered in the Changbai Mts. and partially distributed in North Korea and the Far East of Russia, and a northwest group mainly distributed in the Qinling, Liupan and Lüliang Mts. in northern China [39]. The historical population dynamics of *Q. mongolica* and *Q. liaotungensis* has not yet been reported. For butternuts, *J. mandshurica* is distributed in northeastern China, *J. cathayensis* occurs in southern China and Taiwan Island, and *J. ailantifolia* is only distributed in Japan [40]. Bai et al. [41] had estimated the divergence time and ancestral effective population size using DIYABC approach and microsatellite data, but the estimation of posterior parameters was greatly influenced by prior value setting [42, 43]. Later, Bai et al. [44] applied the PSMC method to the whole genome resequencing data to infer butternut population demography, but PSMC cannot reveal the population changes within the last ten thousand years because few coalescent events are expected to have occurred during this time period [9].

The main goal of this study was to reconstruct the recent demographic histories of the above-mentioned five temperate tree species with microsatellite data. To that end, we applied two methods implementing a GSM, i.e., MIGRAINE, which is based on the maximum likelihood method, and VarEff, which is based on approximate likelihood. Then, we assessed the effects of microsatellite mutation models and the multistep mutation proportions in the GSM on population demography inference and compared the advantages and disadvantages of the two programs. Our research can help to understand the recent demographic histories of temperate trees as well as shed some light on how microsatellite data can be used to analyze recent population size changes.

Results

Effective population size fluctuation

Both MIGRAINE and VarEff were used to estimate the current and past effective population sizes and the N_e changing times in the two species of oaks and three species of Asian butternuts (see Table 1 and Additional file 1: Table S1). MIGRAINE and VarEff showed similar trends in the population dynamic curves for the five species, in which they experienced a rapid population expansion after LGM (Figs. 1, 2). However, the specific expansion times estimated from MIGRAINE and VarEff were different. Considering MIGRAINE can only deal with simple model and provide a wide confidence interval (Table 1), it would give a less precise value than

Table 1 Comparison of population parameters estimated using MIGRAINE and VarEff

Populations (sample size)	Methods	p_{GSM}	θ	θ_{anc}	$D = G \times \mu$	$N_{ratio} = \theta/\theta_{anc}$	$N (\times 10^4)$	$N_{anc} (\times 10^4)$	$T_{years} (kya)$
Northeast <i>Q. liaotungensis</i> (139)	MIGRAINE	0.53	18.58	5.37	0.67	3.46	0.46	0.13	33.65
	VarEff		12.50	4.79		2.61	0.31	0.12	
Northwest <i>Q. liaotungensis</i> (240)	MIGRAINE	0.58	21.47	2.48	1.32	8.67	0.54	0.06	66.15
	VarEff		17.40	2.46		7.08	0.43	0.06	
<i>Q. mongolica</i> (502)	MIGRAINE	0.62	20.88	3.52	1.19	5.93	0.52	0.09	59.45
	VarEff		12.23	3.24		3.78	0.31	0.08	
<i>J. cathayensis</i> (596)	MIGRAINE	0.64	14.76	0.64	2.95	23.21	0.37	0.02	88.41
	VarEff		17.44	1.11		15.72	0.44	0.03	
<i>J. mandshurica</i> (399)	MIGRAINE	0.62	15.65	4.35	0.99	3.60	0.39	0.11	29.64
	VarEff		19.30	4.01		4.81	0.48	0.10	
<i>J. ailantifolia</i> (107)	MIGRAINE	0.62	18.12	1.29	1.80	14.08	0.45	0.03	54.03
	VarEff		19.75	3.41		5.79	0.49	0.09	

The estimates of past and current population sizes and times in years (T_{years}) obtained by MIGRAINE and VarEff are converted with a fixed mutation rate of 10^{-3} mutation per locus per generation and a generation time of 50 years for oaks, 30 years for Asian butternuts. The population size calculated by VarEff is median. θ , scaled current effective population size by mutation rate; θ_{anc} , scaled ancestor effective population size by mutation rate; $N_{ratio} = \theta/\theta_{anc}$; N , current effective population size (individual number); N_{anc} , ancestor effective population size (individual number); G , time measured by generations; μ , mutation rate per locus per generation

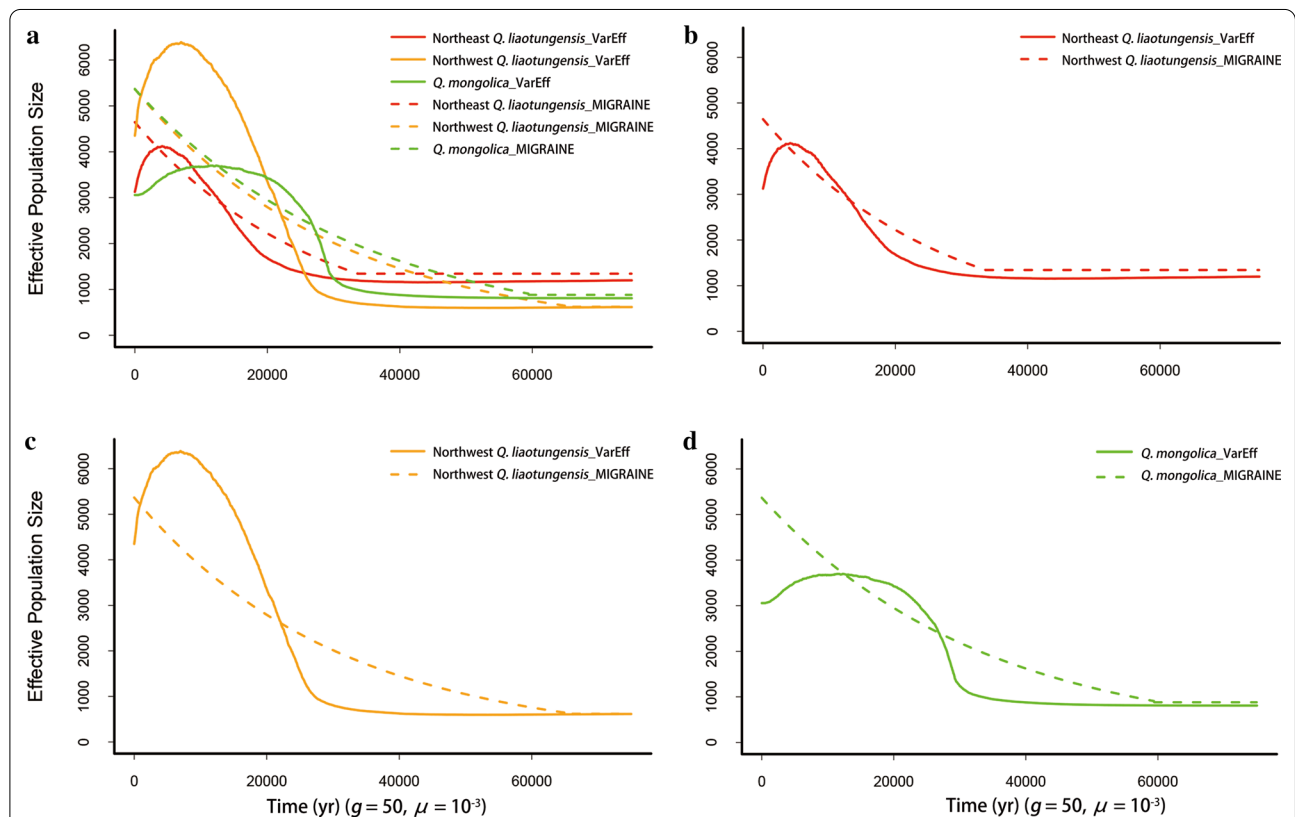
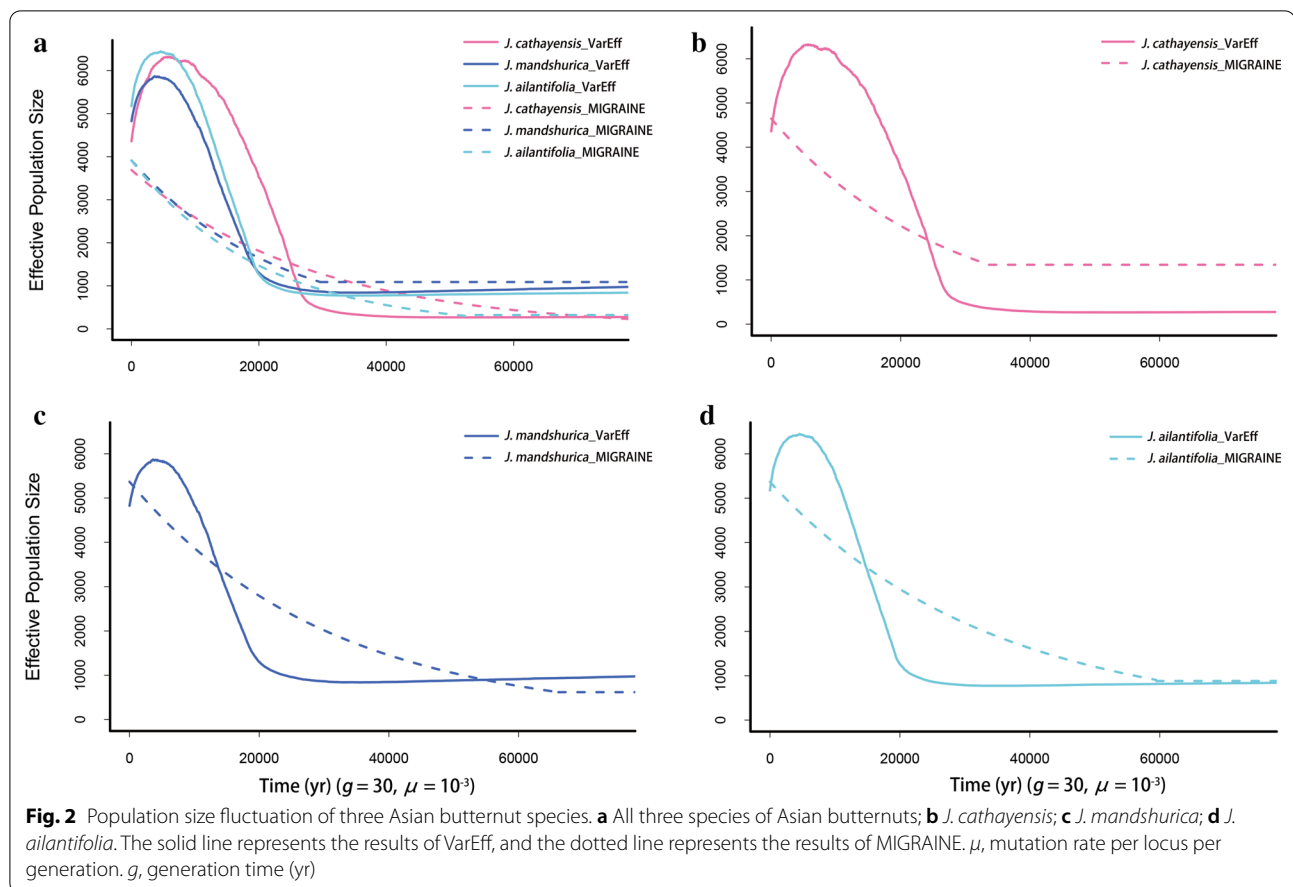


Fig. 1 Population size fluctuation of two oak species. **a** All two species of oaks; **b** Northeast *Q. liaotungensis*; **c** Northwest *Q. liaotungensis*; **d** *Q. mongolica*. The solid line represents the results of VarEff, and the dotted line represents the results of MIGRAINE. μ , mutation rate per locus per generation. g , generation time (yr)



VarEff. So, in the following, we focus our discussion on the VarEff results.

With a fixed mutation rate of 1×10^{-3} per locus per generation [16, 17] and a generation time of 50 years for oaks [45, 46] and 30 years for Asian butternuts [44], all five species began to expand at the time interval of 25–10 kya. Within oaks, the northwestern group of *Q. liaotungensis* had a high increase in population size ($N_e \approx 4,300$), while the northeastern group of *Q. liaotungensis* and *Q. mongolica* showed a moderate increase ($N_e \approx 3,100$). *Q. mongolica* remained stable size after the expansion at 25 kya (Fig. 1d). Within butternuts, *J. cathayensis* began to expand at 24 kya, while *J. mandshurica* and *J. ailantifolia* expanded at 18 kya (Fig. 2b). The effective population size of *J. cathayensis* ($N_e \approx 300$) was much smaller than *J. mandshurica* and *J. ailantifolia* before expansion.

VarEff suggested all five species had experienced a population decline after their expansions. However, as pointed out by Nikolic et al. [33], it should be very cautious to make such a conclusion, because VarEff cannot distinguish well between ongoing expansion and finished expansion without N_e changes afterward. Both scenarios would be detected as a population expansion followed by a recent population decline.

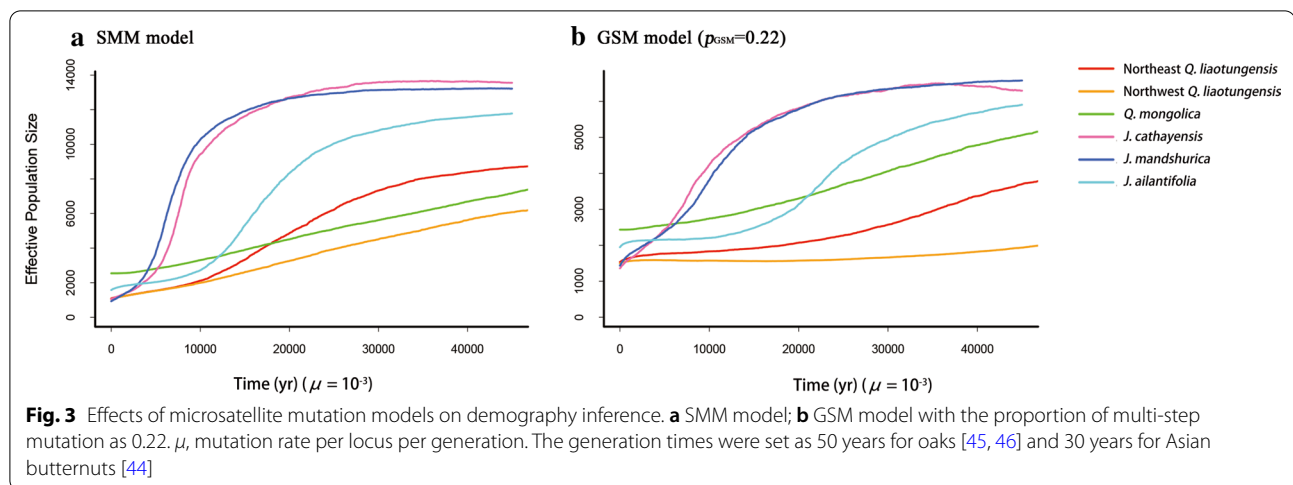
Effects of microsatellite mutation models on demographic inference

We used MIGRAINE to estimate the multistep mutation proportions (p_{GSM}) in the GSM for the five species. The point estimates for p_{GSM} were quite high, at the range of 0.53–0.63 for oaks and 0.62–0.64 for butternuts (Table 1). Then, we implemented VarEff to assess the effects of p_{GSM} on the demographic inference of these trees. When p_{GSM} was set as 0.22, all the five species were detected to experience population declines (Fig. 3b). Moreover, if we set mutation model as SMM instead of GSM, we also got a result of population decline (Fig. 3a). These findings indicate that we will get a wrong signal of population dynamics if we misuse the mutation model or incorrectly set the p_{GSM} parameter.

Discussion

Historical population dynamics is driven by climate fluctuations

The repeated Quaternary glacial cycles had a profound impact on the geographic history of temperate plants, especially during LGM (20–18 kya) [2–4]. During glacial periods, temperate trees were either forced to retreat southward [47] or restricted to a few northern refugia



[48–52] with dramatic population declines. During interglacial periods, they recolonized their modern distribution areas with population expansion. After analyzing the microsatellite data of the five tree species with two complementary programs, we found that they all experienced a strong and recent population expansion since the LGM (Figs. 1, 2, Table 1). This finding accords well with what ecological niche modeling predicted for Asian butternuts [41] and oaks [53], that is, the present distribution of butternuts and oaks are much wider and larger than that in LGM (see Fig. 6 of [41] and Fig. 4 of [53]). Since the expansion times of the five species were highly coincident with each other, we believe that their population dynamical changes are most likely attributed to climate fluctuations since the LGM.

Among the five species, *J. cathayensis* has experienced the most serious bottleneck, leading to a smallest effective population size before expansion. This indicated that *J. cathayensis* may be more susceptible to environmental changes than the other species. According to VarEff, *J. mandshurica* and *J. ailantifolia* had similar population sizes before and after their expansions, but different from that of *J. cathayensis*, which is consistent with the results of PSMC [44] that *J. cathayensis* decreased faster to an extremely small size than *J. mandshurica* and *J. ailantifolia* before LGM. The plot of *J. cathayensis* is different from those of *J. mandshurica* and *J. ailantifolia* before and after LGM, and *J. mandshurica* and *J. ailantifolia* were highly coincident with each other for most of time (Fig. 2). Among the three species of oak, population dynamics of the two groups of *Q. liaotungensis* were very similar, but different from *Q. mongolica* which maintained a constant size after the expansion. This may be because *Q. mongolica* is more acclimated to cold but not to the dry warming climate [54].

Effects of the microsatellite mutation model on inferring population dynamics

The mutation model of microsatellite is important for the estimation of population parameters such as the number of migrants, divergence time and N_e [17]. Although an increasing amount of data has shown that most microsatellite loci conform to a GSM rather than an SMM [55, 56], many programs for analyzing microsatellite data still assumed SMM as the only model. For example, as one of the most commonly used software in the last two decades, MSVAR includes only SMM model [22]. However, the misuse of SMM might generate severe bias in the inference of demographic history [30]. This is because multistep mutations in the GSM model can produce a pattern of allele length distribution very similar to that of population decline under the SMM model [31, 57]. Moreover, when inferring the ancient evolution events, the error rate of population parameter estimation with microsatellites is much higher than that with SNPs if an SMM or an infinite allele model (IAM) is assumed [14]. Therefore, when we are not sure about the microsatellite mutation process in a species, we should be very cautious to choose SMM as the mutation model [15]. In our study, the proportions of multistep mutations in the GSM calculated by MIGRAINE ranged from 0.50 to 0.65 in the oaks and 0.60 to 0.65 in the Asian butternuts, indicating GSM model was more suitable for microsatellites, at least for cool-temperate tree species. Bai et al. [41] estimated a similar value for Asian butternuts by using DIYABC program. We also found that different settings for the proportions of multistep mutations may lead to completely opposite trends of population size change, suggesting that correctly setting this parameter of the GSM (Fig. 3) is very important.

There are thousands of studies that have employed microsatellite markers in the last two decades, and hence it is worth continuing to mine the information from these existing data. As we noted earlier, microsatellites are well suited to detect recent population changes due to their high mutation rate and neutrality [14, 15]. Therefore, they can provide a good complement to the programs analyzing whole genome data which often cannot estimate recent population dynamics [9]. The development of bioinformatic methods in recent years will allow us to understand more about the distribution and mutation mechanisms of microsatellites [17]. For all of these reasons, microsatellites would remain an effective and cost-efficient marker to study population demography in the genomic age [58].

Comparison of MIGRAINE vs. VarEff

MIGRAINE is a software that relies on the importance sampling of gene genealogies and the coalescent theory under the maximum likelihood framework to estimate current and past N_e , as well as the time of changes in population size. It is more flexible and robust than other programs because it can accommodate several demographic models, e.g., isolation by distance model (*IBD*), the single population dynamic variation model (*OnePopVarSize*), the founder model (*OnePopFounderFlush*) and several microsatellite mutation models, e.g., SMMs, GSMs, and infinite sites mutation models (ISMs). In addition, MIGRAINE can estimate the multistep mutation proportion (p_{GSM}) in the GSM, which is a key parameter for other programs (e.g., VarEff). However, MIGRAINE is time consuming, and when running models of *OnePopVarSize* or *OnePopFounderFlush*, more runs (usually > 2000) per sampling point are needed. Furthermore, MIGRAINE assumes that a single isolated population has undergone a single past size change, which definitely oversimplifies actual complex population dynamics.

Compared to MIGRAINE, VarEff relies on an approximation likelihood algorithm to detect population demography. So, it need not presuppose any scenario about demographic history, such as monotonous growth or decline. VarEff is less dependent on the priors and can evaluate the effects of different mutation models on the results [33]. VarEff has included several microsatellite mutation models, such as SMMs, two-phase models (TPMs) and GSMs, but needs to set the multistep mutation proportion parameter in the GSM. For VarEff, the main disadvantages are that it cannot infer ancient population changes (e.g., $G \times \mu > 20$) and would give a false population decline when a SMM model was wrongly chosen or when substantial gene flow exists.

In view of their advantages and disadvantages, we advocate that researchers should apply MIGRAINE and VarEff in combination when inferring population

demography with microsatellite data, which can effectively reduce errors caused by inappropriate model selection and prior parameter settings.

Conclusions

Even though whole-genome sequencing is prevalent today, the low mutation rate of SNPs and the limited sampling size are two main hurdles preventing us from inferring recent population dynamics accurately. In this context, microsatellites may be a good choice for reconstructing population dynamics due to their high mutation rates and selective neutrality. Moreover, researchers have accumulated much more microsatellite data in last two decades, which can be reused to help infer most recent demographic events. Indeed, using microsatellites data we have successfully revealed expansions among five cool-temperate tree species in East Asia after the Last Glacial Maximum. This confirms previous hypothesis that the distribution range of temperate trees have expanded with the climate warming during interglacial period [41, 53]. Additionally, we have shown that when inferring demographic history with microsatellite data, both the microsatellite mutation model and the parameter settings can have a significant influence on the result. As a rule of thumb, it is better to assume a GSM rather than a SMM and to apply multiple programs to carry out microsatellite analysis.

Methods

Data source

All data were from previous studies; the data on oaks came from Zeng et al. [39] and the data on Asian butternuts from Bai et al. [41]. In both studies, 19 microsatellite markers were used respectively to generate polymorphism data.

Based on previous studies, the investigated areas covering the whole range of *Q. mongolica* and most distribution areas of *Q. liaotungensis* in China, included 502 individuals from 17 *Q. mongolica* populations, 139 individuals from five populations of northeast *Q. liaotungensis* and 240 individuals from eight populations of northwest *Q. liaotungensis*. The sample distribution range of Asian butternuts also covered the entire distribution area of the three species, including 596 individuals from 25 *J. cathayensis* populations, 399 individuals from 14 *J. mandshurica* populations and 107 individuals from five *J. ailantifolia* populations (Fig. 4).

Data analysis

MIGRAINE analysis of historical population dynamics

We explored the demographic history of the above five species using MIGRAINE 0.5.4 (<http://kimura>.

univ-montp2.fr/~rousset/Migraine.htm), which is based on importance sampling of gene genealogies under a maximum likelihood framework. It is extended for GSM which notably allows departure from the strict SMM with the parameter p_{GSM} for the geometric distribution of mutation sizes. The present and ancestral scaled population sizes in mutation rate ($\theta = 4N\mu$, $\theta_{anc} = 4N_{anc}\mu$), and the occurrence time of the past change scaled by mutation rate ($D = G \times \mu$) were calculated. Here, N_{anc} is the ancestral effective population size and N is the current effective population size; μ is the mutation rate per locus per generation; G is the time measured by generations. The detection of significant past change in the population is based on the population size ratio ($\theta_{ratio} = \theta/\theta_{anc}$). If $\theta_{ratio} < 1$, the population has experienced contraction; if $\theta_{ratio} > 1$, the population has experienced expansion. Also, we estimated the multistep mutation proportion of the GSM (p_{GSM}) using MIGRAINE.

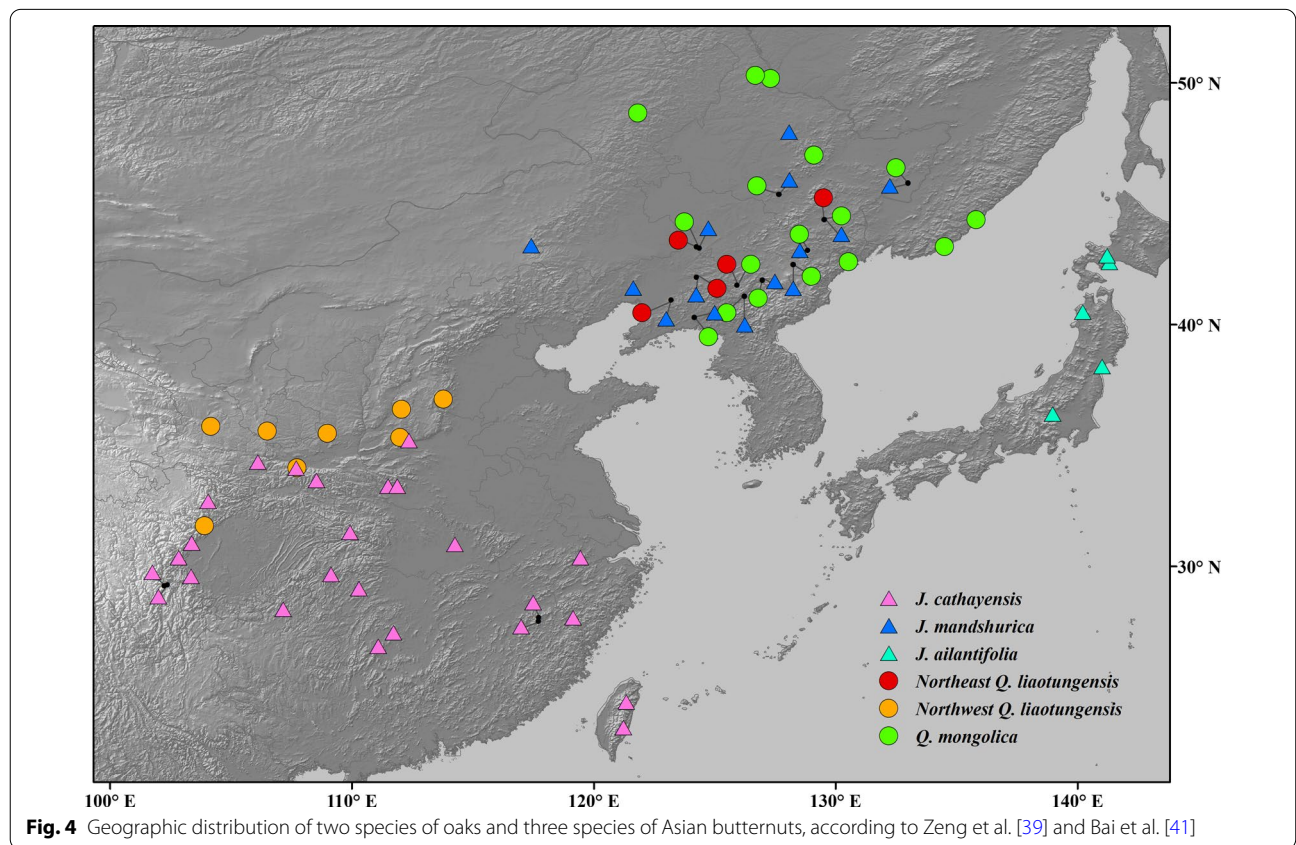
The *OnePopVarSize* model was used, which considers a single isolated population with a unique past size change. To ensure enough points with high likelihoods for the smoothing procedure, we considered relatively high values of parameter *NRunsPerPoint*. For the data of northeast *Q. liaotungensis*, northwest *Q. liaotungensis*, *Q. mongolica* and *J. mandshurica*, we considered 2,000

trees, with 200 points in each iteration and a total of 16 iterations. For the data of *J. cathayensis* and *J. ailantifolia*, MIGRAINE was run using 20,000 trees, with 200 points in each iteration and a total of 16 iterations.

VarEff analysis of historical population dynamics

Since MIGRAINE infers population dynamics under a model of a single panmictic population with one exponential change in population size, the highly simplified demographic model may be insufficient to characterize the actual situation. So, we also used VarEff [33], a program estimating the past changes in N_e by using approximate likelihoods under a MCMC approach. It has been found to be especially useful for providing evidence of transient changes in population size in the past. The VarEff method was implemented in the R package VarEff (<https://qgsp.jouy.inra.fr>). By using the functions in the VarEff package, we extracted several global statistics of N_e (arithmetic and harmonic means, mode, median, and quantiles). Since the median of the posterior distribution was found to be the most robust estimator [33], we visualized it at different times in the past.

Using the GSM as the microsatellite mutation model, we first set the multistep mutation proportion, C (which is the same as parameter p_{GSM} in MIGRAINE)



as that estimated by MIGRAINE. In order to compare with the results of MIGRAINE, we set the parameter *TMAX* (Length of the period for which the distributions of N_e in the past) in VarEff with reference to the estimates of parameter *D* in MIGRAINE. For the parameter settings related to MCMC, *NumberBatch*, *LengthBatch*, and *SpaceBatch* were set to be 10,000, 1, and 100, respectively, and *burnin* was set to 10,000.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-021-01805-w>.

Additional file 1: Table S1. Detailed population parameters estimated by VarEff. θ ($= 4N\mu$), the scaled current effective population size in mutation rate; θ_{anc} ($= 4N_{anc}\mu$), the scaled ancestor effective population size in mutation rate; θ/θ_{anc} (*H*), the ratio of harmonic means of the effective population size; θ/θ_{anc} (*M*), the ratio of medians of the effective population size. *N*, the current effective population size (individual number); N_{anc} , the ancestral effective population size (individual number); μ , the mutation rate per locus per generation.

Acknowledgements

We thank two anonymous reviewers for feedback and comments that improved this manuscript. We thank Chen J, Zhang WP and Ding YM for their helpful suggestions and valuable comments on the manuscript. We thank Wu Y and Liang Y for their help in applying R software.

Authors' contributions

DYZ and WNB conceived the ideas, designed methodology, WNB and YFZ provided data, YC analyzed data; YC drafted the article, DYZ, WNB and YFZ revised it critically for important intellectual content. All authors contributed critically to the drafts. All authors read and approved the final manuscript.

Funding

This work was financially supported by the National Key R&D Program of China (2017YFA0605100), the National Natural Science Foundation of China (31421063 and 41671040), and the "111" Program of Introducing Talents of Discipline to Universities (B13008). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

The datasets analyzed during the current study (SSR genotypes and geographical locations of collections) are available in the Dryad Digital Repository, <https://doi.org/10.5061/dryad.83bk3j9mz>.

Declarations

Ethics approval and consent to participate

All the plant materials were sampled from natural populations in North and Northeast China and no specific permission was needed to collect such samples. This study was conducted in accordance with local legislation and the Convention on the Trade in Endangered Species of Wild Fauna and Flora.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹State Key Laboratory of Earth Surface Process and Resource Ecology and Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing 100875,

China. ²State Key Laboratory of Tree Genetics and Breeding, Chinese Academy of Forestry, Beijing 100091, China.

Received: 6 November 2020 Accepted: 23 April 2021

Published online: 18 May 2021

References

- Nie GZ, Liu JQ, Guo ZT. The major stratigraphic boundaries and climatic events in Weinan loess section since 0.15Ma B. P.: based on chronological evidences. *Quat Sci.* 1996;16(3):221–231. <http://www.dsjyj.com.cn/CN/Y1996/V16/I3/221> (in Chinese with English abstract).
- Hewitt G. The genetic legacy of the quaternary ice ages. *Nature.* 2000;405(6789):907–13. <https://doi.org/10.1038/35016000>.
- Hu FS, Hampe A, Petit RJ. Paleoeecology meets genetics: deciphering past vegetational dynamics. *Front Ecol Environ.* 2009;7(7):371–9. <https://doi.org/10.1890/070160>.
- Hewitt GM. Genetic consequences of climatic oscillations in the quaternary. *Philos Trans R Soc B.* 2004;359(1442):183–95; discussion 195.
- Melo WA, Freitas CG, Bacon CD, Collevatti RG. The road to evolutionary success: insights from the demographic history of an Amazonian palm. *Heredity (Edinb).* 2018;121(2):183–95. <https://doi.org/10.1038/s41437-018-0074-1>.
- Evans LM, Allan GJ, DiFazio SP, Slavov GT, Wilder JA, Floate KD, Rood SB, Whitham TG. Geographical barriers and climate influence demographic history in narrowleaf cottonwoods. *Heredity (Edinb).* 2015;114(4):387–96. <https://doi.org/10.1038/hdy.2014.115>.
- Zhou L, Bawa R, Holliday JA. Exome resequencing reveals signatures of demographic and adaptive processes across the genome and range of black cottonwood (*Populus trichocarpa*). *Mol Ecol.* 2014;23(10):2486–99. <https://doi.org/10.1111/mec.12752>.
- Lin RC, Yeung CK, Li SH. Drastic post-LGM expansion and lack of historical genetic structure of a subtropical fig-pollinating wasp (*Ceratosolen* sp. 1) of *Ficus septica* in Taiwan. *Mol Ecol.* 2008;17(23):5008–22.
- Li H, Durbin R. Inference of human population history from individual whole-genome sequences. *Nature.* 2011;475(7357):493–6. <https://doi.org/10.1038/nature10231>.
- Schiffels S, Durbin R. Inferring human population size and separation history from multiple genome sequences. *Nat Genet.* 2014;46(8):919–25.
- Liu X, Fu YX. Exploring population size changes using SNP frequency spectra. *Nat Genet.* 2015;47(5):555–9.
- Salmona J, Heller R, Lascoux M, Shafer A. Inferring demographic history using genomic data. In: Rajara OP, editor. *Population Genomics*. Springer International Publishing: Berlin; 2017. p. 511–37.
- Patton AH, Margres MJ, Stahlke AR, Hendricks S, Lewallen K, Hamede RK, Ruiz-Aravena M, Ryder O, McCallum HI, Jones ME, et al. Contemporary demographic reconstruction methods are robust to genome assembly quality: a case study in Tasmanian Devils. *Mol Biol Evol.* 2019;36(12):2906–21.
- Freeland JR, Anderson S. *Molecular ecology*. 3rd ed. New Jersey: Wiley; 2018.
- Bai WN, Zhang DY. Current status and future directions in plant phylogeography. *Chin Bull Life Sci.* 2014;26(2):125–37. <https://doi.org/10.13376/j.cbils/2014020> (in Chinese with English abstract).
- Selkoe KA, Toonen RJ. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol Lett.* 2006;9(5):615–29. <https://doi.org/10.1111/j.1461-0248.2006.00889.x>.
- Hoshino AA, Bravo JP, Nobile PM, Morelli KA. Microsatellites as tools for genetic diversity analysis. In: Caliskan M, editor. *Genetic diversity in microorganisms*. Rijeka: InTech; 2012. p. 149–70.
- Pokhriyal B, Thorat K, Limaye D, Joshi Y, Kadam V, RD. Microsatellite markers – a novel tool in molecular genetics. *Int J Res Pharm Chem.* 2012;2(2):397–412.
- Zane L, Bargelloni L, Patarnello T. Strategies for microsatellite isolation: a review. *Mol Ecol.* 2002;11(1):1–16. <https://doi.org/10.1046/j.0962-1083.2001.01418.x>.
- Cornuet JM, Luikart G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics.* 1996;144(4):2001–14.

21. Garza JC, Williamson EG. Detection of reduction in population size using data from microsatellite loci. *Mol Ecol*. 2001;10(2):305–18. <https://doi.org/10.1046/j.1365-294x.2001.01190.x>.
22. Girod C, Vitalis R, Leblois R, Freville H. Inferring population decline and expansion from microsatellite data: a simulation-based evaluation of the Msvar method. *Genetics*. 2011;188(1):165–79. <https://doi.org/10.1534/genetics.110.121764>.
23. Beaumont MA. Detecting population expansion and decline using microsatellites. *Genetics*. 1999;153(4):2013–29.
24. Beerli P. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*. 2006;22(3):341–5. <https://doi.org/10.1093/bioinformatics/bti803>.
25. Felsenstein J. Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. *Genet Res*. 1992;59:139–47.
26. Griffiths RC, Tavare S. Sampling theory for neutral alleles in a varying environment. *Philos Trans R Soc B*. 1994;344(1310):403–10.
27. Emerson BC, Paradis E, Thébaud C. Revealing the demographic histories of species using DNA sequences. *Trends Ecol Evol*. 2001;16(12):707–16.
28. Ohta T, Kimura M. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population*. *Genet Res*. 2007;89(5–6):367–70. <https://doi.org/10.1017/S0016672308009531>.
29. Bhargava A, Fuentes FF. Mutational dynamics of microsatellites. *Mol Biotechnol*. 2010;44(3):250–66. <https://doi.org/10.1007/s12033-009-9230-4>.
30. Faurby S, Pertoldi C. The consequences of the unlikely but critical assumption of stepwise mutation in the population genetic software, MSVAR. *Evol Ecol Res*. 2012;14(7):859–879. <https://www.researchgate.net/publication/257924528>.
31. Leblois R, Pudlo P, Neron J, Bertaux F, Reddy Beeravolu C, Vitalis R, Rousset F. Maximum-likelihood inference of population size contractions from microsatellite data. *Mol Biol Evol*. 2014;31(10):2805–23. <https://doi.org/10.1093/molbev/msu212>.
32. Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol Biol Evol*. 1999;16(12):1791–8. <https://doi.org/10.1093/oxfordjournals.molbev.a026091>.
33. Nikolic N, Chevalet C. Detecting past changes of effective population size. *Evol Appl*. 2014;7(6):663–81. <https://doi.org/10.1111/eva.12170>.
34. Savolainen O, Pyhäjärvi T, Knürr T. Gene flow and local adaptation in trees. *Annu Rev Ecol Syst*. 2007;38(1):595–619. <https://doi.org/10.1146/annurev.ecolsys.38.091206.095646>.
35. Bai WN, Liao WJ, Zhang DY. Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. *New Phytol*. 2010;188(3):892–901. <https://doi.org/10.1111/j.1469-8137.2010.03407.x>.
36. Zeng YF, Liao WJ, Petit RJ, Zhang DY. Exploring species limits in two closely related Chinese oaks. *PLoS ONE*. 2010;5(11):1–15.
37. Bai WN, Wang WT, Zhang DY. Contrasts between the phylogeographic patterns of chloroplast and nuclear DNA highlight a role for pollen-mediated gene flow in preventing population divergence in an East Asian temperate tree. *Mol Phylogenet Evol*. 2014;81:37–48. <https://doi.org/10.1016/j.ympev.2014.08.024>.
38. Wang LM, Ren XW, Liu YQ. Geographic distribution of deciduous oaks in China. *J Beijing For Coll*. 1985;2:57–69.
39. Zeng YF, Liao WJ, Petit RJ, Zhang DY. Geographic variation in the structure of oak hybrid zones provides insights into the dynamics of speciation. *Mol Ecol*. 2011;20(23):4995–5011. <https://doi.org/10.1111/j.1365-294X.2011.05354.x>.
40. Aradhya MK, Potter D, Simon CJ. Origin, evolution, and biogeography of *Juglans*: a phylogenetic perspective. In: 5th International Walnut Symposium: Nov 09–13, 2004; Sorrento, Italy. 2006:85–94.
41. Bai WN, Wang WT, Zhang DY. Phylogeographic breaks within Asian butternuts indicate the existence of a phytogeographic divide in East Asia. *New Phytol*. 2016;209(4):1757–72. <https://doi.org/10.1111/nph.13711>.
42. Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin JM, Estoup A. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*. 2014;30(8):1187–1189. <https://doi.org/10.1093/bioinformatics/btt763>.
43. Navascues M, Leblois R, Burgarella C. Demographic inference through approximate-Bayesian-computation skyline plots. *PeerJ*. 2017;5:13. <https://doi.org/10.7717/peerj.3530>.
44. Bai WN, Yan PC, Zhang BW, Woeste KE, Lin K, Zhang DY. Demographically idiosyncratic responses to climate change and rapid Pleistocene diversification of the walnut genus *Juglans* (Juglandaceae) revealed by whole-genome sequences. *New Phytol*. 2018;217(4):1726–36. <https://doi.org/10.1111/nph.14917>.
45. Gregorius HR, Degen B, König A. Problems in the analysis of genetic differentiation among populations – a case study in *Quercus robur*. *Silvae Genet*. 2007;56(1–6):190–9.
46. Leroy T, Rougemont Q, Dupouey JL, Bodenès C, Lalanne C, Belsier C, Labadie K, Le Provost G, Aury JM, Kremer A, et al. Massive postglacial gene flow between European white oaks uncovered genes underlying species barriers. *New Phytol*. 2020;226(4):1183–97.
47. Bennett KD, Tzedakis PC, Willis KJ. Quaternary refugia of North European trees. *J Biogeogr*. 1991;18(1):103–15. <https://doi.org/10.2307/2845248>.
48. Qian H, Ricklefs RE. Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature*. 2000;407(6801):180–2. <https://doi.org/10.1038/35025052>.
49. Harrison SP, Yu G, Takahara H, Prentice IC. Diversity of temperate plants in East Asia. *Nature*. 2001;413(6852):129–30. <https://doi.org/10.1038/35093166>.
50. Qiu YX, Fu CX, Comes HP. Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of quaternary climate and environmental change in the world's most diverse temperate flora. *Mol Phylogenet Evol*. 2011;59(1):225–44. <https://doi.org/10.1016/j.ympev.2011.01.012>.
51. Ye J, Yuan Y, Cai L, Wang X. Research progress of phylogeographic studies of plant species in temperate coniferous and broadleaf mixed forests in Northeastern China. *Biodivers Sci*. 2017;25(12):1339–49. <https://doi.org/10.17520/biods.2017265> (in Chinese with English abstract).
52. Stewart JR, Lister AM. Cryptic northern refugia and the origins of the modern biota. *Trends Ecol Evol*. 2001;16(11):608–13.
53. Zeng YF, Wang WT, Liao WJ, Wang HF, Zhang DY. Multiple glacial refugia for cool-temperate deciduous trees in northern East Asia: the Mongolian oak as a case study. *Mol Ecol*. 2015;24(22):5676–91. <https://doi.org/10.1111/mec.13408>.
54. Li T, He X-Y, Chen Z-J. Tree-ring growth responses of Mongolian oak (*Quercus mongolica*) to climate change in southern Northeast: a case study in Qianshan Mountains. *Chin J Appl Ecol*. 2014;25(7):1841–8. <https://doi.org/10.13287/j.1001-9332.2014.0124> (in Chinese with English abstract).
55. Ellegren H. Microsatellites: simple sequences with complex evolution. *Nat Rev Genet*. 2004;5(6):435–45.
56. Seyfert AL, Cristescu ME, Frisse L, Schaack S, Thomas WK, Lynch M. The rate and spectrum of microsatellite mutation in *Caenorhabditis elegans* and *Daphnia pulex*. *Genetics*. 2008;178(4):2113–21.
57. Jentzsch IMV, Bagshaw ATM, Buschiazio E, Merkel A, Gemmill NJ. Evolution of Microsatellite DNA. In: eLS. Wiley J, Sons, editors. Chichester: Wiley; 2013.
58. Hodel RG, Segovia-Salcedo MC, Landis JB, Crowl AA, Sun M, Liu X, Gitzendanner MA, Douglas NA, Germain-Aubrey CC, Chen S, et al. The report of my death was an exaggeration: a review for researchers using microsatellites in the 21st century. *Appl Plant Sci*. 2016;4(6):1–13. <https://doi.org/10.3732/apps.1600025>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.