**Genetic Structure of Norway Spruce Ecotypes Studied by SSR Markers**

Zuzana Bínová, Jiří Korecký *, Jakub Dvořák, Jan Bílý, Dagmar Zádrapová, Václav Jansa and Milan Lstiburek

Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 165 21 Prague, Czech Republic; binova@fld.czu.cz (Z.B.); dvorakjakub@fld.czu.cz (J.D.); bily@fld.czu.cz (J.B.); zadrapova@fld.czu.cz (D.Z.); jansav@fld.czu.cz (V.J.); lstiburek@fld.czu.cz (M.L.)

* Correspondence: korecky@fld.czu.cz

Received: 6 December 2019; Accepted: 13 January 2020; Published: 16 January 2020

**Abstract:** Norway spruce is a widespread and economically highly important tree species in Central Europe which occurs there in different morphotypic forms (also known as ecotypes). Previously established common garden experiments indicated that the morphological differentiation is most likely genetically determined. The genetic structure of Norway spruce morphological variants might be an indicator (marker) of specific sustainability in forest ecosystems. In this study, we investigated 436 individuals from autochthonous populations belonging to three different ecotypes. The main aim was to evaluate a level of genetic intra and interpopulation diversity among the low, medium and high-elevation ecotypes using both expressed sequence tag simple sequence repeats (EST – SSR) and genomic SSR markers. Sixteen highly polymorphic microsatellite loci folded in two newly designed multiplexes were used to depict the genetic structure of targeted trees. Important allele frequency parameters, such as the mean expected (0.722, SE = 0.061) and observed (0.585, SE = 0.062) heterozygosity and mean effective number of alleles (N_e = 5.943, SE = 1.279), were estimated. The low genetic differentiation detected among different ecotypes (F_st = 0.008) was further discussed and clarified.

**Keywords:** Picea abies (L.) Karst; microsatellite genotyping; genetic variation; three forms

---

**1. Introduction**

Despite the ongoing decline, Norway spruce *Picea abies* (L.) Karst. remains the most widespread tree species in central Europe, which demonstrates its economic as well as ecological importance [1–3]. In the last decades, the European spruce bark beetle *Ips typographus* (L.) has caused an extensive population decline of European Norway spruce forests. This unfortunate fact is caused by many factors but has been accelerated mainly by climate change [4]. However, Norway spruce is still a preferred species for reforestation for its easy cultivation and growth performance, particularly for completely deforested sites.

The observed morphological variability of Norway spruce is attributable to its phenotypic adaptation to diverse climatic conditions. Locally adapted populations are more resilient to abiotic factors such as snow load or wind exposure [5–9]. Besides the morphological characteristics, Norway spruce ecotypes are distinguishable by their geographic origin, climatic conditions, and type of vegetation in the area [6,8,10–12]. In the Czech Republic, all three ecotypic forms are distributed predominantly along the vertical geographic gradient. The low-elevation ecotype, *acuminata*, is characterized by its wide crowns, toothed cone scales, and comb-like branches (Figure S1). The high-elevation (upper tree line) ecotype, *obovata*, typically has a narrow crown, round cone scales, and flat branches (Figure S2). The medium-elevated ecotype, *europaea*, has rhombic toothed cone
Forests 2020, 11, 110 2 of 15

scales. The crown is mostly conical, with intermediate width and brush-type branches that are usually shorter, denser and hang down (Figure S3) [13]. Since the beginning of the 20th century, the scientific interest in the heritable variability of trees’ crown shape has grown [13]. Swedish cross-pollination experiments carried out in the 1940s proved the heritable variability of crown phenotypic characteristics [14]. Furthermore, several provenance trials were established in Germany in the 1972 to improve the knowledge of the variation in traits, such as the resistance to damages caused by heavy snow cover among various provenances [15]. Although most damages occurred in clumps indicating occurrence unrelated to provenance, it was shown that the elevation of the seed origin has a significant effect on growth and yield. Historically, Gruber [16], Sylvestre [13] (Germany), and Samek [6] (Czech Republic) independently anticipated the phenotypic resilience and potential heritability of crown shape. Geburek et al. [17] considered the possibility of different crown types being associated with a temperature–elevation gradient. While in low-elevated ecotypes and medium-elevated ecotypes the selective pressure is comparably low and the gene flow very intense, the adaptive differentiation seems to be weaker; in high-elevated ecotypes, gene flow is more restricted, and environmental conditions cause high selective pressure. Principles of environmental heterogeneity in space and over time generally require a high adaptation potential in long-living organisms and different strategies in various climate conditions [18]. The pure ecotopic forms are most likely genetically encoded, but certain modifications may be caused by environmental signals and thus increasing fitness leads to an adaptation. Therefore, knowledge of the magnitude and extent of a species’ genetic diversity is key to understanding the justified selection and breeding activities [19].

Microsatellites or simple sequence repeats (SSRs) are highly informative genetic markers of a codominant Mendelian inheritance. Due to their advantages, such as their high level of detected polymorphism, reproducibility, and transferability across related species, as well as their occurrence predominantly in non-coding DNA sequences, microsatellites have been extensively used for the analysis of populations’ genetic diversity [11,20–26], gene flow [27], parentship analysis [28], or for the construction of linkage maps [29–31] in Norway spruce. In these studies, both genomic (gSSRs) and expressed sequence tag SSR (EST-SSR) markers are frequently used without any significant difference among population-genetic parameters being obtained, which is due to a theoretically explainable basis [32,33]. To our knowledge, no studies have focused on the genetic differentiation among distinctive tree forms (ecotypes). Caré et al. [34] focused on a high-elevation ecotype; however, they compared this morphotype to neighboring allochthonous stands without any effort to reveal the genetic basis of other ecotypic groups. Conifers are typically outbreeding, which is characterized by high gene flow and facilitated by long-distance pollen and seed dispersal. This results in a sizeable effective population size and usually low inter-population genetic differentiation [35]. Additionally, conifers are known for their highly heterozygous genomes, although their nucleotide substitution rates are lower than those of most angiosperms, perhaps because of their long lifespan, lasting from decades to centuries [36–39]. In general, Norway spruce exhibits high levels of genetic variability within populations and low levels of differentiation among populations [23,24,40–45].

In our study, we assessed the genetic differentiation among three distinctive ecotypic groups of Norway spruce sampled on autochthonous stands in the Czech Republic. We used two newly assembled multiplexes of gSSR and EST-SSR microsatellite markers for molecular data mining. More specifically, we addressed the following questions: (i) What is the level of genetic diversity and differentiation within and among Norway spruce populations? (ii) Is there a distinctive genetic structure among the studied ecotypes based on microsatellite genotypes?
2. Materials and Methods

2.1. Study Site and Sampling

Plant material: Initially, 450 individuals were sampled across three geographic regions (150 individuals per area). Due to genotyping failure, the number of analyzed trees was slightly reduced (Table 1, column N). Each region represents a particular ecotypic form within a given elevation range (Table 1, column E), which is characterized by average annual temperature (Table 1, column T) and average annual precipitation (Table 1, column P). Each geographic area consists of several smaller sub-areas; within these, targeted trees were chosen in transects or clusters which were at least 25 m apart from each other (Figures S4–S7). All forest stands are putatively autochthonous and naturally regenerated without known artificial reforestation. The high-elevation ecotype (HE) of Norway spruce originated in the Krkonoše (Giant Mountains) national park (50°44′57.3″ N; 15°25′09.6″ E) at an elevation of over 1100 m above sea level (a.s.l.). The medium-elevation ecotype (ME) was sampled in the Jizera Mountains natural reservation area (50°48′25.9″ N; 15°13′05.9″ E) at an elevation range from 650 to 850 m a.s.l, whereas the low-elevation form (LE) came from the national park České Švýcarsko (Bohemian Switzerland, 50°51′56.6″ N; 14°25′23.9″ E), which is under 400 m a.s.l.

Table 1. Climatological and stand-specific data.

<table>
<thead>
<tr>
<th>Ecotypic Form</th>
<th>Elevation</th>
<th>N</th>
<th>E, m a.s.l.</th>
<th>T, °C</th>
<th>P, mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>acuminata</td>
<td>Low</td>
<td>149</td>
<td>190–400</td>
<td>8.2</td>
<td>752</td>
</tr>
<tr>
<td>europaea</td>
<td>Medium</td>
<td>148</td>
<td>650–850</td>
<td>5.9</td>
<td>821</td>
</tr>
<tr>
<td>obovata</td>
<td>High</td>
<td>141</td>
<td>1100–1250</td>
<td>4.9</td>
<td>946</td>
</tr>
</tbody>
</table>

The elevation range of sampled individuals (E, meters above sea level), average annual air temperature (T), average annual precipitation (P), and the number of individuals genotyped (N). Climatic values were long-term average measurements [46]. Elevation data were derived from GPS coordinates.

2.2. DNA Extraction and SSR Genotyping

DNA extraction: From each sample, approximately 50–80 mg of unsterilized fresh needle tissue was cut into small pieces, frozen in liquid nitrogen and ground in a mixer mill MM400 (Retsch, Haan, Germany) for 3 min at 30 Hz. The genomic DNA was extracted by the DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions with the minor modification in the lysis step (60 min of heating phase). The DNA quality and quantity were evaluated with a NanoDrop® 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and DNA was subsequently diluted to a working concentration of 20 ng/µL.

Microsatellite analysis: We initially tested 68 pre-selected dinucleotide and trinucleotide microsatellite markers from various sources [22,26,29,47–54] in singleplex polymerase chain reactions (PCRs). As a result of this laborious activity, we set up two multiplexes containing altogether 16 polymorphic and highly reproducible, fluorescently labeled primers (Table 2). The final primer sequences selection was also accounted for PCR amplification compatibility among primers. An issue of various annealing temperatures among selected ones was considerably diminished when using Type-it® Microsatellite PCR kit (Qiagen, Hilden, Germany) for PCR amplification.
**Table 2.** Features of microsatellite primers assembled into multiplexes. SSR: simple sequence repeats.

<table>
<thead>
<tr>
<th>SSR Locus</th>
<th>EMBL Accession Number</th>
<th>Primer Sequences</th>
<th>Repeats</th>
<th>Allele Size Range</th>
<th>Fluorescent Dye</th>
<th>Source Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multiplex 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>paGB3</td>
<td>AJ133748</td>
<td>AGTGATTAACCTCTGACCAC</td>
<td>(AT)\textsubscript{11}</td>
<td>112–130</td>
<td>PET</td>
<td>[52]</td>
</tr>
<tr>
<td>WS0022.B15 *</td>
<td>CA480899</td>
<td>TTTTGAGTTCGCAAGAATG</td>
<td>(AG)\textsubscript{12}</td>
<td>174–214</td>
<td>NED</td>
<td>[47]</td>
</tr>
<tr>
<td>SpAGD1</td>
<td>G31824</td>
<td>GTCAAAAACCTTGAAAGCCA</td>
<td>(AO)\textsubscript{15}</td>
<td>122–182</td>
<td>NED</td>
<td>[48]</td>
</tr>
<tr>
<td>SpAGC1</td>
<td>G31822</td>
<td>TTCACCTTAGGCCGAAGACC</td>
<td>(TC\textsubscript{3}TT</td>
<td>TC\textsubscript{10}</td>
<td>73–123</td>
<td>VIC</td>
</tr>
<tr>
<td>SpAG2</td>
<td>G31821</td>
<td>GTCTCTTGATCTACCTGACCA</td>
<td>(TC)\textsubscript{7}</td>
<td>87–115</td>
<td>NED</td>
<td>[48]</td>
</tr>
<tr>
<td>EATC1E03</td>
<td>AJ296702</td>
<td>CACCTATTCTCAAAGGCTAAA</td>
<td>(CAT)\textsubscript{3}</td>
<td>121–143</td>
<td>6-FAM</td>
<td>[51]</td>
</tr>
<tr>
<td>WS00716.F13 *</td>
<td>CA480905</td>
<td>TCAAGTAAATGGACAAAGGACAAGATACA</td>
<td>(GA)\textsubscript{10}</td>
<td>206–252</td>
<td>6-FAM</td>
<td>[47]</td>
</tr>
<tr>
<td><strong>Multiplex 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WS0073.A08 *</td>
<td>CA480903</td>
<td>TGCTCTCTATTTCGCCGCTTC</td>
<td>(AT)\textsubscript{14}</td>
<td>199–217</td>
<td>VIC</td>
<td>[47]</td>
</tr>
<tr>
<td>PA_28</td>
<td>GT887962</td>
<td>GGCCGAAAGTGGCTACTGCTA</td>
<td>(TC)\textsubscript{18}</td>
<td>151–172</td>
<td>VIC</td>
<td>[54]</td>
</tr>
<tr>
<td>PAAAC3</td>
<td>AJ131104</td>
<td>CGTACTGCAATTTTCTCCA</td>
<td>(CA)\textsubscript{16}</td>
<td>240–320</td>
<td>VIC</td>
<td>[49]</td>
</tr>
<tr>
<td>WS0092.A19 *</td>
<td>CA480888</td>
<td>TTGGTTTGTGCTGCTGAAAGAA</td>
<td>(AC)\textsubscript{3}</td>
<td>213–237</td>
<td>NED</td>
<td>[47]</td>
</tr>
<tr>
<td>WS0023.B03 *</td>
<td>CA480900</td>
<td>AGAGCTGGGGTGACAAGATT</td>
<td>(AT)\textsubscript{10}</td>
<td>163–211</td>
<td>NED</td>
<td>[47]</td>
</tr>
<tr>
<td>PA_33</td>
<td>GT884592</td>
<td>GTTCTGGAGGAGGAGGAGTACAG</td>
<td>(CGG)\textsubscript{8}</td>
<td>99–108</td>
<td>6-FAM</td>
<td>[54]</td>
</tr>
<tr>
<td>WS0019.F12 *</td>
<td>CA480896</td>
<td>AACACCGTCATCTCTGCTCGG</td>
<td>(AT)\textsubscript{13}</td>
<td>355–387</td>
<td>6-FAM</td>
<td>[47]</td>
</tr>
<tr>
<td>Pa_56</td>
<td>GT884695</td>
<td>ATCGTCTGATGCTGATCCAC</td>
<td>(AGGTTG)\textsubscript{3}</td>
<td>124–139</td>
<td>PET</td>
<td>[54]</td>
</tr>
<tr>
<td>PAAAC23</td>
<td>AJ131109</td>
<td>TTGGGGCGCCTGAATATACAG</td>
<td>(GT)\textsubscript{14}</td>
<td>181–315</td>
<td>PET</td>
<td>[49]</td>
</tr>
</tbody>
</table>

* EST-SSRs, otherwise genomic SSRs, European Molecular Biology Laboratory (EMBL), all forward primers were fluorescently labelled (PET, NED, 6-FAM, VIC) at the 5’ primer end.
SSRs amplification was performed in two multiplex polymerase chain reactions (Mastercycler® nexus, Eppendorf, Germany) in 12.5 µL of total reaction volume. Each reaction contained 1 µL of DNA (c = 20 ng/µL) and an appropriate mixture of SSR primers for each multiplex (Table 2) in a concentration of 0.12 µM for each forward and reverse primer and 6.25 µL of Type-it® solution (Qiagen, Hilden, Germany). Multiplex 1 (paGB3, WS0022.B15, WS00716.F13, SpAGD1, SpAGC1, SpAG2, EATC1E03) was amplified according to the following protocol: an initial incubation (15 min at 95 °C), 40 cycles (denaturation for 30 s at 95 °C followed by annealing for 45 s at 60 °C and extension phase for 30 s at 72 °C), terminated by a final extension (30 min at 60 °C). Reaction conditions of multiplex 2 (WS0073.H8, WS0092.A19, WS0023.B02, WS0019.F22, WS0015.I04, PAAC23, PAAC3, Pa_28, Pa_33) were as follows: 35 cycles (denaturation for 45 s at 95 °C followed by annealing for 60 s at 59 °C and extension for 45 s at 72 °C), and final extension (30 min at 72 °C). The fluorescently labelled PCR products, along with a size standard (GMCGT500 LIZ), were processed on a 3500 Series Genetic Analyzer® (Applied Biosystems, Foster City, CA, USA). Allele identification and genotyping were performed using the GeneMarker® Fragment Analysis software (SoftGenetics, State College, PA, USA). Final allele binning was checked manually to minimize genotyping errors.

2.3. Data Analysis

Allele parameters such as the number of alleles per locus (k), expected (H_e) and observed (H_o) heterozygosity, and the polymorphic information content (PIC) were estimated using CERVUS v. 3.0 [55–57]. Deviation from the Hardy–Weinberg equilibrium, the significance of genetic differentiation (F_is inbreeding coefficient and pairwise F_st values), Mantel test (9999 permutations), the mean number of alleles per locus (N_a), the effective number of alleles (N_e) and an analysis of molecular variance (AMOVA) were performed using the software GenAlEx v. 6.5 [58]. The significance of null alleles was evaluated using Microchecker [59]. Two methods of genetic population structure analysis were performed, which model-based Bayesian clustering using software STRUCTURE v. 2.3.4 [60] (admixture model, correlated allele frequencies, 50,000 burn-in and 100,000 repeats after burn-in). We evaluated the true number of clusters K following the methods used by Evanno et al. [61] (iterations for K = 1–10) implemented in the web-based tool STRUCTURE HARVESTER [62]. An alternative method of genetic clustering—Discriminant Analysis of Principal Components (DAPC, R package “adegenet”) [63–65]—was also applied. For DAPC analysis, an optimal number of eigenvalues taken for analysis was determined (PCA eigenvalues = 8, DA eigenvalues = 10).

3. Results

3.1. Allele Frequencies and Statistic Parameters of Individual Loci

All examined samples (n = 436) were genotyped on 16 loci, possessing a low level of allelic dropout (8.39 × 10⁻³). The number of detected alleles per locus varied substantially between 4 and 36 (mean 17.44 ± 10.54), totaling 279 allelic variants across all loci (Table 3). Out of these, 180 alleles were rare (i.e., with a frequency rate lower than 0.05). The highest counts of identified alleles per locus (36) were found at PAAC3 and PAAC23. The total mean number of alleles was estimated to be 17.44 (standard deviation (SD) =10.541, standard error (SE) = 2.974).
Table 3. Locus specific parameters of genetic diversity (CERVUS).

<table>
<thead>
<tr>
<th>Locus</th>
<th>k</th>
<th>H_o</th>
<th>H_e</th>
<th>PIC</th>
<th>HW</th>
<th>F(Null)</th>
</tr>
</thead>
<tbody>
<tr>
<td>paGB3</td>
<td>8</td>
<td>0.624</td>
<td>0.779</td>
<td>0.747</td>
<td>***</td>
<td>0.1105</td>
</tr>
<tr>
<td>WS0022.B15 †</td>
<td>20</td>
<td>0.843</td>
<td>0.869</td>
<td>0.858</td>
<td>NS</td>
<td>0.0139</td>
</tr>
<tr>
<td>SpAGD1</td>
<td>30</td>
<td>0.742</td>
<td>0.953</td>
<td>0.950</td>
<td>***</td>
<td>0.1269</td>
</tr>
<tr>
<td>SpAGC1</td>
<td>25</td>
<td>0.580</td>
<td>0.610</td>
<td>0.600</td>
<td>NS</td>
<td>0.0345</td>
</tr>
<tr>
<td>SpAG2</td>
<td>17</td>
<td>0.857</td>
<td>0.892</td>
<td>0.881</td>
<td>NS</td>
<td>0.0182</td>
</tr>
<tr>
<td>EATC1E03</td>
<td>8</td>
<td>0.276</td>
<td>0.38</td>
<td>0.351</td>
<td>***</td>
<td>0.1599</td>
</tr>
<tr>
<td>WS00716.F13 †</td>
<td>23</td>
<td>0.874</td>
<td>0.934</td>
<td>0.929</td>
<td>ND</td>
<td>0.0335</td>
</tr>
<tr>
<td>WS0073.H08 †</td>
<td>7</td>
<td>0.660</td>
<td>0.670</td>
<td>0.608</td>
<td>NS</td>
<td>0.0068</td>
</tr>
<tr>
<td>PA_28 †</td>
<td>7</td>
<td>0.766</td>
<td>0.687</td>
<td>0.637</td>
<td>***</td>
<td>−0.0663</td>
</tr>
<tr>
<td>PAAC3</td>
<td>36</td>
<td>0.457</td>
<td>0.879</td>
<td>0.868</td>
<td>***</td>
<td>0.3145</td>
</tr>
<tr>
<td>WS0092.A19 †</td>
<td>13</td>
<td>0.722</td>
<td>0.794</td>
<td>0.767</td>
<td>*</td>
<td>0.0513</td>
</tr>
<tr>
<td>WS0023.B03 †</td>
<td>25</td>
<td>0.378</td>
<td>0.859</td>
<td>0.846</td>
<td>***</td>
<td>0.3904</td>
</tr>
<tr>
<td>PA_33 †</td>
<td>4</td>
<td>0.088</td>
<td>0.101</td>
<td>0.098</td>
<td>ND</td>
<td>0.0737</td>
</tr>
<tr>
<td>WS0019.F22 †</td>
<td>16</td>
<td>0.355</td>
<td>0.629</td>
<td>0.577</td>
<td>***</td>
<td>0.2785</td>
</tr>
<tr>
<td>Pa_56 †</td>
<td>4</td>
<td>0.679</td>
<td>0.637</td>
<td>0.567</td>
<td>NS</td>
<td>−0.0346</td>
</tr>
<tr>
<td>PAAC23</td>
<td>36</td>
<td>0.457</td>
<td>0.879</td>
<td>0.868</td>
<td>***</td>
<td>0.3145</td>
</tr>
</tbody>
</table>

Number of allele in each locus (k), observed heterozygosity (H_o), expected heterozygosity (H_e), polymorphic information content (PIC), deviations from Hardy–Weinberg equilibrium (HW, level of significance: *** α = 0.001, * α = 0.05, NS—not significant, ND—not performed), estimation of null allele frequency (F(Null)), significant estimation in bold, SD—standard deviation, SE—standard error, † EST-SSRs, otherwise genomic SSRs.

Seven microsatellite loci—SpAGD1, WS00716.F13, WS0022.B15, SpAG2, PAAC3, WS0023.B03 and PAAC23—showed a polymorphic information content (PIC) higher than 0.85. The mean PIC for all loci was estimated to be 0.697 (SD = 0.223, SE = 0.063). Parameters of null alleles were estimated in CERVUS (ranging from −0.0663 (locus Pa_28) to 0.3904 (WS0023.B03)), and their statistical significance was checked in Microchecker. We detected seven loci with a significant occurrence of null alleles; these loci only partially overlap with highly polymorphic loci. Values of observed heterozygosity (H_o) ranged from 0.088 to 0.874, with the highest values found in the loci WS00716.F13 (0.874), SpAG2 (0.857) and WS0022.B15 (0.843). In contrast, the lowest values were detected at loci PA_33 (0.088) and EATC1E03 (0.276).

3.2. Discrimination Power of Assembled SSRs Markers

In order to back up a high determination power of a given microsatellite battery, we performed the Probability of Identity analysis (PI, GenAlEx). The graphical outcome (Figure 1) depicted an increasing discrimination power (decreasing Probability of Identity) to the increase of the effect of cumulative markers. For unrelated trees, the minimally sufficient number of markers for individual differentiation was considered to be 4 (PI = 7.2 × 10^{-4}), while for related individuals, the minimal amount of loci for statistically significant distinction was estimated to be 9 (PI = 7.9 × 10^{-4}). The increasing number of markers further diminishes the PI. For an entire set of 16 loci, the PI for related individuals of 7.9 × 10^{-7} and PI for unrelated individuals of 1.9 × 10^{-18} confirmed the high discrimination power of the assembled collection of markers.
The mean number of alleles ($N_a$) was closely similar between low-elevation (13.813) and high-elevation ecotypes (13.125), whereas in the medium-elevation ecotype, the mean number of alleles was higher (15.625). The number of effective alleles ($N_e$) varied between 5.594 for the high-elevation ecotype to 6.249 for the low-elevation ecotype. The mean value across all populations was estimated as $N_e = 5.984$, with an SE of 1.279. The mean values of population genetic diversity according to Shannon’s information index ($I$) ranged from 1.730 to 1.803. The fixation indexes for all ecotypes were positive, ranging from 0.128 to 0.175; this indicated a heterozygote deficiency, which is probably attributable to an occurrence of null alleles at several loci (Table 3). Also, the values of expected heterozygosity being exceeded for all ecotypic groups (LE: 0.715, ME: 0.701, HE: 0.698) showed heterozygosity (LE: 0.609, ME: 0.599, HE: 0.596).
All p-values indicated non-significant differentiation among population groups. The genetic distance across all populations expressed as a mean $F_{st}$ value was estimated to be 0.008 (SE 0.002). Estimation of overall gene flow ($N_{m} = 56.656$) implied a high gametic exchange among Norway spruce stands.

Molecular data were further processed with the model-based Bayesian clustering tool (software STRUCTURE, version 2.3.4). The true number of clusters estimated according to Evanno et al. [61] indicated $K = 1$. Therefore, for graphical output (Figure 3), we opted an original geographical distribution of $K = 3$ as an input parameter. In accordance with the parameters of F-statistics, we observed a low genetic differentiation among ecotypic groups.

![Figure 3. Bar plot analysis in the software STRUCTURE (K = 3). The probability of individuals are clustered into three populations. Top: individuals sorted in the original order; bottom: individuals within each population sorted by the estimated membership coefficient Q. Each individual corresponds to a colored column.](image)

An alternative method of population structure analysis, the Discriminant Analysis of Principal Components (DAPC), was used to depict genetic composition. The DAPC scatterplot demonstrated individual clustering inside and between populations. As the graphical layout documented (Figure 4), similar to the STRUCTURE outputs, all of the ecotypes showed a clear pattern of grouping, and individuals overlapped their ecotypic origin to a great extent. AMOVA dissected the molecular variance chiefly into within-individual (74%) and among-individual (23%) variance. Only 3% of the variance is attributable to among-population variance.

Table 4. Genetic population parameters of ecotypic groups (GenAlEx).

<table>
<thead>
<tr>
<th>Populations</th>
<th>$N$</th>
<th>$N_a$</th>
<th>$N_e$</th>
<th>$I$</th>
<th>$H_o$</th>
<th>$H_e$</th>
<th>$F_{is}$</th>
<th>$p$-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-elevation</td>
<td>Mean</td>
<td>143</td>
<td>13.813</td>
<td>6.249</td>
<td>1.803</td>
<td>0.609</td>
<td>0.715</td>
<td>0.137</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>2.199</td>
<td>1.258</td>
<td>0.200</td>
<td>0.050</td>
<td>0.055</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>Medium-elevation</td>
<td>Mean</td>
<td>150</td>
<td>15.625</td>
<td>5.984</td>
<td>1.749</td>
<td>0.599</td>
<td>0.701</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>2.115</td>
<td>1.321</td>
<td>0.201</td>
<td>0.058</td>
<td>0.056</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td>High-elevation</td>
<td>Mean</td>
<td>143</td>
<td>13.125</td>
<td>5.594</td>
<td>1.730</td>
<td>0.596</td>
<td>0.698</td>
<td>0.175</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>1.805</td>
<td>1.258</td>
<td>0.188</td>
<td>0.070</td>
<td>0.056</td>
<td>0.072</td>
<td></td>
</tr>
<tr>
<td>All individuals</td>
<td>Mean</td>
<td>436</td>
<td>14.211</td>
<td>5.943</td>
<td>1.760</td>
<td>0.601</td>
<td>0.705</td>
<td>0.146</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>2.041</td>
<td>1.279</td>
<td>0.197</td>
<td>0.054</td>
<td>0.055</td>
<td>0.054</td>
<td></td>
</tr>
</tbody>
</table>

$N$, sample size (number of sampled trees); $N_e$, mean number of alleles per locus; $N_a$, effective number of alleles; $I$, mean Shannon genetic diversity index; $H_o$, observed heterozygosity; $H_e$, expected heterozygosity; and $F_{is}$, mean fixation index with corresponding $p$-value based on permutation. All values except for $N$ (sample size) are given with standard errors (SE).
4. Discussion

In our study, the estimated parameters of mean heterozygosity ($H_o = 0.601, H_e = 0.705$) are comparable to previously published studies [34,66–70]; similarly, in our case, these are accompanied by a low level of genetic differentiation among populations. Similar values were observed despite the variety of experimental and sampling design, targeting various features and traits of interest. Our mean value of expected heterozygosity ($H_e = 0.705$) is similar to that in German ($H_e = 0.74$) [34] and Lithuanian populations ($H_e = 0.71$) [66]. Higher mean values were determined in Serbia ($H_e = 0.89$) [69], Tyrol ($H_e = 0.87$) [67], Slovenia ($H_e = 0.92$) [44], and the Czech Republic (0.86) [68]. Lower values were detected in another study of Serbian populations ($H_e = 0.62$) [69] and in Bosnia and Herzegovina ($H_e = 0.68$) [70]. As argued by Unger et al. [67], the homozygote excess ($H_o > H_e$) found in several studies [20,44,69–71] might be attributable to various evolutionary effects, but the occurrence of null alleles was presumed to be the most significant factor. As in the studies cited above, we decided not to omit loci with detected null alleles from the analysis.

In Norway spruce populations, the intrapopulation genetic diversity (based on F-statistics) is commonly higher than the variation among populations [72,73]. Our overall mean $F_{st}$ was estimated to be 0.008, which is similar to other findings. $F_{st}$ values range from 0.002 (Austria) [67] to 0.014 in Bosnia and Herzegovina [70].

Despite numerous studies of the genetic structure of Norway spruce, there are only a few studies which have utilized primers designed by the authors themselves. Hence, primers developed by Pfeiffer et al. [48] and Scotti et al. [51] are very frequently used [43,71,74]. Also, the primer sequences published by Fluch et al. [54], Rungis [47] and Besnard et al. [52] have been utilized frequently [20,34,66,69,70,75–77]. The maximal total number of SSR markers used in studies so far (11 primer pairs) was found in Cvjetkovic et al. [70], who exploited three sources [47,51,54]; in contrast, Caré et al. [34] yielded five primer sources [48,49,51,52,75]. Hence, our study increased the total number of SSR markers used for the analysis of population structure (16 primer pairs). We finally exploited six sources of primer sequences [47–49,51,52,54]. When considering the number of primer pairs pooled into one PCR reaction in previous publications, this varied from 1 (single) [66,71] to 3 [34,43,69,74,78] and 6 primer pairs assembled into one multiplex reaction [44,70]. Here, we present a more efficient...
genotyping protocol, combining 16 markers into just two multiplexes: multiplex 1 contains seven markers, whereas multiplex 2 was established from nine markers. As approved by the probability of identity analysis, a given number of loci is sufficient (PI for related individuals $7.9 \times 10^{-7}$, PI for unrelated $1.9 \times 10^{-18}$) to produce reliable results of population differentiation studies [79]. Moreover, the multiplexes assembled a large number of primers, improving the economic and time-consuming aspects of genotyping. We combined both genomic and EST-SSR markers (Table 2) into multiplexes. Our theoretical assumptions indicated the comparable usability of these two marker types, without any significant detection of differences in genetic population parameters [32]. Briefly, this is explained by a random distribution of SSRs across the spruce genome and non-selective effects of EST-SSRs for a polygenic trait (e.g., morphotypes). An advantage of EST-SSRs markers is generally the higher transferability across spruce [47,52,75–77,80,81] and related species [32,82–84], which is known as cross-amplification. While comparing the genomic and EST-SSR markers applied in this study, we did not detect a significant difference between them (based on $H_0$, $\alpha = 0.05$). Out of all 16 markers, we identified five which have the most significant effect on genetic clustering (Figure 1). As two of them are EST-SSRs (WS0092.A19, WS0019.F22) and three belong to genomic SSRs (paGB3, EATC1E03, PAAC3), we did not detect an apparent effect of different SSR marker types on clustering.

The nature of the outputs obtained may be due to two causes: either the inexpediency of the analytic method used, or the real genetic structure of forest stands. It is undeniable that the species composition of forest stands in Central Europe has been strongly altered by human impact for centuries. Since the 17th century, there has been a documented translocation and planting of non-local reproductive material throughout all the countries of Central Europe. Therefore, it is challenging to identify purely autochthonous stands, although our targeted locations were placed in protected areas and sites which are difficult to access, such as steep slopes where minimal human impact is assumed. Besides artificial migration and assisted gene flow through the extensive translocation of seeds, a high level of natural gene flow, which is well documented in Norway spruce [85], could also have significant effects on the genetic diversity of spruce stands [6,8,16]. The animal and wind dispersal of seeds seems to have a higher impact on gene flow than elevation gradients themselves [20,86]. Owing to the low average weight of pollen grain, pollen moved hundreds of kilometers, causing an intensive gene flow across long distances. Piotti et al. [74] concluded that only 11% of juvenile trees originated from adjacent paternal trees due to the high gene flow rate from distant areas. The estimated value of the migration coefficient ($Nm = 56.656$) indicates a high gene flow in our populations. The geographical distance of ecotypic units is in a range of 50 to 100 km, which is a similar distance to those in the study by Nowakowska et al. [71] in Poland ($Nm = 18.687$) and Cvjetković et al. [70] in Bosnia and Herzegovina ($Nm = 10.406$). A higher gene flow rate was found among populations in Lithuania ($Nm = 71.4$) by Verbylaite et al. [66], which is probably due to a closer geographical distance (i.e., a 20–80 km range) between them.

5. Conclusions

In our study, we aimed to unveil a genetic pattern of Norway spruce morphotype forms, which are also known as ecotypes. Our investigation utilized the presumably selective neutral genetic markers of microsatellites (simple sequence repeats, SSRs). Although there have previously been several studies evaluating the genetic structure of both natural and artificially planted spruce populations, none of them have aimed at a differentiation of three distinctive ecotypes. Moreover, for effective DNA fingerprinting, we developed a new microsatellite genotyping protocol by pooling compatible primer pairs into two multiplex reactions. As proved by our estimations (Probability of Identity values), the 16 loci genotyping protocol generates molecular data with a high discrimination power, and hence, these are potentially employable for various applied research activities, such as for a practical evaluation of alleged origin sources of seeds/seedlings.

Although we focused on a distinct trait of interest (ecotypic forms), we obtained values of genetic estimators of population which are comparable with the findings of the other authors. The
interpopulation differentiation was generally low, without any significant genetic grouping among targeted units of trees. Besides a potential influence of long-distance gene flow among spruce populations, this implies that an ecotypic differentiation is arguably caused by natural selection and other adaptive processes. While Norway spruce is the most abundant tree species in central Europe and its ecotypic forms apparently possess a different level of resistance against pests and pathogens, as well as shift of climatic conditions, we consider this study to be a stepping stone for further research efforts aiming to achieve applicable recommendations for forest management and forestry practice. In future studies, we plan to utilize genetic markers in a non-neutral manner (single nucleotide polymorphisms, SNPs) with the potential to reveal the genetic background of ecotypic clustering. More specifically, we will take advantage of any high-throughput genotyping method which generates a massive amount of non-neutral markers, such as the Genotyping by Sequencing approach [87]. Besides a better understanding of genetic substantiality and adaptation processes, we believe that using an alternative genotyping platform will very likely provide more informative data and thus will help to elucidate additional research questions such as the effect of phenological barriers on intrinsic gene flow among Norway spruce populations.

Supplementary Materials: The following are available online at http://www.mdpi.com/1999-4907/11/1/110/s1, Figure S1: Low-elevated ecotype (LE), Figure S2: High-elevated ecotype (HE), Figure S3: Medium-elevated ecotype (ME), Figure S4: Sampling design (LE), Figure S5: Sampling design (ME), Figure S6: Sampling design (HE West), Figure S7: Sampling design (HE East)

Author Contributions: Conceptualization, Z.B. and J.K.; Formal analysis, Z.B., J.K., J.B. and J.D.; Methodology, Z.B., J.K., J.B., J.D., D.Z. and V.J.; Resources, Z.B., D.Z. and V.J.; Software, Z.B., J.K. and J.D.; Supervision, J.K.; Writing—original draft, Z.B. and V.J.; Writing—review & editing, J.K., J.D. and M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the grant “Building up an excellent scientific team and its spatiotechnical background focused on mitigation of the impact of climatic changes to forests from the level of a gene to the level of a landscape (EXTEM-FK) at the FFWS CULS Prague”, No. CZ.02.1.01/0.0/0.0/15_003/0000433 financed by OP RDE and project of National Agency of Agriculture research, Czech Republic (NAZV, No. QK1910480).

Acknowledgments: We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (http://www.uerra.eu) and the data providers in the ECA&D project (https://www.ecad.eu).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References


44. Westergren, M.; Bozic, G.; Kraigher, H. Genetic diversity of core vs. peripheral Norway spruce native populations at a local scale in Slovenia. *iFor. Biogeosci. For.* **2018**, *11*, 104–110. [CrossRef]


63. Barker, M.; Rayens, W. Partial least squares for discrimination. *J. Chemom.* 2003, 17, 166–173. [CrossRef]


67. Máchová, P.; Trčková, O.; Cvrčková, H. Use of nuclear microsatellite loci for evaluating genetic diversity of selected populations of *Picea abies* (L.) karsten in the Czech Republic. *For. Trees* 2018, 9, 92. [CrossRef]


71. Porth, I.; El-Kassaby, Y. Assessment of the genetic diversity in forest tree populations using molecular markers. *Diversity* 2014, 6, 283–295. [CrossRef]


87. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *PLoS ONE* 2011, 6, e19379. [CrossRef] [PubMed]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).