

# European Invasion of North American *Pinus strobus* at Large and Fine Scales: High Genetic Diversity and Fine-Scale Genetic Clustering over Time in the Adventive Range

Bohumil Mandák<sup>1,2\*</sup>, Věroslava Hadincová<sup>1</sup>, Václav Mahelka<sup>1</sup>, Radka Wildová<sup>1,3‡</sup>

**1** Institute of Botany, Academy of Sciences of the Czech Republic, Průhonice, Czech Republic, **2** Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Prague, Czech Republic, **3** Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, United States of America

## Abstract

**Background:** North American *Pinus strobus* is a highly invasive tree species in Central Europe. Using ten polymorphic microsatellite loci we compared various aspects of the large-scale genetic diversity of individuals from 30 sites in the native distribution range with those from 30 sites in the European adventive distribution range. To investigate the ascertained pattern of genetic diversity of this intercontinental comparison further, we surveyed fine-scale genetic diversity patterns and changes over time within four highly invasive populations in the adventive range.

**Results:** Our data show that at the large scale the genetic diversity found within the relatively small adventive range in Central Europe, surprisingly, equals the diversity found within the sampled area in the native range, which is about thirty times larger. Bayesian assignment grouped individuals into two genetic clusters separating North American native populations from the European, non-native populations, without any strong genetic structure shown over either range. In the case of the fine scale, our comparison of genetic diversity parameters among the localities and age classes yielded no evidence of genetic diversity increase over time. We found that SGS differed across age classes within the populations under study. Old trees in general completely lacked any SGS, which increased over time and reached its maximum in the sapling stage.

**Conclusions:** Based on (1) the absence of difference in genetic diversity between the native and adventive ranges, together with the lack of structure in the native range, and (2) the lack of any evidence of any temporal increase in genetic diversity at four highly invasive populations in the adventive range, we conclude that population amalgamation probably first happened in the native range, prior to introduction. In such case, there would have been no need for multiple introductions from previously isolated populations, but only several introductions from genetically diverse populations.

**Citation:** Mandák B, Hadincová V, Mahelka V, Wildová R (2013) European Invasion of North American *Pinus strobus* at Large and Fine Scales: High Genetic Diversity and Fine-Scale Genetic Clustering over Time in the Adventive Range. PLoS ONE 8(7): e68514. doi:10.1371/journal.pone.0068514

**Editor:** Sebastien Lavergne, CNRS/Université Joseph-Fourier, France

**Received:** November 2, 2012; **Accepted:** May 29, 2013; **Published:** July 10, 2013

**Copyright:** © 2013 Mandák et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was supported by grant no. IAA600050711 from the Grant Agency of the Academy of Sciences of the Czech Republic, partly supported by grant no. P504/11/0402 from the Grant Agency of the Czech Republic and realized in the frame of the long-term research development project no. RVO 67985939. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: mandak@ibot.cas.cz ‡ Current address: Cary Institute of Ecosystem Studies, Millbrook, NY, USA

## Introduction

Many studies of plant invasions have focused on comparative aspects of ecology, with emphases on understanding either the properties of species that determine their invasive potential [1–3] or the properties of communities that determine their resistance to invasion [4,5]. The outcome of an invasion may be influenced by abiotic factors such as resource availability [6] or type and frequency of disturbance [7], and by biotic characteristics such as propagule supply [8,9] or the abundance of natural enemies, competitors or mutualists in the host community [10,11]. Although some general theories of invasibility have been put forward [6,12,13], it is not easy to come up with general explanations as to why some invasions succeed while other fail [3,10,14,15]. This is because of the numerous factors influencing

populations, species, plant communities, and whole ecosystems [14,15].

Even if introduced species encounter physical environmental (e.g. climatic or edaphic) conditions in their adventive ranges that are similar to those in their native ranges, they face new biotic environments [14]. Therefore, the success of invasive species may largely depend on their ability to evolve in response to new environments, although adaptation may not always be a prerequisite for successful establishment of introduced populations [16–18]. The potential for local adaptation would seem to be limited by loss of genetic diversity during the introduction process and subsequent range expansion in newly colonized areas [19]. In particular, during colonization the founder effect may reduce genetic diversity in newly established populations relative to the source population [20].

Several mechanisms by which invasive populations can retain sufficient genetic diversity to enable adaptation have been suggested. First, rapid population expansion can occur immediately after an introduction, allowing retention of genetic diversity [21] and facilitating later adaptation to a local environment. Second, polyploidization and hybridization phenomena may rapidly produce novel diversity in the introduced area [22]. Finally, when a species is introduced to a new range from several genetically different populations, further recombination within the adventive range may dramatically increase genetic diversity outside the native range [23–27]. This population admixture among genetically distinct lineages has been predicted to contribute to invasion success by directly increasing fitness through hybrid vigour or by enhancing evolutionary potential within populations [28].

The time taken for introduced populations to acquire sufficiently high frequencies of adaptive genotypes might explain the substantial time lags (of years or decades) between initial establishment and the manifestation of invasiveness. These time lags have been well documented [29–32] and could be due to these evolutionary processes as well as ecological dynamics such as the time needed to spread to favourable habitat or to reach population sizes capable of producing abundant offspring [33].

Although many studies have reported multiple introductions of invasive species, the impacts of admixture on the fitness of invasive species have rarely been studied. Those studies that have been done have all shown a strong influence of the resulting genetic diversity on colonization success [28,34–37]. Population-level genetic diversity appears important for long-term stability, especially if adaptive genetic variation is maintained along with neutral variation [38]. Following population establishment, genetic diversity may increase with population age due to admixture of populations derived from different introduction sources. These processes have been studied using theoretical models [39], but the amount of empirical information we have about the relative importance of spread rate and population age for the distribution of genetic diversity within and among invasive populations is still limited. It is nevertheless important to understand the distribution of genetic diversity along invasion routes, especially if we are to better understand invasion dynamics. However, we are usually unable to observe the invasion process as it happens because we cannot study the originally introduced individuals which had started the invasion, often many years earlier. Their progeny, which can slowly adapt to the new environment, also disappear over time in most cases. There is one prominent exception, however: trees. Invasive tree species represent a unique opportunity to study, in ideal cases, representatives of all generations that took part in an invasion within an area. The originally introduced trees and their progeny represent various stages of the lag phase. Thanks to the long life history of trees, a continuous record can be available at study sites, so the entire invasion history can be readily determined by using selectively neutral markers.

It has been shown that not only genetic diversity *per se* but also its distribution in space is an important aspect of population dynamics ([40] and references therein). On the one hand, it has been shown that spatial genetic structure (hereafter SGS) within populations is influenced by patterns of seed dispersal [41], mating system, [42] mating pattern [43], colonization events [44–46], competition [47,48], demographic structure [49] and microhabitat heterogeneity [50]. On the other hand, SGS may itself have an effect on population dynamics by affecting the level of inbreeding and microenvironmental adaptation [50], mating pattern [43], or patterns of viability selection [51,52]. Therefore, in newly established invading populations, characterizing both the patterns

of fine-scale genetic diversity and SGS and their changes over time provides an important tool especially for shedding light on the processes of invasion when genetic reorganization can be expected. Analyses of both genetic diversity and SGS changes over time in tree species are quite rare, however [40].

From this standpoint, *Pinus strobus* in the Czech Republic is one of the best examples of an invasive tree under study [53]. Large amounts of this species' seeds intended for cultivation have been imported from North America since the end of the 18<sup>th</sup> century [54]. Historical records provide evidence of about tens of kilograms of *P. strobus* seeds introduced to different parts of the Czech Republic at the turn of the 18<sup>th</sup> and 19<sup>th</sup> century [54]. At present, this long-lived species is highly invasive in several mainly sandstone areas of the country [55] (further denoted as invasive populations; *sensu* Richardson *et al.* [56]), but non-invasive in other parts (further denoted as naturalized populations [56]). It is now a component not only of planted mixed forests but also of other forests, as well as occurring in sparsely vegetated rocky sites. In Central Europe, many sandstone areas are protected for their unique environment, and large-scale regeneration of an alien tree species therefore poses a problem of serious conservation concern [53].

We used *P. strobus* populations in the Czech Republic (Central Europe) and North America to test theoretical predictions concerning genetic diversity of an invasive species in its native and adventive ranges, and evolution in its adventive range on large as well as fine scales. In studying the pattern of large-scale genetic diversity, we specifically asked (1) whether there are any differences in population genetic diversity and structure between the native and the adventive ranges and between invasive and naturalized populations within the adventive range; (2) whether there is any indication of bottleneck in introduced populations in the adventive range, and (3) whether populations of *P. strobus* are derived from independent introduction events and from which areas. In a subsequent study we investigated the pattern of fine-scale genetic diversity and SGS and their changes over time within four highly invasive populations in the species' adventive range. Such a fine-scale study might help clarify the species' population dynamics at different localities over the course of its invasion history. In particular we: (1) investigated the pattern of genetic diversity within and among different age classes and in different highly invasive populations; and (2) asked if there is any SGS that is maintained within and among different age classes in different invasive populations.

## Materials and Methods

### Ethics Statement

The study did not involve endangered or protected species. To perform our study in the Bohemian Switzerland National Park, a collaboration agreement was made between the Institute of Botany of the Academy of Sciences of the Czech Republic and the National Park authorities. Other forests on the Czech side are managed by the Forests of the Czech Republic, State Enterprise, i.e. forests managed under public ownership, for which no special permission is required to collect small samples of needles from an alien invasive species. We did not procure special permissions for collecting in any of the US localities, as we collected just several needles from a very common tree species. We could confirm that none of the collection sites were privately owned.

### Study Species

*Pinus strobus* L. (eastern white pine, also called Weymouth pine) is a conifer native to northeast North America [57], with its

current distribution there shown in Fig. 1. In its native range, it is common in moist cool forests, near streams and rivers, and on rocky or sandy nutrient-poor and well drained sites [57]. The majority of *P. strobus* stands were deforested by the late 19<sup>th</sup> century [58], and therefore, not many old-growth stands are left in its native area. However, because it is among the most rapidly growing northern conifers, it has been used extensively for reforestation projects [57].

*Pinus strobus* is highly invasive in some sandstone areas of the Czech Republic [53]. The first report of the introduction of a large amount of this species' seeds intended for cultivation in the Czech Republic is from 1784 [54]. In the sandstone area, it was first cultivated in 1798 (Elbe River Sandstones – district Přední and Zadní Doubice). *Pinus strobus* was introduced into mixed conifer forests to increase community diversity and inhibit pest infestation in species-poor forests on nutrient poor sandy soils and sandy loams [54]. The ability of the species to regenerate easily was noticed from the beginning of its cultivation [54,59], but its spread into surrounding forests was not spotted until the 1950s, when self-sown individuals were observed in the area of the Elbe River Sandstones [53]. A massive expansion took place in the same area in the 1990s, when the species entered the predominantly acidophilous pine (*Dicrano-Pinetum*) and oak-pine (*Vaccinio vitis-idaeae-Quercetum*) forests [53], which are typical plant communities of protected sandstone areas.

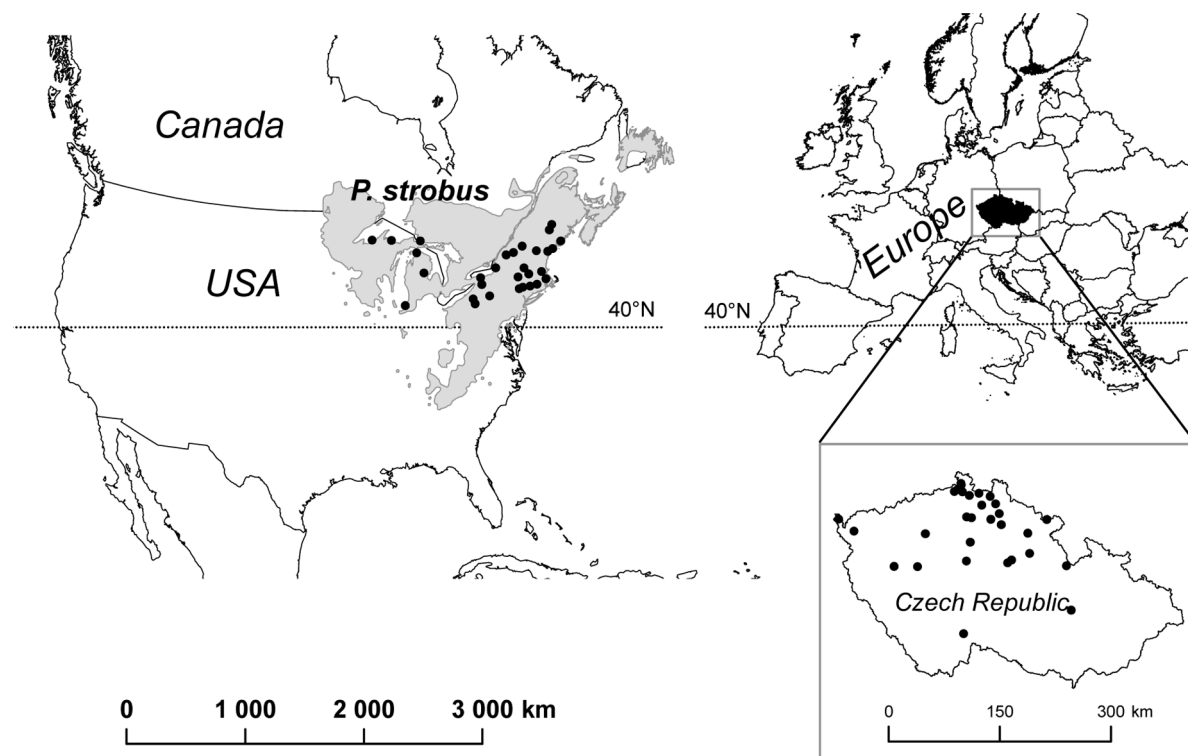
*Pinus strobus* is an outcrossing, wind-pollinated species whose winged seeds are dispersed by wind. In the adventive area, individual trees can be fertile at the age of twenty years, with the substantial production of seeds at the age of 50 years and older. Münzbergová *et al.* [60] found that younger, middle-aged trees (20–50 years old) contributed to the total seed set by only about 5% of all seeds and trees younger than 20 years have nearly no fertility. White pine produces many seedlings at all localities

observed in sandstone areas and on sandy and sandy-loam soils [53,61]. However, high mortality (50%) of seedlings younger than 5 years substantially reduces the number of individuals passing into the sapling stage. The mortality of saplings (5–20 years old), middle-aged (20–50 years old) and old ( $\geq 50$  years old) trees is low (5%, 0.75% and 0.9%, respectively; [60]).

*Pinus strobus* possesses high seed and pollen dispersal capacity. Wind-dispersed seeds released from the upper canopy can travel several hundreds meters from the source tree, with the maximum recorded distance of a self-sown tree from the potential seed source found to be 757 m. However, the observed median distance for 50% of self-sown trees was only 52 m and only 1% of self-sown trees reached 502 m [62]. Some pine pollen has been reported to travel tens of kilometers [63,64].

### Locality Selection and Sampling Procedures

**General approach – determination and definition of age classes.** In both large- and fine-scale studies, we aimed at obtaining samples from different phases of the invasion process. We thus sampled trees of several consecutive age classes which we define, and hereafter use, as follows: (1) “old trees” – trees 50 years old and older. While these trees are canopy trees (in the adventive range mostly planted); (2) “mature trees” – between 21 and 49 years old trees; (3) “saplings” – trees of age between 6 and 20 years; and (4) “seedlings” – trees not older than 5 years. Individual age classes were determined on the basis of a combination of mortality rate and fecundity estimated in a related study ([60], see Study species in Materials and Methods section). The age of old trees was either extracted from forest inventory data contained in forest management plans or estimated by counting whorls, which are formed annually [65]. Age determination based on whorl counting is precise up to the age of 50 years, where the Pearson correlation coefficient between the number of whorls (or branch



**Figure 1. Geographical location of *Pinus strobus* study populations.**  
doi:10.1371/journal.pone.0068514.g001

scars) and number of rings above the root collar was found to be 0.979 in the Czech Republic [62]. The age of trees in the other age classes was estimated by whorl-counting, which is more precise in younger age classes.

**Large-scale study.** We sampled 60 localities, 30 in the native and 30 in the adventive range. Based on a comparative morphological study of seedlings, Musil [59,66] stated that populations of *P. strobus* occurring in Europe were derived from a relatively small area near two important commercial ports (New York and Boston) along the eastern edge of its native range. Because the distribution of *P. strobus* in North America is quite broad [57], we sampled both in this area from which the species was most likely introduced and from other areas from which introduction was not as likely. A detailed list of individual locations is presented in Table 1. Localities were selected to be placed in close proximity to old-growth trees based on a survey of old-growth forests in the eastern USA done by Davis [67]. We sampled forests with the average age of the oldest trees about 83 years with the exception of one location where we sampled trees that were more than 300 years old (population A57). The sampling in the adventive range was done only within the Czech Republic (Fig. 1, Table 1), with an effort to sample all the areas in the country in which the species occurs. Samples were collected in July and August 2007 in the native range and from April 2007 to September 2008 in the adventive range.

We collected several needles from each of 20 plants at each locality wherever possible. Some localities, however, had such small numbers of trees that we could only sample fewer individuals there (Table 1). At each locality, we aimed at sampling trees of all age cohorts where it was possible. Because various authors [68,69] have detected a significant fine-scale genetic structure in native American populations at 15 m intervals, we sampled at this interval. Since little is known about the fine-scale genetic structure of *P. strobus* populations in Central Europe, we sampled plants at the same interval in the adventive range to ensure sampling consistency. Within a population, we would locate individuals to be sampled by creating a 60 by 45 m rectangular grid with the gridlines spaced at 15 m intervals, and then collecting from those trees located at the intersections of the gridlines. If no *P. strobus* tree was located at one of these points, we would use the plant nearest to the point but not further than 1 m away in cases of trees not older than 10 years, 2.5 m away in cases of trees not older than 50 years, and 5 m away in cases of trees older than 50 years. If the resulting area surrounding a point did not contain any usable tree, we skipped that point and extended our grid by 15 m to generate a replacement point. Where possible, we would continue sampling until we collected samples from twenty trees. In total we collected 1194 samples.

In the Czech populations, we later used the number of points in our grid not occupied by trees for discrimination between invasive and naturalized populations (sensu Richardson *et al.* [56]). A population was considered invasive if more than 90% of points contained a self-sown individual; otherwise it would be classified as naturalized (see Table 1). Populations classified as invasive clearly differed also visually from naturalized ones in that the former would include seedlings and saplings nearly completely covering available space, while the naturalized populations were predominantly composed of old trees with only a few self-sown individuals. For each population in the adventive range we also collected data on elevation above sea level, long term average annual precipitation and air temperature [70] and soil types (taken from National geoportal INSPIRE, <http://geoportal.gov.cz>).

**Fine-scale study.** We chose four localities in different parts of the Czech Republic, all of them being highly affected by the

invasion of *Pinus strobus*. The four sampling areas were (see Table 1): (1) Bynovec – locality E28, (2) Hrad – locality E1, (3) Ralsko – locality E22 and (4) Sopřech – locality E30 (for details on individual localities see Table 1). To obtain samples from different phases of the invasion process, we sampled trees belonging to the consecutive age classes as described above, i.e. old trees, mature trees, saplings, and seedlings.

In each of the four populations, we intended to collect 50 samples of each age class from regularly distributed plots, if possible. To ensure sampling consistency, we used the same sampling strategy at all localities. At each locality, we sampled all old trees because their number was always limited. Afterwards, in the core of each stand, we created a 40 by 100 m grid with gridlines spaced at 10 m intervals. We then took one sample of each age class within each square (except for old trees), i.e., 50 samples from each age class were taken. Saplings were originally divided into old (50 samples) and young saplings (50 samples) that were later merged due to our inability to precisely define these groups on the basis of mortality and fecundity parameters. In most cases, we were unable to collect all 250 samples at each locality, in particular due to the lack of old trees. In total, we collected 954 samples and recorded the spatial coordinates of each. The positions of all trees were recorded using a post-processing DGPS technique (GPS Trimble Pathfinder Pro XRS). All GPS positions were corrected to achieve sub-meter accuracy using data from two-reference base stations located in distance of up to 50 km (data provided by Czech office for surveying mapping and cadastre; <http://www.cuzk.cz>).

## Molecular Methods

**DNA extraction.** The needles were stored in CTAB. DNA from the *P. strobus* samples was isolated as described in Štorchová *et al.* [71], with only needles crushed in liquid nitrogen. The quality and yield of isolated DNA was checked on agarose gels and then precisely measured for DNA concentration using a biophotometer (Eppendorf, Germany). All samples were then diluted to a 15–30 ng/μl concentration, which is suitable for PCR with labeled microsatellite primers.

**Microsatellite analysis.** We analysed genetic variation at ten nuclear microsatellite loci in both the large-scale genetic diversity study (i.e., 1194 samples) and the fine-scale genetic diversity study (i.e., 954 samples). These loci were RPS1b, RPS2, RPS12, RPS25b, RPS34b, RPS39, RPS50, RPS84, RPS118b and RPS127 [72,73]. DNA amplification was carried out in three multiplex PCRs (multiplex 1: RPS1b, RPS2, RPS12 and RPS39; multiplex 2: RPS25b, RPS34b, RPS50 and RPS127; multiplex 3: RPS84 and RPS118b). DNA was amplified using the QIAGEN Multiplex PCR kit (QIAGEN, Germany) in a total reaction volume of 5 μl of PCR mix plus 5 μl of mineral oil to avoid PCR mix evaporation, containing 15–30 ng of DNA, 0.1–0.5 μM of each primer, and 2.5 μl of Master Mix (QIAGEN). To improve the quality of the PCR product, we added Q-solution (QIAGEN) and 4 mM MgCl<sub>2</sub> in Multiplex 3. PCR amplifications were conducted in a Mastercycler (Eppendorf) under the following conditions: an initial denaturation step of 15 min at 95°C followed by 40 cycles of 30 s at 94°C, 90 s at 60°C, 60 s at 72°C and a final extension of 10 min at 72°C.

PCR products were electrophoresed in an ABI PRISM 3130 sequencer (Applied Biosystems, USA). One microlitre of PCR product was mixed with 0.2 μl of GeneScan-500 LIZ (Applied Biosystems) and 12 μl of formamide (Applied Biosystems). Allele sizes were determined using GeneMapper version 4.0 software (Applied Biosystems). The raw data are available by request from the authors. An individual was declared null (nonamplifying at a

**Table 1.** Summary of genetic diversity within 60 populations of *Pinus strobus* based on ten microsatellite loci.

Pop.	Locality	Latitude	Longitude	N	A	H <sub>S</sub>	R <sub>S</sub>	H <sub>O</sub>	H <sub>E</sub>	f(F <sub>IS</sub> )
E1	Doubice, Hrad	N50°53'26"	E14°24'58"	20	5.2	0.597	4.467	0.479	0.579	0.198*
E2	Doubice, Dravčí Skály	N50°54'25"	E14°24'25"	20	4.9	0.543	4.005	0.441	0.525	0.188*
E3	Hradec Králové	N50°10'13"	E15°52'34"	20	5.1	0.547	4.180	0.444	0.529	0.188*
E4	Klokočí	N50°36'20"	E15°12'46"	20	4.5	0.550	3.845	0.433	0.533	0.214*
E5	Brandýs-Stará Boleslav	N50°13'02"	E14°44'05"	20	5.5	0.569	4.478	0.433	0.550	0.240*
E6	Řečany nad Labem	N50°01'38"	E15°28'59"	20	5.2	0.614	4.362	0.500	0.594	0.185*
E7	Bílchov	N50°14'38"	E13°52'08"	21	5.2	0.543	4.302	0.372	0.524	0.314*
E8	Říčany	N49°59'10"	E14°42'25"	20	5.0	0.520	4.118	0.418	0.503	0.195*
E9	Jemčina	N49°06'30"	E14°50'04"	20	4.9	0.533	4.113	0.394	0.512	0.261*
E10	Adršpach	N50°36'08"	E16°07'33"	20	4.1	0.465	3.473	0.363	0.449	0.219*
E11	Studené	N50°04'20"	E16°35'19"	20	4.5	0.524	3.841	0.377	0.505	0.280*
E12	Blatce	N50°30'53"	E14°36'05"	19	5.0	0.544	4.274	0.459	0.526	0.156*
E13	Pařezská Lhota	N50°28'28"	E15°16'43"	20	6.0	0.517	4.689	0.385	0.499	0.255*
E14	Příhrazy	N50°31'35"	E15°03'54"	20	5.0	0.546	4.175	0.437	0.528	0.198*
E15	Aš	N50°15'23"	E12°12'25"	20	5.4	0.585	4.443	0.452	0.565	0.227*
E16	Kynšperk nad Ohří	N50°08'32"	E12°31'42"	21	4.7	0.586	4.042	0.441	0.568	0.248*
E17	Cvikov	N50°46'31"	E14°36'02"	20	5.4	0.545	4.482	0.433	0.528	0.206*
E18	Liberec	N50°48'04"	E14°59'51"	21	5.7	0.563	4.523	0.467	0.547	0.171*
E19	Bezděz	N50°30'44"	E14°41'46"	20	5.2	0.570	4.315	0.489	0.551	0.141*
E20	Petrovice	N50°48'58"	E14°46'34"	20	5.4	0.573	4.338	0.473	0.555	0.174*
E21	Česká Kamenice	N50°48'36"	E14°27'35"	20	4.7	0.479	3.793	0.367	0.463	0.234*
E22	Stráž pod Ralskem, Ralsko	N50°40'53"	E14°51'36"	20	5.2	0.585	4.500	0.455	0.565	0.221*
E23	Jablonec nad Nisou	N50°43'06"	E15°07'03"	20	5.8	0.575	4.546	0.396	0.555	0.312*
E24	Dvůr Králové nad Labem	N50°24'45"	E15°47'50"	20	5.2	0.526	4.131	0.400	0.509	0.241*
E25	Zbiroh	N49°50'03"	E13°48'45"	20	5.1	0.560	4.075	0.472	0.544	0.157*
E26	Plzeň	N49°47'43"	E13°22'37"	20	5.3	0.596	4.443	0.482	0.577	0.191*
E27	Horní Štěpánov	N49°32'37"	E16°46'12"	20	4.8	0.557	4.070	0.399	0.537	0.284*
E28	Huntířov, Bynovec	N50°48'01"	E14°18'08"	20	5.1	0.560	4.121	0.455	0.543	0.187*
E29	Srbská Kamenice	N50°50'01"	E14°22'08"	20	5.6	0.527	4.589	0.446	0.511	0.154*
E30	Sopřeč	N50°04'26"	E15°32'54"	20	4.8	0.547	4.075	0.429	0.528	0.216*
A31	NY, Millbrook	N41°47'38"	W73°43'59"	20	5.0	0.591	4.145	0.461	0.572	0.221*
A32	NY, New Paltz	N41°45'42"	W74°09'46"	20	5.2	0.476	4.097	0.422	0.462	0.114*
A33	NY, Albany	N42°35'27"	W74°00'28"	20	5.3	0.546	4.378	0.368	0.526	0.326*
A34	PA, Wellsboro	N41°41'48"	W77°27'18"	20	5.5	0.552	4.381	0.422	0.534	0.236*
A35	PA, Clarington	N41°20'27"	W79°08'00"	20	5.6	0.522	4.409	0.352	0.504	0.326*
A36	PA, Warren	N41°42'04"	W79°14'56"	20	4.4	0.516	3.782	0.433	0.497	0.161*
A37	NY, Castile	N42°36'28"	W78°01'30"	20	5.2	0.587	4.386	0.442	0.568	0.246*
A38	NY, Bergen	N43°05'24"	W78°02'00"	20	5.3	0.574	4.335	0.507	0.557	0.117*
A39	NY, Pulaski	N43°33'43"	W76°11'40"	20	5.9	0.526	4.515	0.400	0.509	0.240*
A40	NY, Tupper Lake	N44°18'41"	W74°43'16"	20	5.9	0.516	4.478	0.411	0.499	0.203*
A41	NY, Wilmington	N44°21'46"	W73°50'32"	20	5.7	0.559	4.575	0.468	0.540	0.162*
A42	VT, Johnson	N44°39'11"	W72°43'38"	20	5.8	0.577	4.673	0.417	0.557	0.277*
A43	NH, Bartlett	N44°04'30"	W71°18'17"	20	5.4	0.499	4.184	0.444	0.485	0.111*
A44	ME, Freeport	N43°49'22"	W70°05'36"	20	5.6	0.537	4.316	0.396	0.519	0.262*
A45	ME, Guilford	N45°15'53"	W69°16'55"	20	5.4	0.560	4.422	0.405	0.542	0.277*
A46	ME, Millinocket	N45°35'47"	W68°48'51"	20	5.2	0.548	4.154	0.467	0.532	0.149*
A47	ME, Seal Cove	N44°16'40"	W68°22'45"	20	4.1	0.484	3.376	0.468	0.471	0.033
A48	ME, New Harbor	N43°54'48"	W69°29'12"	20	5.5	0.567	4.459	0.461	0.548	0.187*
A49	MA, Chelmsford	N42°33'31"	W71°20'55"	20	5.3	0.536	4.359	0.473	0.521	0.117*

**Table 1. Cont.**

Pop.	Locality	Latitude	Longitude	<i>N</i>	<i>A</i>	<i>H<sub>S</sub></i>	<i>R<sub>S</sub></i>	<i>H<sub>O</sub></i>	<i>H<sub>E</sub></i>	<i>f(F<sub>IS</sub>)</i>
A50	MA, Taunton	N41°58'42"	W71°04'52"	20	5.5	0.548	4.311	0.466	0.532	0.150*
A51	VT, Manchester	N43°07'41"	W73°07'05"	20	5.9	0.559	4.659	0.487	0.543	0.128*
A52	MA, Shelburne Falls	N42°37'30"	W72°46'13"	20	5.8	0.551	4.653	0.453	0.533	0.179*
A53	CT, Mansfield Center	N41°45'60"	W72°10'47"	20	6.3	0.600	4.898	0.478	0.582	0.203*
A54	CT, Burlington	N41°46'14"	W72°57'13"	20	5.0	0.497	4.070	0.443	0.483	0.109*
A55	MI, Alger	N44°04'04"	W84°10'58"	20	5.2	0.559	4.306	0.489	0.543	0.125*
A56	MI, Pellston	N45°32'54"	W84°41'58"	20	5.2	0.602	4.458	0.433	0.581	0.280*
A57	MI, White Pine	N46°44'89"	W89°42'48"	20	6.2	0.592	4.843	0.453	0.572	0.235*
A58	MI, Marquette	N46°37'21"	W87°28'20"	20	5.5	0.56	4.490	0.486	0.543	0.132*
A59	MI, Sault Ste. Marie	N46°21'23"	W84°08'22"	20	6.1	0.536	4.884	0.432	0.518	0.195*
A60	MI, Bridgman	N41°54'17"	W86°36'25"	12	4.6	0.501	4.302	0.367	0.475	0.268*
Mean					5.3	0.549	4.302	0.437	0.531	0.204
SD					0.5	0.033	0.300	0.038	0.032	0.061

Pop. = Population – the first letter of each population name refers to the continent in which the population was found, i.e. E = Europe (Czech Republic) and A = North America; populations in bold were classified as invasive within the adventive range; GPS coordinates of sampling sites are in WGS84; *N* – number of individuals sampled from each population; *A* – average number of alleles per locus; *H<sub>S</sub>* – mean gene diversity over all loci; *R<sub>S</sub>* – mean allelic richness; *H<sub>O</sub>* – observed heterozygosity; *H<sub>E</sub>* – expected heterozygosity; *f(F<sub>IS</sub>)* – inbreeding coefficient according to Weir and Cockerham [77]. Populations deviating from Hardy-Weinberg equilibrium at  $P < 0.05$  are marked by asterisk. SD – standard deviation. The fine scale studies were performed at localities E1, E22, E28 and E30.

doi:10.1371/journal.pone.0068514.t001

locus) and treated as missing data after at least two amplification failures. The diagnostic results using MICRO-CHECKER [74] found no evidence of stuttering or large allele drop-out for any of the loci. However, the potential occurrence of null alleles was detected in loci RPS25b and RPS34b, which mostly corresponded to the deviation from Hardy-Weinberg equilibrium.

### Statistical Analysis – Large-scale Study

**Genetic diversity measures.** Summary data for SSR loci, including the average number of alleles per locus (*A*), mean gene diversity overall loci (*H<sub>S</sub>*), mean allelic richness (*R<sub>S</sub>*) (here allelic richness is a metric that uses a rarefaction index to take into account differences in sample size [75,76]) and Weir & Cockerham's parameter *f(F<sub>IS</sub>)* [77], a measure of deviation from random mating within a population, were calculated using FSTAT [76]. Observed (*H<sub>O</sub>*) and expected (*H<sub>E</sub>*) heterozygosities were calculated using the program Arlequin [78], and deviation from the Hardy-Weinberg equilibrium was determined on the basis of 10,000 permutations in FSTAT. Sequential Bonferroni corrections were applied to adjust *P* value according to Rice [79].

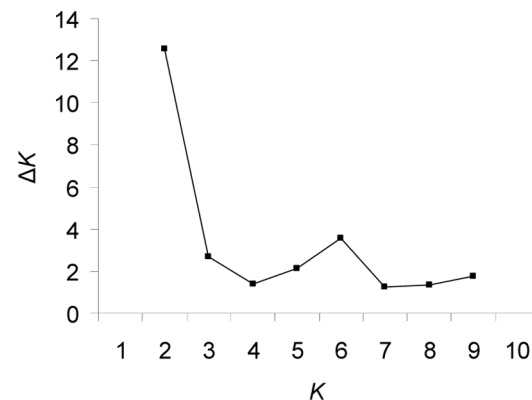
Weir & Cockerham's [77] estimates of Wright's [80] *F* statistics were generated for all loci. Significant deviations from the null expectation of  $F = 0$  were determined by 5000 bootstrap replicates using FSTAT. In the bootstrap analysis, *F* (corresponding to Wright's *F<sub>IT</sub>*) was estimated by alleles permuted across the entire dataset, *f(F<sub>IS</sub>)* was estimated by the permutation of alleles within populations, and  $\theta(F_{ST})$  was estimated by the permutation of alleles among populations. The allele frequency-based *F<sub>ST</sub>* was used rather than the allele size-based *R<sub>ST</sub>* of Slatkin [81], which is derived specifically under the assumptions of the generalized stepwise-mutation model (SMM). *R<sub>ST</sub>* and *F<sub>ST</sub>* values are not expected to differ greatly for short-term differentiation of populations within species [81], but *F<sub>ST</sub>*-based estimates of differentiation are considered more reliable when fewer than 20 loci are used [82].

Comparisons of genetic diversity parameters between groups (i.e. native North American and non-native European populations

or invasive and naturalized non-native populations) were performed with FSTAT with 10,000 permutations.

We used the Mantel test to assess the model of isolation-by-distance using genetic distance for pairs of populations [83] and geographic distance among these populations (calculated using FSTAT). To test differences in allele frequencies, we used the exact test for population differentiation [84]. This analysis uses a contingency table approach (Fisher's  $R \times C$ -test) [85] to determine whether significant differences in allele frequencies exist among groups of individuals (calculated using TFPGA, <http://www.marksgeneticssoftware.net>).

**Bottleneck detection.** We tested for an excess of heterozygosity, which reveals the loss of rare alleles in a bottlenecked population relative to that expected under the mutation-drift equilibrium (neutrality) for an observed level of heterozygosity [86], using microsatellite frequency data. In a population which



**Figure 2. Estimation of the number of genetic clusters using Bayesian assignment technique.** Second-order rate of change in the probability between successive runs ( $\Delta K$ ) as a function of *K* (number of clusters).

doi:10.1371/journal.pone.0068514.g002



has recently been reduced in size, both the number of alleles (allelic diversity) and heterozygosity would likely be reduced. However, allelic diversity would be reduced faster than heterozygosity, resulting in deficiency in the observed number of alleles relative to the number of alleles expected from the observed heterozygosity, providing the basis for this test. Microsatellite data were processed using the BOTTLENECK program [86,87]. We considered the two-phase model (TPM), which is probably closer to the true mode of mutation at most microsatellite loci [87]. The proportion of alleles attributed to a stepwise-mutation model (SMM) under TPM was set to 85%, with a variance of 5. Ten thousand iterations were run. The one-tailed Wilcoxon signed rank test for heterozygote excess was applied as a test of significance [86,87], and the distribution of allele frequencies was tested against the L-shaped distribution, as expected under the mutation-drift equilibrium [88].

**Population structure and identification of sources.** We used Structure version 2.3.3 software [89] to estimate the number of genetic clusters ( $K$ ) and to fractionally assign individuals sampled in North America and Europe to the inferred groups. We applied the model which allows population admixture with prior population information [90], and correlated allele frequency [89]. The model with prior population information was used because of a weak population structure [90]. The number of clusters ( $K$ ) was set at each value from one through ten, and the simulation was run ten times at each  $K$  value to confirm the repeatability of the results. Each run comprised a burn-in period of 150,000, followed by 300,000 Markov chain Monte Carlo (MCMC) steps. Of the various ways to estimate the 'true' number of genetic clusters using Bayesian assignment techniques [89,91,92], we used the one [91] in which the quantity  $\Delta K$  is employed as an *ad hoc* estimator of the second order rate of change of the  $\ln P(K|X)$  to furnish an initial estimate of  $K$  (see Fig. 2). Per Evanno *et al.* [91], our modelling demonstrated that the peak (modal) value(s) of  $\Delta K$  were good estimates of  $K$  in simulations with a range of known population sizes and types and numbers of loci (but see [92]). The output of structure analyses was visualized using CLUMPP [93] and DISTRICT [94] software.

**Ecological factors.** The effects of ecological factors (elevation above sea level, average annual precipitation and temperature and soil types) on *P. strobus* population regeneration ability were analysed by Logit models implemented in the Generalized Linear/Nonlinear Models module of the software package STATISTICA 9.1 (<http://www.statsoft.com>).

### Statistical Analysis – Fine-scale Study

**Genetic diversity measures.** Genetic diversity parameters were calculated as described in the large-scale study (see above). Comparisons of genetic diversity parameters among populations were performed with FSTAT with 10,000 permutations. Pairwise  $F_{ST}$  values were used to compare differences in gene frequencies among different age classes within each locality with FSTAT, and a Bonferroni correction was used to adjust observed  $P$  values for multiple comparisons.

Genetic variation at the level of populations and age classes was investigated with a nested analysis of molecular variance (age classes nested within populations; AMOVA – [78]). Levels of significance were determined by computing 1,000 random permutation replicates.

**Spatial Genetic Structure (SGS).** Fine-scale genetic structure was assessed by spatial autocorrelation analysis of genetic relatedness between pairs of individuals as described in [95,96] and defined as a ratio of probabilities of identity by descent. Assuming migration-drift equilibrium, the impact of mutations at

small spatial scales is negligible relative to that of gene flow, and the ratio of probabilities of identity in state is similar to that of probabilities of identity by descent [97]. A pairwise estimate of genetic correlation,  $f_{ij}$  (co-ancestry coefficient; sensu Kalisz *et al.* [96]) measures the frequency of allele sharing between two individuals relative to average frequency of allele sharing based on the frequencies of the alleles in the sampled population [98]. Co-ancestry ( $f_{ij}$ ) was estimated between all pairs of individuals within each age class following the methods of Loiselle *et al.* [95] and Kalisz *et al.* [96]. To analyse the relationships between pairwise physical distances and co-ancestry ( $f_{ij}$ ) coefficient, we calculated  $f_{ij}$  for 5 m distance intervals up to 100 m. Mean values of  $f_{ij}$  were obtained for distance intervals of 5 m by averaging over all pairs of individuals located within that interval. Spatial autocorrelation analysis was also computed for all individuals within each population irrespectively of age class and among age classes within each population to test whether the spatial structure of individual age classes were dependent on each other. When  $f_{ij} = 0$ , there is no significant correlation among individuals at the spatial scale of interest; when  $f_{ij} > 0$ , individuals in a given distance class are more closely related than expected by chance; and when  $f_{ij} < 0$ , individuals within a given distance class are less related than expected by chance with respect to the local population. Co-ancestry coefficients were computed using SPAGeDi software [99].

We also evaluated the strength of isolation by distance using  $S_p$  statistics, calculated as  $-b/(1-F_{(1)})$ , where  $F_{(1)}$  is the  $f_{ij}$  for the first distance class.  $F_{(1)}$  can be considered an approximation of the kinship between pairs of neighbours, provided the first distance class contains enough pair of individuals to obtain a reasonably precise  $F_{(1)}$  value [42].

## Results

### Large-scale Study

**Genetic diversity, population structure and regional differentiation.** While tests for genotypic linkage disequilibrium were all non-significant after applying sequential Bonferroni correction, tests of deviation from the Hardy-Weinberg equilibrium (Table 1) were all significant with the exception of locality A47. We identified 178 alleles at ten microsatellite loci, with an average of 17.8 alleles per locus across all populations. Populations in the adventive range had a very similar average number of alleles per locus and population to that of native ones ( $A = 5.1$  vs. 5.4)

**Table 2.** Comparison of population genetic parameters between native and adventive range.

Diversity measure	Native	Adventive	$P$
$R_S$	4.227	4.377	0.052
$H_O$	0.432	0.440	0.418
$H_S$	0.550	0.543	0.452
$f(F_{IS})$	0.215	0.191	0.129
$\theta(F_{ST})$	0.025	0.030	0.513

Statistical comparison of allelic richness ( $R_S$ ), observed heterozygosity ( $H_O$ ), gene diversity ( $H_S$ ), inbreeding coefficient  $f(F_{IS})$ , and levels of differentiation among populations  $\theta(F_{ST})$  for the native and adventive ranges of *Pinus strobus*.

Probability values for differences between the native and the adventive range are given for two-sided  $t$ -test, after 10,000 permutations. The analysis was performed using FSTAT software.

doi:10.1371/journal.pone.0068514.t002

**Table 3.** Genetic variation and structure for 10 polymorphic loci identified in 60 *Pinus strobus* populations.

Locus	Total			Native range			Adventive range		
	$f(F_{IT})$	$f(F_{IS})$	$\theta(F_{ST})$	$f(F_{IT})$	$f(F_{IS})$	$\theta(F_{ST})$	$f(F_{IT})$	$f(F_{IS})$	$\theta(F_{ST})$
RPS1b	<b>0.089</b>	<b>0.078</b>	<b>0.013</b>	<b>0.104</b>	<b>0.103</b>	0.002	<b>0.074</b>	0.055	<b>0.020</b>
RPS25b	<b>0.615</b>	<b>0.589</b>	<b>0.064</b>	<b>0.624</b>	<b>0.600</b>	<b>0.062</b>	<b>0.605</b>	<b>0.578</b>	<b>0.062</b>
RPS39	<b>0.261</b>	<b>0.250</b>	<b>0.014</b>	<b>0.276</b>	<b>0.271</b>	0.006	<b>0.253</b>	<b>0.236</b>	<b>0.023</b>
RPS84	<b>0.039</b>	0.013	<b>0.026</b>	0.041	0.032	<b>0.009</b>	0.016	0.001	<b>0.015</b>
RPS2	0.002	-0.021	<b>0.022</b>	-0.032	-0.040	<b>0.008</b>	0.031	0.011	<b>0.020</b>
RPS12	<b>0.197</b>	<b>0.175</b>	<b>0.027</b>	<b>0.168</b>	<b>0.154</b>	<b>0.015</b>	<b>0.216</b>	<b>0.196</b>	<b>0.025</b>
RPS34b	<b>0.318</b>	<b>0.286</b>	<b>0.045</b>	<b>0.394</b>	<b>0.369</b>	<b>0.039</b>	<b>0.219</b>	<b>0.190</b>	<b>0.035</b>
RPS50	<b>0.059</b>	<b>0.024</b>	<b>0.036</b>	<b>0.028</b>	0.003	<b>0.025</b>	<b>0.071</b>	<b>0.045</b>	<b>0.027</b>
RPS118b	<b>0.322</b>	<b>0.306</b>	<b>0.023</b>	<b>0.409</b>	<b>0.402</b>	<b>0.012</b>	<b>0.228</b>	<b>0.208</b>	<b>0.024</b>
RPS127	<b>0.093</b>	<b>0.063</b>	<b>0.031</b>	0.055	0.019	<b>0.037</b>	<b>0.130</b>	<b>0.112</b>	<b>0.020</b>
Over all loci	<b>0.228</b>	<b>0.202</b>	<b>0.033</b>	<b>0.234</b>	<b>0.215</b>	<b>0.025</b>	<b>0.215</b>	<b>0.191</b>	<b>0.030</b>

$F$ ,  $f$ ,  $\theta$  = Weir and Cockerham's estimates of Wright's  $F$  statistics ( $F_{IT}$ ,  $F_{IS}$  and  $F_{ST}$ , respectively), which represent deviations from Hardy-Weinberg expectations over all populations, deviations within individual populations and the proportion of total genetic diversity partitioned among populations. Significant deviations ( $P < 0.05$ ) from the null expectation of  $F = 0$  are indicated in bold.

doi:10.1371/journal.pone.0068514.t003

(Table 1). No genetic diversity measures differed significantly between the native and adventive ranges, with the exception of allelic richness ( $R_S$ ), where the difference was marginally significant (Table 2), suggesting slightly higher allelic richness in the adventive range. These data show that the genetic diversity found within the relatively small adventive range, surprisingly, equals the diversity found within the sampled area in the native range, which is about thirty times larger (Fig. 1). Only 135 alleles were observed in the adventive range compared to 159 alleles found in native populations. The difference is due to the higher number of rare alleles in the native range that are absent in the adventive range. Therefore, there were more unique genotypes (i.e. combinations of alleles at individual loci) present in the native range (31.3% of individuals had at least one unique genotype at a locus compared to 26.4% in the adventive range). The value of  $f(F_{IS})$  was lowest in native population A47 (0.033), and highest in native populations A33 and A35 (0.326), which indicates the existence both of populations in Hardy-Weinberg equilibrium and of populations showing excesses of homozygotes, suggesting some level of inbreeding (Table 1).

To investigate the pattern of inbreeding and population differentiation further, we analysed genetic variation and structure for individual microsatellite loci and over all loci separately for native and non-native populations. The mean value of  $f(F_{IS})$  was quite high and significantly different from zero in most cases within both ranges (Table 3), with a tendency of populations in the adventive range to be less inbred than native populations. Significant differentiation among populations in both ranges was detected, with slightly higher  $\theta(F_{ST})$  values reached by populations in the adventive range (Table 3). However, neither  $f(F_{IS})$  nor  $\theta(F_{ST})$  differed significantly between the native and adventive ranges (Table 2). When we compared allele frequencies between the native and adventive ranges, we obtained significant differences at all individual loci (data not shown) as well as over all loci ( $\chi^2 = 269.26$ , d.f. = 20,  $p < 10^{-4}$ ).

There was no consistent association between genetic and geographic distances in both the native and the adventive range, as indicated by the Mantel tests ( $r = -0.003$ ,  $R^2 = 0.001$ ,  $P = 0.947$  and  $r = -0.081$ ,  $R^2 = 0.66$ ,  $P = 0.088$ , respectively).

**Bottleneck.** We found no evidence for an excess of heterozygotes in any of the populations tested using the TMP mutation model. Allele frequencies followed an L-shaped distribution; that is, no skews of allele frequencies towards intermediate values were observed in any of the populations, with such a finding expected from a non-bottlenecked population at mutation-drift equilibrium. This indicates that the populations in the adventive range were not exposed to strong bottlenecks in their recent history (data not shown). We note, however, that the heterozygote excess/deficit test has been known to miss bottlenecks, as it was accurate in only 50–75% of test studies [88].

**Population structure and identification of sources.** In the Bayesian clustering analysis,  $\Delta K$  indicated that two clusters best explained the genetic structuring of *P. strobus* populations (Fig. 2). Assignments to more than two clusters did not provide unequivocal and easily interpretable results, as many clusters were mixed together, indicating that quite a high amount of genetic material from genetically very diverse population(s) was introduced from North America to Europe. Hence, populations were assigned to two clusters corresponding mainly to the native and the adventive ranges (Fig. 3).

**Regeneration ability of *Pinus strobus* within the adventive range.** The regeneration ability (in terms of the categorization scheme described in Materials and Methods) of the studied Czech populations was not affected by elevation above sea level, precipitation, temperature, or soil types, as indicated by Logit models (data not shown).

Invasive and naturalized populations differed in gene diversity ( $H_S$ ), although the differences were very weak (Table 4). Otherwise, they did not differ significantly in allelic richness ( $R_S$ ), observed heterozygosity ( $H_O$ ), inbreeding  $f(F_{IS})$  or fixation  $\theta(F_{ST})$  coefficients (Table 4). Furthermore, invasive and naturalized populations differed significantly in allele frequencies ( $\chi^2 = 65.33$ , d.f. = 20,  $p < 10^{-4}$ ).

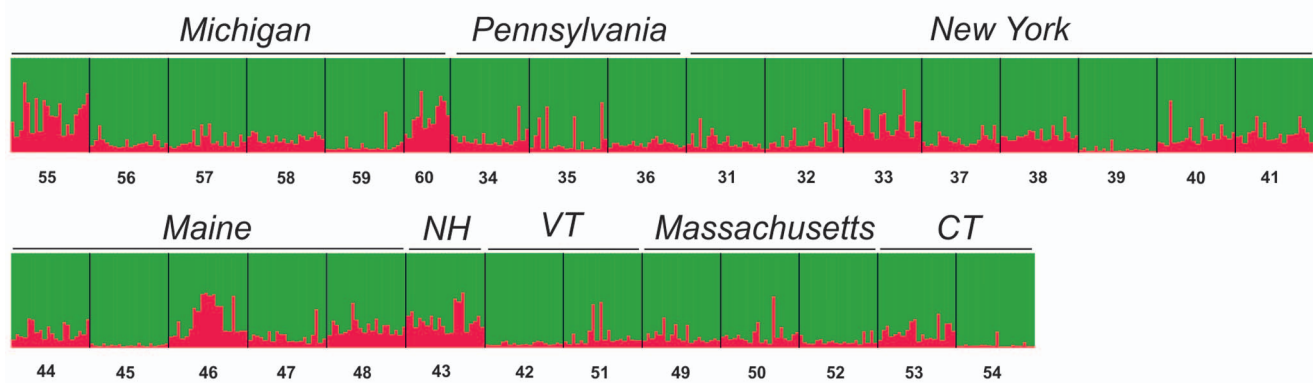
### Fine-scale Study

**Genetic diversity of the invasive populations.** A total of 954 individuals from four localities were analysed across four age classes (Table 1). All 10 microsatellite loci analysed were polymorphic. While our comparison of genetic diversity param-

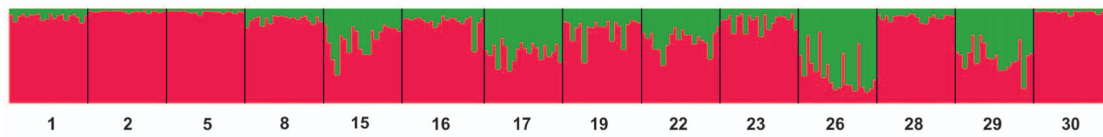


## Native range

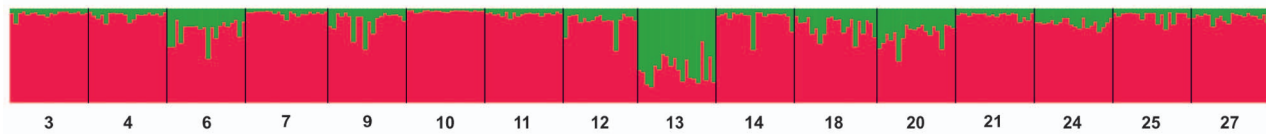
K = 2



## Invasive populations within the adventive range



## Naturalized populations within the adventive range



**Figure 3. Results of assignment test.** Percentage assignment of each North American and European individual (represented by vertical bars) to each of the two clusters (represented by different colors) inferred by the program Structure. Site codes (see Table 1) indicate the geographical location of individuals along the x-axis.  
doi:10.1371/journal.pone.0068514.g003

eters among four localities of *Pinus strobus* (two-sided *t*-test, after 10,000 permutations; the analysis performed using FSTAT) showed similarities for allelic richness ( $R_S$ ,  $p = 0.796$ ), gene diversity ( $H_S$ ,  $p = 0.130$ ) and fixation indices [ $\theta(F_{ST})$ ,  $p = 0.977$ ], there was a significant difference in observed heterozygosity ( $H_O$ ) and inbreeding coefficient [ $f(F_{IS})$ ,  $p = 0.004$  and  $p = 0.005$ , respective-

ly]. Especially the locality Hrad showed a lower level of heterozygosity and consequently higher inbreeding (Table 5).

Regression analyses of parameters of genetic diversity, i.e. allelic richness ( $R_S$ ), observed heterozygosity ( $H_O$ ), gene diversity ( $H_S$ ) and inbreeding coefficient ( $F_{IS}$ ) over time, i.e. from old trees to seedlings, with the dependent factor being the genetic diversity parameter and the independent one being time, were not statistically significant (data not shown).

A hierarchical analysis of molecular variance (AMOVA) revealed that a majority of the genetic diversity was partitioned within individuals (79.76%,  $P < 10^{-6}$ ). Only 19.25% ( $P < 10^{-6}$ ) was partitioned among individuals within age classes and 0.007% ( $P = 0.446$ ) was partitioned among age classes within localities. The remaining 0.97% ( $P < 10^{-6}$ ) of the variation was due to differences among localities.

Pairwise  $F_{ST}$  values revealed that allele frequencies were not significantly different among individual age classes at all localities but one (Hrad). Allele frequencies at the locality Hrad significantly differed between seedlings and mature trees ( $P = 0.01$ ), seedlings and old trees ( $P = 0.05$ ) and between saplings and mature trees ( $P = 0.05$ ) (data not shown).

**Spatial Genetic Structure (SGS).** Our results indicated differences in SGS for different age classes within the investigated populations (Figs. 4, 5). In old trees, i.e. trees planted by foresters that founded the populations more than 100 years ago, we found no positive autocorrelation in any distance class or in any population. Mature trees showed positive autocorrelations at 10 and 25 m distances at the localities Bynovec and Hrad,

**Table 4. Comparison of population genetic parameters between invasive and naturalized populations in the adventive range.**

Diversity measure	Invasive	Naturalized	P
$R_S$	4.330	4.137	0.073
$H_O$	0.447	0.423	0.081
$H_S$	0.565	0.540	<b>0.034</b>
$F_{IS}$	0.210	0.217	0.734
$F_{ST}$	0.019	0.032	0.256

Statistical comparison of allelic richness ( $R_S$ ), observed heterozygosity ( $H_O$ ), gene diversity ( $H_S$ ), inbreeding coefficient ( $F_{IS}$ ), and levels of differentiation among populations ( $F_{ST}$ ) for the invasive ( $N = 14$ ) and naturalized populations ( $N = 16$ ) of *Pinus strobus* within its adventive range in Europe. Probability values for differences between invasive and naturalized populations are given for two-sided *t*-test, after 10,000 permutations. The analysis was performed using FSTAT software. Significant differences are indicated in bold.  
doi:10.1371/journal.pone.0068514.t004

**Table 5.** Summary of population genetic characteristics of *Pinus strobus* populations at which fine scale genetic study was performed.

Locality	Age class	No.	A	R <sub>S</sub>	H <sub>O</sub>	H <sub>S</sub>	f(F <sub>IS</sub> )
Bynovec	Old trees	50	6.0	4.794	0.468	0.569	0.179
	Mature trees	50	6.1	4.813	0.482	0.568	0.152
	Saplings	100	6.5	4.748	0.480	0.572	0.163
	Seedlings	50	5.7	4.738	0.473	0.568	0.169
	Overall	250	6.1	4.773	0.476	0.569	0.166
Hrad	Old trees	40	6.8	5.303	0.463	0.564	0.181
	Mature trees	50	7.6	5.340	0.434	0.565	0.234
	Saplings	100	6.5	4.790	0.412	0.565	0.272
	Seedlings	50	5.4	4.522	0.403	0.536	0.249
	Overall	240	6.6	4.989	0.428	0.558	0.234
Ralsko	Old trees	22	5.9	5.454	0.468	0.587	0.208
	Mature trees	49	5.7	4.580	0.474	0.557	0.150
	Saplings	100	6.5	4.694	0.435	0.552	0.212
	Seedlings	50	5.3	4.477	0.447	0.536	0.168
	Overall	221	5.9	4.801	0.456	0.558	0.185
Sopřeč	Old trees	43	5.6	4.726	0.472	0.567	0.170
	Mature trees	50	6.6	5.212	0.448	0.559	0.200
	Saplings	100	6.8	4.738	0.461	0.553	0.168
	Seedlings	50	6.1	4.750	0.455	0.563	0.194
	Overall	243	6.3	4.856	0.459	0.561	0.183

Number of individuals of *Pinus strobus* collected and analysed (No.) at different localities in the Czech Republic, average number of alleles per locus (A), allelic richness (R<sub>S</sub>), observed heterozygosity (H<sub>O</sub>), gene diversity (H<sub>S</sub>) and inbreeding coefficient f(F<sub>IS</sub>) according to [77] for 10 microsatellite loci of four recognized age classes – “old trees” (50 years old and older), “mature trees” (between 21 and 49 years old), “saplings” (between 6 and 20 years old) and “seedlings” (trees not older than 5 years) computed by Arlequin and FSTAT. All populations significantly deviated from the Hardy-Weinberg equilibrium calculated using 50,000 permutations in FSTAT.

doi:10.1371/journal.pone.0068514.t005

respectively (Figs. 4C, I). Saplings were significantly structured in all tested localities (Figs. 4, 5, Table 6). In most cases, positive autocorrelations were in distance classes between 5 and 25 m. Populations of seedlings were not so strongly structured as saplings and showed positive autocorrelations at three localities, at two of them (Hrad and Ralsko) in 15 m distance classes and in 5 m distance class at the other (Sopřeč) (Figs. 4, 5, Table 6). The slope of regressions (*b*) between the kinship coefficient (*f<sub>ij</sub>*) and the log of the physical distance between individuals as well as *S<sub>p</sub>* statistics corroborate these results (Table 6). When we computed the SGS irrespectively of age classes, we obtained positive autocorrelations at all localities on distances usually up to 25 m (Figs. 4, 5). Similarly, there were positive autocorrelations among individuals of different age classes at three localities (Figs. 4, 5, Table 6), suggesting dispersal distances of *Pinus strobus* seeds at different localities to be around 15 m (Figs. 4, 5).

## Discussion

This study reports for the first time the patterns of both large-scale (between continent comparison) and fine-scale (spatial genetic structure within populations from outside of the species' native distribution range) genetic diversity in invasive tree *Pinus strobus*. So far, relevant studies have focused on the species' native

distribution range and have dealt with population genetic structure [100–105] or spatial genetic structure [47,68,69,106,107]. The main findings of the present study are:

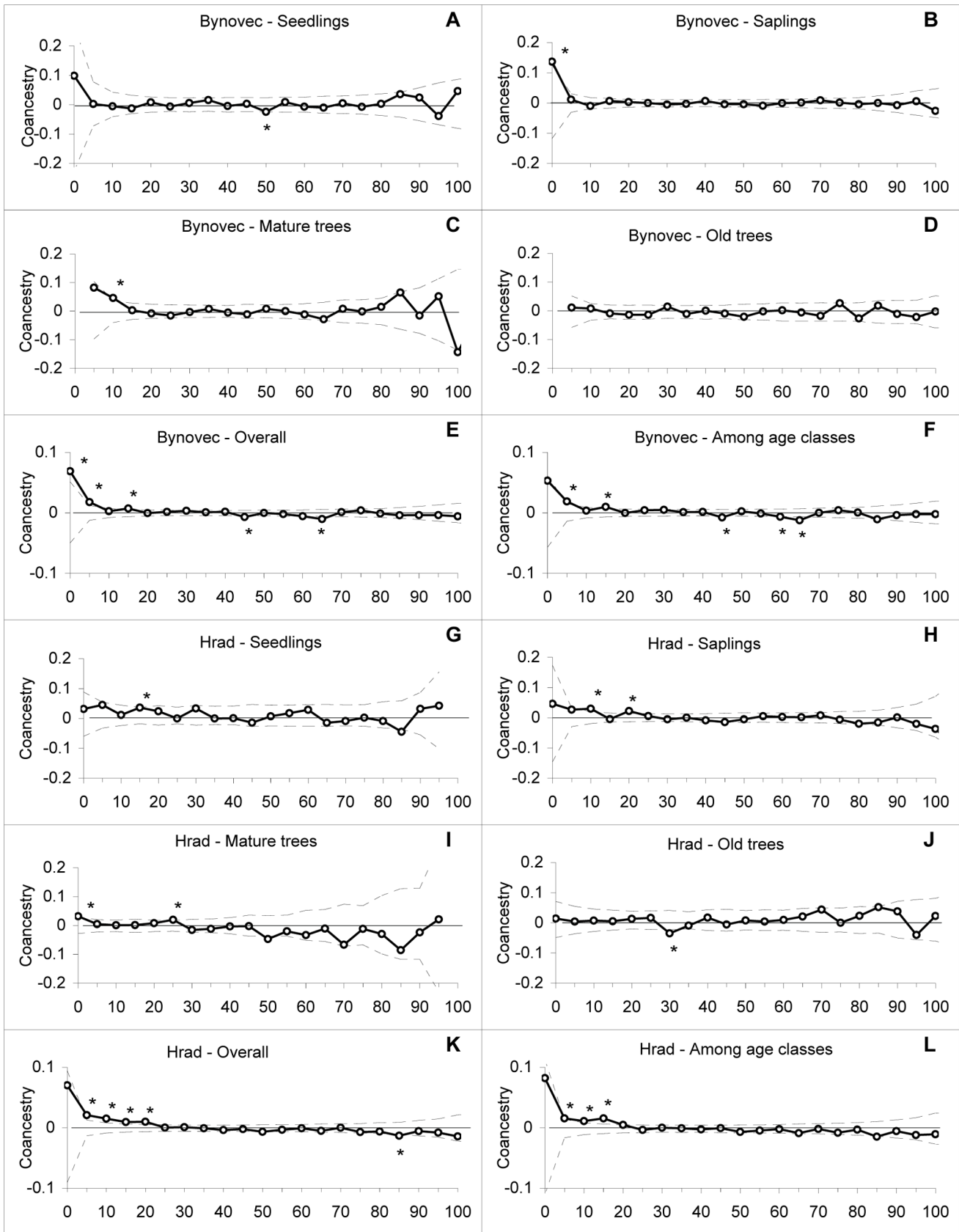
- (1) a comparably high level of genetic diversity present in the native and the adventive ranges, with an absence of strong large-scale genetic structure in both distribution ranges (Table 2). In other words, the amount of genetic diversity detected over a significant part of the native range was also present in the relatively tiny area of the species' adventive range in the Czech Republic (see Fig. 1).
- (2) The genetic diversity of trees introduced to the adventive localities approximately one hundred years ago was high enough to ensure the maintenance of high genetic diversity. On the fine scale we found no evidence for any increase in genetic diversity parameters over time; however, we observed the development of fine-scale genetic clustering over time.

## Large-scale Study

**Genetic diversity in the native vs. the adventive range.** The finding of comparably high levels of genetic diversity has two important implications. First, the amount of genetic diversity present in the adventive range is so high that it is difficult to envisage a severe demographic bottleneck having occurred either during or after the introduction to Europe. This finding is consistent with the supply of large amounts of seeds from diverse sources since the beginning of the introduction process (see also below).

Second, even though theory predicts high genetic diversity in the adventive range to be a consequence of multiple introductions, the pattern of genetic diversity in the native range can provide clearer insight into the invasion history. *Pinus strobus* populations in the native range show a pattern of genetic diversity distribution that is less structured than that of many invasive plant species studied so far (but see [108]). Hence, high genetic diversity in the adventive range could be the result of an introduction from one genetically very diverse source or multiple introductions from different sources, or a combination of both. The question arises whether such high genetic diversity could have been generated *in situ* after the introduction of the species in the adventive range. Analysis of fine-scale genetic diversity (see below) within the four localities in the adventive range shows how genetic diversity changed through the invasion process. We did not find any increase over time in the genetic diversity parameters measured, further supporting our explanation that the presence of a high genetic diversity in the adventive range is due to massive introduction of genetically highly diverse material since the beginning of the invasion process. Based on the absence of a difference in genetic diversity between the native and the adventive ranges, together with the lack of structure in the native range, we propose that in this species, genetic structure in the native range did not play a significant role for invasion success in the adventive range as in invasive species having highly genetically structured populations [24,27].

**Population structure and identification of sources.** Our Bayesian assignment test identified two genetic clusters (Fig. 2). They weakly separated North American native populations from the European, non-native populations (Fig 3). A different picture is presented by the less supported schemes of three or more genetic clusters, in that most individuals were not strongly assigned to any single genetic cluster. Therefore, non-native populations either descended from a cluster that is not represented within the current sample from North America or allele frequencies were strongly



**Figure 4. Correlograms of  $f_{ij}$  coefficients for the localities Bynovec and Hrad.** Correlograms of  $f_{ij}$  coefficients for the locality Bynovec within (A) seedlings, (B) saplings, (C) mature trees, (D) old trees, (E) overall data and (F) among all age classes; for the locality Hrad within (G) seedlings, (H) saplings, (I) mature trees, (J) old trees, (K) overall data and (L) among all age classes. The solid line plots the observed data, and the dotted lines

indicate the 95% confidence interval deduced from 10,000 permutations of individual multilocus genotypes within each age class, overall age classes or among all age classes. Values with asterisks (\*) above and below the confidence envelopes indicate a greater or lesser genetic structure among individuals at different distances than expected of individuals chosen at random ( $P < 0.005$ ). doi:10.1371/journal.pone.0068514.g004

shaped during the introduction of genotypes to Europe, resulting in the creation of a novel cluster as a result of invasion. Although we cannot exclude the first scenario despite our sampling design covering a significant part of the native range, our data rather support the latter explanation. Allele frequencies differ between the native and the adventive ranges at all individual loci as well as over all loci, indicating introduction of different individuals in different frequencies.

Interestingly, we did not detect any strong structure even within the native range. There are basically two non-exclusive explanations that can clarify the pattern. (1) *Pinus strobus*, as do some other conifers, shows low population differentiation and considerably high pollen-mediated gene flow [105,109–114]. If a conifer species spreads its pollen effectively over long distances, then low genetic structure would be detected, i.e. pollen-mediated gene flow alone could explain the patterns observed in the native range. (2) The second explanation is linked to the use of *P. strobus* in the timber industry, resulting in disruption of the genetic integrity of original populations due to admixture of populations from different sources. Historically, by the early 1600s, the first colonists were quick to make use of *P. strobus* forest resources in the eastern part of North America, and by the early 1900s the area of destruction of the primeval white pine forests reached from the Atlantic seaboard through the Great Lake states [115,116]. Some of this cleared land was reforested using seeds and seedlings from different sources, which could have resulted in mixing of genetic material and obscuring the current genetic structure of the species in its native range.

The very low genetic structure of the native populations is worth considering in the light of Taylor and Keller's study [108] of effects of genetic structure in the native range on the genetic diversity of invasive populations in the adventive range. They found that greater phylogeographic structure in a species' native range can increase opportunities for admixture among previously isolated lineages in its adventive range, influencing the evolutionary potential there. However, if mixing of individuals from separate native populations occurs prior to introduction, this would weaken the phylogeographic structure, but allow each subpopulation to contain a large portion of the genetic diversity of the whole population. In this scenario, if a sufficient number of genetically diverse individuals are introduced, it obviates the need for multiple introductions. Then, if population regeneration is ensured and inbreeding prevented, as commonly done by foresters, a loss of genetic diversity might not be apparent in the adventive range.

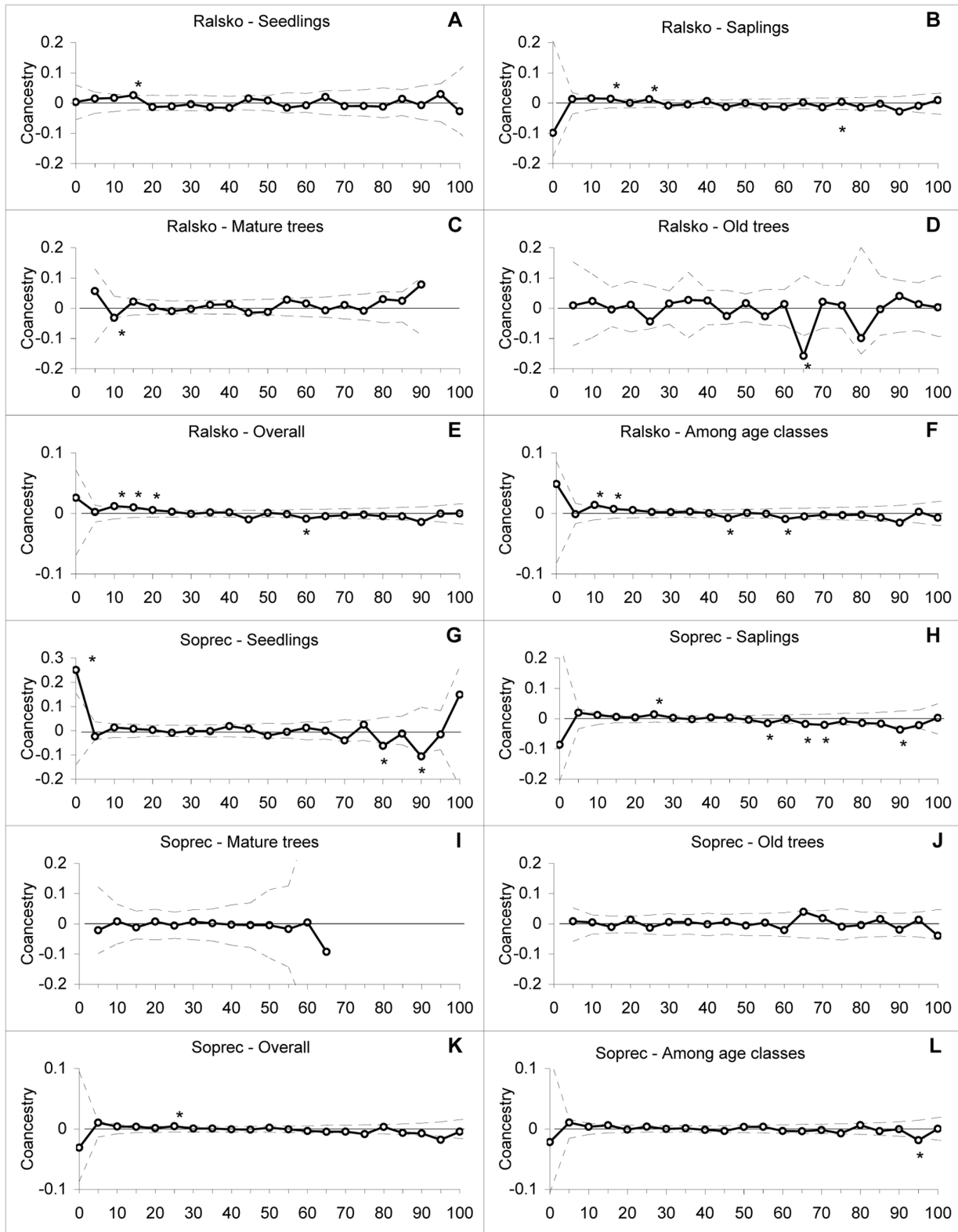
In the case of the introduction of *P. strobus* to the Czech Republic, there is historical evidence of large-scale importations of seeds, which were likely sufficient to largely maintain the genetic diversity of the mixed populations from which they came. For example, there is a historical record that in 1784 10 kg (if the average weight of a *P. strobus* seed is 0.0196 g, than 10 kg is equivalent to introduction of 510,204 seeds) of *P. strobus* seeds were bought and sown in forests in the north-western part of the Czech Republic. These seeds came from Germany with a declared origin in England. Another example is the historical record concerning nearly 32 kg (approximately 1,632,653 seeds) of white pine seeds imported directly from North America in 1800 and planted in the south-eastern part of the Czech Republic [54]. Thus, we know that introductions of new individuals were massive from the very beginning of the cultivation process.

Moreover, we did not see any significant correlation between genetic and geographical distance in either Europe or North America. This pattern is in contrast to studies of other species in which isolation by distance was found in their native ranges and was attributed to a long history of genetic isolation [25]. Our results suggest that present population structure in both the native and the adventive range of *P. strobus* may have been also affected by recent human activities, mainly by the transport of propagules.

**Evolution within the adventive range.** A comparative study between invasive and naturalized populations within the adventive range allows us to assess the consequences of evolutionary events occurring over time. Even if the majority of genetic diversity was introduced from the native to the adventive range, not all introduced populations of *P. strobus* in the Czech Republic can be classified as invasive. Some populations are represented only by old trees and almost completely lack sapling and seedling stages. The test of whether this behavior is influenced by environmental conditions did not reveal any significant relationships. Invasive and naturalized populations only differed in gene diversity ( $H_s$ ) (see Table 4). Although there is a significant difference, it is so weak that it seems to be biologically irrelevant. However, we obtained a highly significant difference in allele frequencies between invasive and naturalized populations. This may have occurred when different alleles were introduced to different populations, and might be an indication of different seed material used for introduction of *P. strobus* to different areas within the Czech Republic or due to pollen dispersal that distributed alleles across populations within small regions.

### Fine-scale Study

**Genetic diversity.** When a species is introduced to a new range from several genetically distinct populations, further recombination within the adventive range may dramatically increase genetic diversity outside the native range [23–25,27]. However, we are not aware of any published study which has examined changes in genetic diversity in an invasive species outside of its native range over time. Such a study would show whether the initial population at the beginning of an invasion needed admixture to reach its present genetic diversity. Trees are an excellent group for making such comparisons, as different age classes present at a locality represent different stages of the invasion process. What disappears very quickly in other plant life forms (i.e. founding genotypes in the lag-phase of the invasion [29,33], is still present in trees and can be readily sampled, aged and genotyped. In this study, we took several highly invasive populations of *Pinus strobus* from different parts of the Czech Republic and explored them in detail to answer this question. By comparing the different age classes in four populations, we found that the initial population of old trees did not suffer from low genetic diversity due to the introduction of a low number of individuals. Conversely, our populations of old trees did not differ from younger age classes, with the exception of the locality Hrad, where we detected lower heterozygosity and a consequently higher inbreeding coefficient ( $F_{IS}$ ). This means that foresters sowed large amounts of seeds that were genetically diverse enough that no admixture was necessary in this case. Even in highly invasive seedlings, which completely fill the space available in the forest, we did not detect any increase in genetic diversity in comparison with



**Figure 5. Correlograms of  $f_{ij}$  coefficients for the localities Ralsko and Sopřeč.** Correlograms of  $f_{ij}$  coefficients for the locality Ralsko within (A) seedlings, (B) saplings, (C) mature trees, (D) old trees, (E) overall data and (F) among all age classes; for the locality Sopřeč within (G) seedlings, (H) saplings, (I) mature trees, (J) old trees, (K) overall data and (L) among all age classes. The solid line plots the observed data, and the dotted lines indicate the 95% confidence interval deduced from 10,000 permutations of individual multilocus genotypes within each age class, overall age classes or among all age classes. Values with asterisks (\*) above and below the confidence envelopes indicate a greater or lesser genetic structure among individuals at different distances than expected of individuals chosen at random ( $P < 0.005$ ).

doi:10.1371/journal.pone.0068514.g005



**Table 6.** Relationships between kinship coefficients and physical distances in *Pinus strobus* populations at which fine scale genetic study was performed.

Locality	Age class	<i>b</i>	CI (95%)	<i>F</i> <sub>1</sub>	<i>Sp</i>
Bynovec	Old trees	-0.0004NS	(-0.009–0.008)	0.0115	0.0004
	Mature trees	-0.0064NS	(-0.013–0.011)	0.0818	0.0070
	Saplings	-0.0014NS	(-0.006–0.005)	0.0106	0.0014
	Seedlings	0.0003NS	(-0.012–0.010)	0.0025	-0.0003
	Among classes	-0.0055*	(-0.002–0.002)	0.0188	0.0056
	Overall	-0.0044*	(-0.002–0.002)	0.0177	0.0044
Hrad	Old trees	0.0060NS	(-0.013–0.011)	0.0144	-0.0061
	Mature trees	-0.0150*	(-0.008–0.007)	0.0318	0.0155
	Saplings	-0.0104*	(-0.006–0.005)	0.0268	0.0107
	Seedlings	-0.0165*	(-0.012–0.011)	0.0319	0.0170
	Among classes	-0.0078*	(-0.003–0.002)	0.0154	0.0079
	Overall	-0.0086*	(-0.002–0.002)	0.0209	0.0087
Ralsko	Old trees	-0.0015NS	(-0.025–0.021)	0.0092	0.0015
	Mature trees	0.0059NS	(-0.013–0.011)	0.0570	-0.0062
	Saplings	-0.0098*	(-0.006–0.005)	0.0133	0.0100
	Seedlings	-0.0061NS	(-0.011–0.009)	0.0033	-0.0033
	Among classes	-0.0064*	(-0.003–0.003)	-0.0012	0.0064
	Overall	-0.0066*	(-0.002–0.002)	0.0024	0.0066
Sopřeč	Old trees	-0.0065NS	(-0.010–0.008)	0.0083	0.0066
	Mature trees	0.0030NS	(-0.012–0.010)	-0.0214	-0.0029
	Saplings	-0.0126*	(-0.005–0.004)	0.0197	0.0129
	Seedlings	-0.0076NS	(-0.013–0.011)	-0.0238	0.0075
	Among classes	-0.0034*	(-0.002–0.002)	0.0106	0.0034
	Overall	-0.0047*	(-0.002–0.002)	0.0104	0.0047

Slope of regression (*b*) between the kinship coefficient (*F*<sub>ij</sub>) and the log of the physical distance between individuals, 95% confidence interval (CI) for regression slope, the kinship between pairs of neighbours at the first distance class (*F*<sub>1</sub>) and the *Sp* statistics estimated for different age classes within individual localities of *Pinus strobus*. NS – not significant, \**P* < 0.05.

doi:10.1371/journal.pone.0068514.t006

other age classes. These results are concordant with studies performed on a large scale study conclusions presented above.

**Spatial Genetic Structure (SGS).** Our results concerning SGS corroborate those presented above. Old trees completely lacked any SGS. SGS increased over time and reached its maximum in the sapling stage, but on different distance classes in different populations with the lowest SGS in Bynovec locality (Figs. 4, 5, Table 6). In the seedling stage, however, we detected only a very weak SGS, probably because other factors such as increasing density of surrounding populations, thinning or seed shadow overlap started to gain importance (see below). SGS of *Pinus strobus* populations has been intensively studied in the native range [47,68,69,107]. In these studies, spatial autocorrelation analyses detected weak positive structuring at 10 to 15 m, which fits the isolation by distance model particularly for old growth population. Jones *et al.* [47] attribute weaker patterns observed in *P. strobus* to the longer dispersal distance of seeds and a historical overlap of seed shadows from adults outside of the plot coupled with an overlap of seed shadows from younger, more recently established reproductive adults. They also discuss the possible influence of disturbance and colonization history, mating system

and ecological factors on SGS and stress the role of thinning processes, which can weaken the initial structure present in seedlings within the parental population.

Troupin *et al.* [40] analysed the change in SGS of reproductive individuals over a span of 30 years in an expanding *Pinus halepensis* population founded by five reproductive individuals. They found no SGS in the early stages of invasion and suggested the random distribution of genotypes could be the result of density-dependent grazing. In our study on invasive *Pinus strobus*, we also did not detect a SGS in early stages, i.e. in old trees. These trees were planted by foresters, and individual gene combinations were introduced randomly. Old populations therefore cannot have any SGS, and the spatial tree distribution is not an outcome of a biological process but of a human activity. However, many studies of conifer species demonstrate a weak SGS in old-age classes in wind-dispersed species as an outcome of biological processes such as consanguineous mating, low level of pollen flow and short-distance seed dispersal without human influence [44,47,69,117,118].

Epperson [119] theoretically demonstrated that when the initial distribution of genotypes is random, the degree of spatial autocorrelation quickly increases. In the same way as in the case of *Pinus halepensis* [40], we found a gradual increase in SGS over time that culminated in the sapling stage, in which we detected SGS up to 25 m. Although the data from the native range document SGS at 10–15 m distances [47,68,69,107], the patchiness of invasive populations outside of the native range was wider but not stronger. Many studies demonstrate a strong SGS in the seedling stage compared to the adult stage for other species [49,120]. This, by contrast, is not the case in invasive populations in the adventive range of *Pinus strobus*. We have observed very weak spatial structuring in the seedling stage at all localities, which should be the result of two processes that are not mutually exclusive, i.e. long-distance seed dispersal and spatial-temporal overlap in seed shadows. Based on these processes, as the population of an invasive tree rapidly increases, propagule pressure increases as well. In combination with the high dispersal capacity of *Pinus strobus* (the species has winged seeds adapted for wind dispersal), seedlings do not form strongly genetically structured populations, as different genotypes are widely mixed due to pollen and seed dispersal over distances longer than 100 m. Hadincová *et al.* [61] and Münzbergová *et al.* [62] showed from different sandstone areas of the Czech Republic that *P. strobus* is able to spread very effectively. They estimated that *P. strobus* can disperse up to 750 m away from the parental source in different localities. As a result, there is great seed shadow overlap from adults within the population as well as adults adjacent to the plot (but which are not included in our analysis). This has the effect of reducing both the level and spatial scale of relatedness in studied populations and is especially pronounced in the seedling stage. On the other hand, 750 m is the maximal detected dispersal distance, a majority of seeds is dispersed up to 100 m apart and that approximately 80% of them are distributed up to 20 m apart. This dispersal pattern may generate fine-scale genetic clustering up to 15 m among age classes.

## Conclusions

In summary, the comparative population genetics of native and adventive *P. strobus* reveals a surprisingly complicated story. The amount of the genetic diversity in the adventive range is attributable to either multiple introductions, similar to findings on other species [23–27] or introductions from a single, genetically very diverse, source, or a combination of both. However, as we did



not find any population genetic structure in the native range, we suggest that population amalgamation probably first happened in the native range, prior to introduction, due to pollen mediated gene flow and/or human transport of propagules in the 18<sup>th</sup> century and later. In such a case, there was no need for multiple introductions from previously isolated populations but only several introductions from genetically diverse populations. Moreover, we also found evidence of differentiation between invasive and naturalized populations in the adventive range. Highly invasive populations were clearly distinguished by different allele frequencies. The fine-scale genetic diversity study further supports these results. The genetic diversity of trees introduced to the adventive localities approximately one hundred years ago was high enough to ensure the maintenance of high genetic diversity. At the fine scale within the invasive populations, we found no evidence for any increase in genetic diversity parameters over time. Furthermore, in invasive populations, we observed the development of fine-scale genetic clustering over time. This occurred at the maximum distance of 25 m, at which old trees completely lacked any spatial genetic structure that increased over time and reached its maximum in the sapling stage. These results support the hypothesis that rather than admixture, a single introduction from a genetically diverse source or multiple introductions from different sources of similar genetic diversity pattern, or a combination of

such events, was likely responsible for the high genetic diversity of *P. strobus* populations in the adventive range.

## Acknowledgments

The authors thank Lucie Lichtenbergová, Martina Hadincová and Petra Caklová for their excellent technical, field and lab assistance, Barbora Bimová and Michaela Blechová for their significant help with lab work and Josef Bryja for his significant help with microsatellite analysis setup. We also thank to the staff of the GIS laboratory of the Institute of Botany, namely to Jan Wild, for providing us with the DGPS technique and helping us with post processing corrections. We thank Kerry Woods for his help to identify old growth forests throughout USA, John Thompson for his assistance with sample collection at Mohonk Preserve, and Cary Institute of Ecosystem studies for logistic support. We thank Jonathan Rosenthal for his help with the fieldwork, insightful comments, and substantial revision of the manuscript. Fred Rooks improved our English. We would like to thank the Forests of the Czech Republic, State Enterprise, who kindly provided us with maps of white pine plantations.

## Author Contributions

Conceived and designed the experiments: BM VH. Performed the experiments: BM VH VM RW. Analyzed the data: BM VH. Wrote the paper: BM VM.

## References

1. Rejmánek M, Richardson DM (1996) What attributes make some plant species more invasive? *Ecology* 77: 1655–1661.
2. Reichard SH, Hamilton CW (1997) Predicting invasions of woody plants introduced into North America. *Conserv Biol* 11: 193–194.
3. Mandák B (2003) Germination requirements of invasive and non-invasive *Atriplex* species: a comparative study. *Flora* 198: 45–54.
4. Tilman D (1997) Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology* 78: 81–92.
5. Herben T, Mandák B, Bimová K, Münzbergová Z (2004) Invasibility and species richness of a community: A neutral model and a survey of published data. *Ecology* 85: 3223–3233.
6. Davis MA, Grime JP, Thompson K (2000) Fluctuating resources in plant communities: a general theory of invasibility. *J Ecol* 88: 528–534.
7. Hobbs RJ, Huenneke LF (1992) Disturbance, diversity, and invasion – implications for conservations. *Conserv Biol* 6: 324–337.
8. Levine JM (2000) Species diversity and biological invasions: relating local process to community pattern. *Science* 288: 852–854.
9. Simberloff D (2009) The role of propagule pressure in biological invasions. *Annu Rev Ecol Syst* 40: 81–102.
10. Crawley MJ (1978) What makes a community invulnerable? In: Gray AJ, Crawley MJ, Edwards PJ, editors. *Colonization, succession and stability*. Oxford: Blackwell Scientific. 429–453.
11. Richardson DM, Allsopp N, D'Antonio CM, Milton SJ, Rejmánek M (2000) Plant invasions – the role of mutualism. *Biol Rev Camb Philos Soc* 75: 65–93.
12. Lonsdale WM (1999) Global pattern of plant invasion and the concept of invasibility. *Ecology* 80: 1522–1536.
13. Tilman D (1999) The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 80: 1455–1474.
14. Williamson M (1996) *Biological Invasions*. London: Chapman & Hall. 244 p.
15. van Kleunen M, Dawson W, Schlaepfer D, Jeschke JM, Fischer M (2010) Are invaders different? A conceptual framework of comparative approaches for assessing determinants of invasiveness. *Ecol Lett* 13: 947–958.
16. Mandák B, Pyšek P, Lysák M, Suda J, Krahulcová A, et al. (2003) Variation in DNA-ploidy levels of *Reynoutria* taxa in the Czech Republic. *Ann Bot* 92: 265–272.
17. Mandák B, Bimová K, Pyšek P, Štěpánek J, Plačková I (2005) Isoenzyme diversity in *Reynoutria* taxa: escape from sterility by hybridization. *Plant Syst Evol* 253: 219–230.
18. Zhang YY, Zhang DY, Barrett SCH (2010) Genetic uniformity characterizes the invasive spread of water hyacinth (*Eichhornia crassipes*), a clonal aquatic plant. *Mol Ecol* 19: 1774–1786.
19. Lambrinos JG (2004) How interactions between ecology and evolution influence contemporary invasion dynamics. *Ecology* 85: 2061–2070.
20. Dlugosh KM, Parker IM (2008) Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Heredity* 11: 701–709.
21. Zenger KR, Richardson BJ, Vachot-Griffin AM (2003) A rapid population expansion retains genetic diversity within European rabbits in Australia. *Mol Ecol* 12: 789–794.
22. Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus of the evolution of invasiveness in plants? *Proc Natl Acad Sci USA* 97: 7043–7050.
23. Gaskin JF, Schaal BA (2002) Hybrid *Tamarix* widespread in U.S. invasion and undetected in native Asian range. *Proc Natl Acad Sci USA* 99: 11256–11259.
24. Kolbe JJ, Glor RE, Rodríguez-Schettino L, Chamizo-Lara A, Larson A, et al. (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* 431: 177–181.
25. Genton BJ, Shykoff JA, Giraud T (2005) High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Mol Ecol* 14: 4275–4285.
26. Marrs RA, Sforza R, Huffbauer RA (2008) Evidence for multiple introductions of *Centauria stoebe micranthos* (spotted knapweed, Asteraceae) to North America. *Mol Ecol* 17: 4197–4208.
27. Rosenthal DM, Ramakrishnan AP, Cruzan MB (2008) Evidence for multiple sources of invasion and intraspecific hybridization in *Brachypodium sylvaticum* (Hudson) Beauv. in North America. *Mol Ecol* 17: 4657–4669.
28. Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion. *J Evol Biol* 23: 1720–1731.
29. Mandák B, Pyšek P (1998) History of spread and habitat preferences of *Atriplex sagittata* (Chenopodiaceae) in the Czech Republic. In: Starfinger U, Edwards K, Kowarik I, Williamson M, editors. *Plant invasions: ecological mechanisms and human responses*. Leiden: Backhuys Publisher. 209–224.
30. Crooks JA, Soulé ME (1999) Lag times in population explosions of invasive species: causes and implications. In: Sandlund OT, Schei PJ, Viken A, editors. *Invasive species and biodiversity management*. Dordrecht: Kluwer Academic Publisher. 103–125.
31. Mandák B (2003) Distribution of four *Atriplex* species with different degrees of invasiveness in the Czech Republic. In: Child LE, Brock JH, Brundu G, Prach K, Pyšek P, Wade PM, Williamson M, editors. *Plant Invasions: Ecological Threats and Management Solutions*. Leiden: Backhuys Publisher. 313–328.
32. Mandák B, Pyšek P, Bimová K (2004) History of the invasion and distribution of *Reynoutria* taxa in the Czech Republic: a hybrid spreading faster than its parents. *Preslia* 76: 15–64.
33. Kowarik I (1995) Time lags in biological invasions with regard to the success and failure of alien species. In: Pyšek P, Prach K, Rejmánek M, Wade PM, editors. *Plant invasions: general aspects and special problems*. Amsterdam: SPB Academic Publishing. 15–38.
34. Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, et al. (2007) Multiple sources, admixture, and genetic variation in introduced Anolis lizard populations. *Conserv Biol* 21: 1612–1625.
35. Kolbe JJ, Larson A, Losos JB, de Queiroz K (2008) Admixture determines genetic diversity and population differentiation in the biological invasion of a lizard species. *Biol Lett* 4: 434–437.
36. Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proc Natl Acad Sci USA* 104: 3883–3888.
37. Chun YJ, Fumanal B, Laitung B, Bretagnolle F (2010) Gene flow and population admixture as the primary post-invasion processes in common

- ragweed (*Ambrosia artemisiifolia*) populations in France. *New Phytol* 185: 1100–1107.
38. Zayed A, Whitfield CW (2008) A genome wide signature of positive selection in ancient and recent invasive expansions of the honey bee *Apis mellifera*. *Proc Natl Acad Sci USA* 105: 3421–3426.
  39. 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: 1309–1321.
  40. Troupin D, Nathan R, Vendramin GG (2006) Analysis of spatial genetic structure in an expanding *Pinus halepensis* populations reveals development of fine-scale genetic clustering over time. *Mol Ecol* 15: 3617–3630.
  41. Berg EE, Hamrick JL (1995) Fine-scale genetic structure of a turkey oak forest. *Evolution* 49: 110–120.
  42. Vekemans X, Hardy OJ (2004) New insight from fine-scale spatial genetic structure analysis in plant populations. *Mol Ecol* 13: 921–935.
  43. Young AG, Merriam HG (1994) Effects of forest fragmentation on the spatial genetic structure of *Acer saccharum* Marsh. (sugar maple) populations. *Heredity* 72: 201–208.
  44. Parker KC, Hamrick JL, Parker AJ, Nason JD (2001) Fine-scale genetic structure in *Pinus clausa* (Pinaceae) populations: Effects of disturbance history. *Heredity* 87: 99–113.
  45. Premoli AC, Kitzberger T (2005) Regeneration mode affects spatial genetic structure of *Nothofagus dombeyi* forests. *Mol Ecol* 14: 2319–2329.
  46. Pardini EA, Hamrick JL (2008) Inferring recruitment history from spatial genetic structure within populations of the colonizing tree *Albizia julibrissin* (Fabaceae). *Mol Ecol* 17: 2865–2879.
  47. Jones FA, Hamrick JL, Peterson CJ, Squiers ER (2006) Inferring colonization history from analyses of spatial genetic structure within populations of *Pinus strobus* and *Quercus rubra*. *Mol Ecol* 15: 851–861.
  48. Mandák B, Bimová K, Mahelka V, Plačková I (2006) How much genetic variation is stored in the seed bank? A study of *Atriplex tatarica* (Chenopodiaceae). *Mol Ecol* 15: 2653–2663.
  49. Chung MY, Nason JD, Epperson BK, Chung MG (2003) Temporal aspects of the fine-scale genetic structure in a population of *Cinnamomum insularimontanum* (Lauraceae). *Heredity* 90: 98–106.
  50. Epperson B (1992) Spatial structure of genetic variation within populations of forest trees. *New Forests* 6: 257–278.
  51. Augspurger CK (1983) Seed dispersal of the tropical tree, *Platypodium elegans*, and the escape of its seedlings from fungal pathogens. *J Ecol* 71: 759–771.
  52. Augspurger CK, Kelly CK (1984) Pathogen mortality of tropical tree seedlings - experimental studies of the effects of dispersal distance, seedling density, and light conditions. *Oecologia* 61: 211–217.
  53. Hadincová V, Köhleinová I, Marešová J (2007) Invasive behaviour of white pine (*Pinus strobus*) in sandstone areas in the Czech Republic. In: Härtel H, Čilek V, Herben T, Jackson A, Williams R, editors. Sandstones Landscapes. Praha: Academia. 219–224.
  54. Nožička J (1965) Zavádění vejmutovky v Českých zemích do r. 1938 (Introduction of the white pine in the Czech countries up to year 1938). *Práce výzkumného ústavu lesnického ČSSR* 31: 41–67.
  55. Pyšek P, Sádlo J, Mandák B (2002) Catalogue of alien plants of the Czech Republic. *Preslia* 74: 97–186.
  56. Richardson DM, Pyšek P, Rejmánek M, Barbour MG, Panetta FD, et al. (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity & Distributions* 6: 93–107.
  57. Wendel GW, Smith HC (1990) *Pinus strobus* L. Eastern white pine. In: Burns RM, Honkala BH, editors. *Silvics of North America*. Vol. 1. Conifers. U.S. Department of Agriculture, Forest Service. 476–488.
  58. Foster DR, O'Keefe JF (2000) *New England Forests through time: insights from the Harvard Forests Dioramas*. Cumbreland, USA: Harvard University Press. 67 p.
  59. Musil I (1971) K otázce původu evropských vejmutovkových porostů (On the question of the origin of European White Pine growths). *Acta Musei Silesiae* 129–134.
  60. Münzbergová Z, Hadincová V, Wild J, Kindlmannová J (2013) Variability in the contribution of different life stages to population growth as a key factor in the invasion success of *Pinus strobus*. *PLoS ONE* 8: e56953.
  61. Hadincová V, Münzbergová Z, Wild J, Šajtar L, Marešová J (2008) Dispersal of invasive *Pinus strobus* in sandstone areas of the Czech Republic. In: Tokarska-Guzik B, Brock JH, Brundu G, Child LE, Daehler CC, Pyšek P, editors. *Plant invasions: Human perception, ecological impacts and management*. Leiden: Backhuys Publisher. 117–132.
  62. Münzbergová Z, Hadincová V, Wild J, Herben T, Marešová J (2010) Spatial and temporal variation in dispersal pattern of an invasive pine. *Biol Invasions* 12: 2471–2486.
  63. Robledo-Arnuncio JJ (2011) Wind pollination over mesoscale distances: an investigation with Scotch pine. *New Phytol* 190: 222–233.
  64. Williams CG (2010) Long distance pollen still germinates after meso-scale dispersal. *Am J Bot* 97: 846–855.
  65. Dovčiak M, Frelich LE, Reich E (2005) Pathway in old-field succession to white pine: seed rain, shade, and climate effects. *Ecol Monogr* 75: 363–378.
  66. Musil I (1971) Provenienční výzkum vejmutovky (*Pinus strobus* L.) na Ostravsku (Local investigation of White Pine (*Pinus strobus*) in Ostrava region). *Acta Musei Silesiae* 97–128.
  67. Davis MB (1993) *Old growth in the East: a survey*. Richmond, VT: Wild Earth.
  68. Epperson BK, Chung MG (2001) Spatial genetic structure of allozyme polymorphisms within populations of *Pinus strobus* (Pinaceae). *Am J Bot* 88: 1006–1010.
  69. Marquardt PE, Epperson BK (2004) Spatial and population genetic structure of microsatellites in white pine. *Mol Ecol* 13: 3305–3315.
  70. Tolasz R (1997) *Climate atlas of Czechia*. Prague: Czech Hydrometeorological Institute.
  71. Štorchová H, Hrdličková R, Chrtěk Jr J, Tetera M, Fitze D, et al. (2000) An improved method for DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. *Taxon* 49: 79–84.
  72. Echt CS, May-Marquardt P (1997) Survey of microsatellite DNA in pine. *Genome* 40: 9–17.
  73. Echt CS, May-Marquardt P, Hseih M, Zahorchak R (1996) Characterization of microsatellite markers in eastern white pine. *Genome* 39: 1102–1108.
  74. van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4: 535–538.
  75. El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theor Appl Genet* 92: 832–839.
  76. Goudet J (1995) Fstat version 1.2: a computer program to calculate F-statistics. *J Hered* 86: 485–486.
  77. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
  78. Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online* 1: 47–50.
  79. Rice WR (1989) Analyzing tables of statistical tests. *Evolution* 43: 223–225.
  80. Wright S (1965) The interpretation of population structure by F-statistics with special regard to system mating. *Evolution* 19: 395–420.
  81. Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457–462.
  82. Gaggiotti OE, Lange O, Rassmann K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Mol Ecol* 8: 1513–1520.
  83. Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.
  84. Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution* 49: 1280–1283.
  85. Sokal RR, Rohlf FJ (1995) *Biometry: the principles and practise of statistics in biological research*. 3rd edn. New York: WH Freeman and Company.
  86. 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.
  87. Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered* 90: 503.
  88. Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J Hered* 89: 238–247.
  89. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
  90. Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 9: 1322–1332.
  91. 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.
  92. Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol Ecol* 15: 1419–1439.
  93. Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.
  94. Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* 4: 137–138.
  95. 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.
  96. Kalisz S, Nason JD, Hanzawa FM, Tonsor SJ (2001) Spatial population genetic structure in *Trillium grandiflorum*: the roles of dispersal, mating, history, and selection. *Evolution* 55: 1560–1568.
  97. Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* 142: 1357–1362.
  98. Cockerham CC (1969) Variance of gene frequencies. *Evolution* 23: 72–84.
  99. Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* 2: 618–620.
  100. Beaulieu J, Simon JP (1994) Genetic structure and variability in *Pinus strobus* in Quebec. *Can J For Res* 24: 1726–1733.
  101. Buchert GP, Rajora OP, Hood JV, Dancik BP (1997) Effects of harvesting on genetic diversity in old-growth Eastern White Pine in Ontario, Canada. *Conserv Biol* 11: 747–758.

102. Rajora OP, DeVerno L, Mosseler A, Innes DJ (1998) Genetic diversity and population structure of disjunct Newfoundland and central Ontario populations of eastern white pine (*Pinus strobus*). *Can J Bot* 76: 500–508.
103. Rajora OP, Rahman MH, Buchert GP, Dancik BP (2000) Microsatellite DNA analysis of genetic effects of harvesting in oldgrowth eastern white pine (*Pinus strobus*) in Ontario, Canada. *Mol Ecol* 9: 339–348.
104. Mehes M, Nkongolo KK, Michael P (2007) Genetic analysis of *Pinus strobus* and *Pinus monticola* populations from Canada using ISSR and RAPD markers: development of genome-specific SCAR markers. *Plant Syst Evol* 267: 47–63.
105. Mehes M, Nkongolo KK, Michael P (2009) Assessing genetic diversity and structure of fragmented populations of eastern white pine (*Pinus strobus*) and western white pine (*P. monticola*) for conservation management. *J Plant Ecol* 2: 143–151.
106. Myers ER, Chung MY, Chung MG (2007) Genetic diversity and spatial genetic structure of *Pinus strobus* (Pinaceae) across an island landscape inferred from allozyme and cpDNA markers. *Plant Syst Evol* 13: 15–30.
107. Walter R, Epperson BK (2004) Microsatellite analysis of spatial structure among seedlings in populations of *Pinus strobus* (Pinaceae). *Am J Bot* 91: 549–557.
108. Taylor DR, Keller SR (2007) Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. *Evolution* 61: 334–345.
109. Latta RG, Linhart YB, Fleck D, Elliot M (1998) Direct and indirect estimates of seed versus pollen movement within a population of ponderosa pine. *Evolution* 52: 61–67.
110. Latta RG, Mitton JB (1999) Historical separation and present gene flow through a zone of secondary contact in ponderosa pine. *Evolution* 53: 61–67.
111. Jorgensen S, Hamrick JL, Wells PV (2002) Regional patterns of genetic diversity in *Pinus flexilis* (Pinaceae) reveal complex species history. *Am J Bot* 89: 792–800.
112. Richardson BA, Brunsfeld J, Klopfenstein NB (2002) DNA from bird-dispersed seed and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Mol Ecol* 11: 215–227.
113. Parchman TL, Benkman CW, Jenkins B, Buerkle CA (2011) Low levels of population genetic structure in *Pinus contorta* (Pinaceae) across a geographic mosaic of co-evolution. *Am J Bot* 98: 669–679.
114. Prus-Glowacki W, Urbaniak L, Bujas E, Curtu AL (2012) Genetic variation of isolated and peripheral populations of *Pinus sylvestris* (L.) from glacial refugia. *Flora* 207: 150–158.
115. Young RA, Giese RL (2003) *Introduction to Forest Ecosystem, Science and Management*. Madison: John Wiley & Sons.
116. Foster D (1992) Land-use history (1730–1990) and vegetation dynamics in central New England. *J Ecol* 80: 753–772.
117. Epperson BK, Allard RW (1989) Spatial autocorrelation analysis of the distribution of genotypes within populations lodgepole pine. *Genetics* 121: 369–377.
118. Gapare WJ, Aitken SN (2005) Strong spatial genetic structure in peripheral but not core populations of Sitka spruce *Picea sitchensis* (Bong.) Carr. *Mol Ecol* 14: 2659–2667.
119. Epperson BK (2005) Estimating dispersal from short distance spatial autocorrelation. *Heredity* 95: 7–15.
120. Hamrick JL, Murawski DA, Nason JD (1993) The influence of seeds dispersal mechanisms on the genetic structure of tropical tree populations. *Vegetatio* 103: 281–297.