

Review

Assessment of the Genetic Diversity in Forest Tree Populations Using Molecular Markers

Ilga Porth and Yousry A. El-Kassaby *

Department of Forest and Conservation Sciences, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada; E-Mail: porth@mail.ubc.ca

* Author to whom correspondence should be addressed; E-Mail: y.el-kassaby@ubc.ca;
Tel.: +1-604-822-1821; Fax: +1-604-822-9102.

*Received: 16 February 2014; in revised form: 20 March 2014 / Accepted: 27 March 2014 /
Published: 4 April 2014*

Abstract: Molecular markers have proven to be invaluable tools for assessing plants' genetic resources by improving our understanding with regards to the distribution and the extent of genetic variation within and among species. Recently developed marker technologies allow the uncovering of the extent of the genetic variation in an unprecedented way through increased coverage of the genome. Markers have diverse applications in plant sciences, but certain marker types, due to their inherent characteristics, have also shown their limitations. A combination of diverse marker types is usually recommended to provide an accurate assessment of the extent of intra- and inter-population genetic diversity of naturally distributed plant species on which proper conservation directives for species that are at risk of decline can be issued. Here, specifically, natural populations of forest trees are reviewed by summarizing published reports in terms of the status of genetic variation in the pure species. In general, for outbred forest tree species, the genetic diversity within populations is larger than among populations of the same species, indicative of a negligible local spatial structure. Additionally, as is the case for plants in general, the diversity at the phenotypic level is also much larger than at the marker level, as selectively neutral markers are commonly used to capture the extent of genetic variation. However, more and more, nucleotide diversity within candidate genes underlying adaptive traits are studied for signatures of selection at single sites. This adaptive genetic diversity constitutes important potential for future forest management and conservation purposes.

Keywords: forest trees; molecular markers; genetic diversity

1. Introduction

Forests constitute an integral part of the world's ecosystems, and approximately 80,000–100,000 different tree species are estimated to cover 31% of land area globally (Food and Agricultural Organization (FAO), United Nations). Especially in the developing world, people heavily rely on forest tree species for their livelihood, including the supply of wood-based fuels, as well as non-timber based products for nutrition, health and income. Forest trees are largely undomesticated and highly heterozygous, due to their outcrossing breeding systems [1] and, therefore, have large effective population sizes [2]. Despite the high number of known species, approximately 450 different forest tree species are actively part of a deliberate domestication process through tree improvement programs (FAO) [3]. The majority of the world-wide forests represent natural forests (93%), with 12% dedicated as conservation forests. A major concern regarding forests health and resilience is the declining in forest genetic diversity as documented as early as 1967 (FAO conference). Genetic diversity serves several important purposes: (a) as a resource for tree breeding and improvement programs to develop well-adapted tree species varieties and to enhance the genetic gain for a multitude of useful traits; (b) to ensure the vitality of forests as a whole by their capacity to withstand diverse biotic and abiotic stressors under changing and unpredictable environmental conditions; and (c) the livelihoods of indigenous and local communities that use traditional knowledge. Rich genetic diversity within and among forest tree species thus provides an important basis for maintaining food security and enabling sustainable development (FAO) [3].

Historically, for plant improvement, three major areas have always been important for molecular marker applications: (a) the determination of genetic diversity within and among populations; (b) verification and characterization of genotypes; and (c) marker-assisted selection (MAS) [4]. In particular, for forest trees that are outcrossing and largely undomesticated plant species, molecular markers have proven to be invaluable tools with applications in: (1) genetic conservation efforts by identification of genetic diversity hotspots; (2) the assembly of breeding populations in newly developed and advanced breeding programs; (3) the monitoring and characterization of population dynamics and gene flow; (4) the proper delineation of species taxonomy for management issues associated with conservation; (5) assessment of gene flow (pollen contamination) in seed orchards and the authentication of “controlled crossings”, the assessment of inbreeding occurrence in breeding programs and studies of mating systems in non-industrial tree species; and (6) genetic fingerprinting in advanced breeding programs for the purpose of quality control to detect misidentified ramets in production and breeding populations [4]. Although tree breeding programs would significantly benefit from an early selection of clones with advantageous trait characteristics (particularly important for late-expressing wood quality traits, [5]), MAS was deemed not feasible for forest trees with limited genetic marker coverage [6]. The main reasons for the infeasibility of MAS as a tool for forest tree improvement are the inherent characteristics specific of forest trees as compared to inbred agricultural crop plants, such as the polygenic nature of most of the economically important traits in forestry, the inconsistency in quantitative trait locus (QTL) marker linkages among families originating from large outcrossed breeding populations and the instability of QTLs from the same genetic material planted across different sites, due to strong genotype-by-environment ($G \times E$) interactions. As highly efficient next generation SNP (single nucleotide polymorphism) genotyping platforms have become available,

genome-wide selection approaches have become feasible for accelerating forest tree breeding [7,8]. Here, we review the status of genetic variability in forest trees as assessed by molecular markers.

2. Marker Types and Their Applications

We begin our review with regards to the genetic diversity in forest tree species with a brief historical retrospect concerning the development of marker types that have been widely employed for studying genetic variability in plants in general. The first, while the most easily accessible types of plant characteristics, are morphological markers that can easily be monitored based on simple inheritance [9]. However, due to serious drawbacks with respect to dominance, the difficulty of distinguishing between multiple alleles or even between different loci [10,11] and trait expression due to environmental and developmental variation ($G \times E$ interaction), their use was substantially reduced with the advent of DNA marker technologies. Another marker type that played an important role in assessing genetic diversity in plants was isozymes [12,13]. Isozymes had a long history in genetic variability studies in forestry, to assess the genetic diversity present within natural forest stands [14,15] or to determine whether domestication practices had led to a reduction in diversity [16–18]. However, the problem of these biochemical marker assays is that they are affected by plant phenological stage and their limited availability, and therefore, they would never allow for a genome-wide scan of variability (as only 0.1% of the total variation is detectable by this technique, [19]). An invaluable alternative offered DNA-based markers, such as restriction fragment length polymorphism (RFLPs) [20–22]. Finally, the possibility to rapidly amplify specific DNA fragments *in vitro* via polymerase chain reaction (PCR) [23] revolutionized the generation of molecular markers, leading to diverse sets of diagnostic DNA-marker systems with or without *a priori* sequence knowledge, such as random amplified polymorphic DNA (a.k.a RAPD) [24], amplified fragment length polymorphism (a.k.a AFLP) [25], simple sequence repeats (a.k.a SSRs or microsatellites) [26], single nucleotide polymorphisms (a.k.a SNPs) [27,28] and variations thereof [29,30]. Important issues are related to the reproducibility of the RAPD marker system [31], other limitations, such as the presence of null alleles in the case of SSR assays that may underestimate heterozygosity [32], or the dominance nature of the RAPD and AFLP marker systems, where heterozygous individuals cannot be distinguished from homozygous ones, and lastly, the inexpensive generation of a vast abundance of highly polymorphic DNA markers to tackle genome-wide genetic diversity studies. Dependent on the study focus, genetic markers were derived from nuclear or organelle sequences; for example, chloroplast- or mitochondrial-derived diagnostic markers [33–35], dependent on the evidence of their maternal inheritance in the species, were used to trace back the colonization history of angiosperm forest tree species and conifers, respectively [36,37]. Although it has been known that variability within protein-coding regions is far less than within non-coding genomic regions, due to lower mutation rates and purifying selection to maintain proper protein functions, the study of polymorphic sites within coding sequences has been deemed more relevant because of their putative functional associations and, in addition, the ease of their interspecific transferability for comparative genetic studies based on sequence conservation. Thus, a major focus in plant studies has been the development of genetic markers prevalently present within such coding regions for high-throughput analysis of many samples using the inexpensive detection method of PCR fragment length polymorphisms (e.g., eco-tilling to circumvent expensive

Sanger resequencing of PCR products, as in the case of SNP detection and genotyping [38]), but that still relied on laborious PCR optimizations (e.g., [39–42]). The substantial and almost exponential drop in whole genome sequencing costs, thanks to the 1000 Human Genome Project, which has stimulated the development of highly cost-efficient high-throughput technologies, has also provided for the plant research community unprecedented opportunities for affordable in-depth characterization of plant genomes that has involved the genome-wide discovery of SSRs and SNPs and the detection of common, as well as rare functional variants by next generation sequencing [43–49].

3. Assessment of Genetic Diversity

A number of evolutionary processes can impact the genetic diversity of natural populations. These are: (a) spontaneously arising mutations; (b) gene flow via migration; (c) inbreeding; (d) natural selection; (e) the Wahlund effect; and (f) random genetic drift [50]. Genetic drift introduces random changes in allele frequencies over generations and becomes important for finite population samples and/or a large number of generations. These random allele frequency changes can, over time, lead to allele fixation or extinction. By all means, genetic drift represents a source of differences in genetic diversity among different populations. On the other hand, gene flow evens out among-population genetic differences, but increases genetic variation within populations, due to the introduction of new alleles. Selection influences within-population diversity, but the effects are dependent on the nature of these selection processes (balancing selection). Furthermore, the effects of natural selection are interwoven with stochastic effects, such as genetic drift. Mutations can counterbalance the loss of allelic diversity; however, natural mutations are rare, and such mutations that turn out to be harmful allelic variants are again removed by purifying selection. The occurrence of a population bottleneck causes a significant reduction in the effective population size and represents a major reason for the loss in allelic diversity, first by the loss of rare alleles, then by the successive loss of heterozygosity in the population [50]. Inbreeding and the presence of a subpopulation structure, where gene flow is prevented by habitat fragmentation (the Wahlund effect), both cause the loss in heterozygosity [50]. This, in turn, results in increased genetic diversity among populations.

3.1. Within-Population Genetic Variation Using Genotype Data

A gene is defined as polymorphic in the population when its most common allele is less frequent than 95% [50]. Genetic diversity can be assessed by estimating the following parameters: the total number of different alleles in the population, the percentage of polymorphic loci, the mean number of alleles per locus, the allelic richness, the within-population genetic diversity, θ , the effective population size, N_e (i.e., θ divided by the per-generation mutation rate), the minor allele frequency (as in the case of biallelic loci), the proportion of heterozygous individuals in the population for a given locus (the expected heterozygosity, H_E ; based on the Hardy-Weinberg expectations that assume the random mating of genotypes), as well as the observed heterozygosity (H_O) and the fixation index, F [50]. Genomic diversity is estimated by genome-wide assessment of genetic diversity using a larger sample of loci at random. An estimate of the genome-wide genetic diversity in a population is then derived by averaging heterozygosity over the multitude of studied loci.

3.2. Between-/Among-Population Genetic Variation Using Genotype Data

Differences in the genetic diversity between/among (sub-)populations are assessed based on the presence of significant allele frequency differences; widely applied metrics to estimate such “genetic differentiation” include, for example, F_{ST} [51,52], θ [53], R_{ST} [54], Φ_{ST} (Φ'_{ST}) [55,56], G_{ST} (G'_{ST}) [57,58], D_{ST} [57], H_{ST} [59] or D [60]. Some measures are marker-dependent; they are based on the assumption of infinite-allele or stepwise mutation models, respectively, and depending on whether biallelic or multi-allelic molecular markers or haplotype data were used in the analysis (F_{ST} ; R_{ST} ; Φ_{ST}). Moreover, the use of fixation measures for result interpretation with regards to genetic differentiation has been found to be problematic when the populations under study exhibited high genetic diversity/heterozygosity (*cf.* G_{ST} ; [58,60]). For such cases, “standardized” genetic differentiation metrics’ have been suggested ([56,58,60]); but, see also the recent publication on the topic by Whitlock [61], who emphasized the continuous use of F_{ST} for intra-specific differentiation estimation when the mutation rate is small (relative to gene flow), while emphasizing the use of Φ_{ST} and R_{ST} when the mutation rate is high (as in the case of SSRs). In any case, for the estimation of population divergence from genotypic data, freely available software packages within the R environment [62] that have these statistics implemented are readily available (*cf.* “mmod”). Furthermore, genetic loci with allelic frequencies significantly different among populations and potentially under selection (“ F_{ST} outlier loci”) can be efficiently detected using multilocus scans that compare the patterns of nucleotide diversity and genetic differentiation (based on the distribution of empirical F_{ST} estimates conditioned on H_E) to the simulated genome-wide selectively-neutral genetic background [63,64].

3.3. Sequence Divergence Using Sequence Alignment Data

Other and additional ways to look at genetic diversity and study mutation and selection events within populations and by comparing different populations involve the characterization of DNA sequences of genes and the diversity of nucleotides as the specific study entities [65–68]. Widely used tests include nucleotide diversity π [50], the estimation of Tajima’s D , Fu’s D , Fay and Wu’s H , Zeng *et al.*’s E [69–72] and the McDonald–Kreitman and HKA (Hudson–Kreitman–Aguade) tests [73,74], respectively. Such tests are implemented in the freely available software package, DnaSP [75]. The combination of results from such analyses has particular value for identifying past population size changes (population expansion or population bottleneck).

4. Forest Tree Population Diversity

One of the first comprehensive reviews on genetic diversity with regards to forest tree populations was published by Hamrick and co-workers [76]. This early work summarized results based on isozymes and is especially valuable, as it compares long-lived forest trees with other life forms of plant species, in total comprising 662 different species with representatively high sample sizes for the analysis of the genetic diversity parameters. Long-lived, woody species showed the highest genetic diversity (including a significantly higher percentage of polymorphic loci and more alleles per locus) among all plant species. Specifically, the genetic diversity within populations was significantly the highest ($H_E = 0.15$) compared to all other plant life forms ($H_E < 0.10$). However, heterogeneity in

genetic diversity exists among woody species taxa, and this is due to the different evolutionary histories of species. For example, species from smaller founder populations, small disjunct populations or those with past population bottlenecks show generally less genetic diversity. *Alseis blackiana*, *Picea glauca*, *Robinia pseudoacacia* and *Pinus sylvestris* showed high diversity. On the other side of the spectrum were *Acacia mangium*, *Pinus resinosa*, *P. torreyana* and *Populus balsamea* with very low diversity [76]. Other studies [77,78] identified additional species with low intra-population diversity: *Ficus carica* and *Thuja plicata*.

While most studies identified high intra-population variation, by contrast, the diversity among populations of long-lived, woody tree species based on the G_{ST} estimate was significantly the lowest ($G_{ST} = 0.08$) compared to the herbaceous and annual life forms ($G_{ST} > 0.25$) [76]. When woody angiosperms were compared to gymnosperms in terms of their intra-population genetic diversity, differences were not significant, yet the latter exhibited a significantly higher percentage of polymorphic allozyme loci, suggestive of a higher proportion of low frequency alleles in gymnosperm species [76]. Angiosperm species showed higher among-population genetic diversity (G_{ST}). Recent research on the conifer genome evolution, which involved orthologous coding sequence alignments for thousands of gymnosperms and angiosperm orthologous coding sequences, respectively, showed, more specifically, an overrepresentation of non-synonymous substitutions in protein-coding genes for conifers compared to angiosperms [79], while the average synonymous mutation rate in angiosperms is significantly higher, suggestive of a higher number of fixed adaptive mutations in conifers. As expected, the extent of the geographical range had a significant impact on genetic diversity within species and among populations [76]. Geographically widespread species showed a significantly higher intra-population genetic diversity estimate compared to locally confined species, but the latter showed higher genetic diversity among populations [76]. However, the “non-significant” inter-population differentiation sometimes reported in these isozyme studies (see above) can mislead the directions of conservation efforts. Other marker types, those that are able to cover a higher portion of the overall genetic variation (such as restriction fragment length polymorphisms of DNA) succeeded in uncovering significant among-population diversity in *Pinus* and *Quercus*, specifically with the application of organellar DNA markers (cf. [80,81]). Differing outcomes for isozymes and organellar DNA studies on population divergence are frequent and were even reported within the same sample as for *Argania spinosa* (L.) Skeels, an important multi-purpose tree in the Moroccan local community [82]. It is also clear that variation at selectively neutral molecular markers commonly used to assess genetic diversity within or among populations may not covary with the phenotypic expression of a particular qualitative or quantitative trait of interest [29], such that population differentiation for adaptive traits (growth, morphology or fitness) is much higher than for isozymes, for example. In any case, the total allelic richness was identified as a more adequate directive than the H_E estimate for conservation purposes, and marker types, such as SSRs or DNA sequence-based data, that are highly polymorphic are required for an accurate estimate [82]. A recent study integrating molecular genetic analysis based on four SSR and five sequence loci along with climate modeling [83] forecasted the long-term decline of the late-successional Australian rainforest conifer, *Podocarpus elatus*, in its southern populations, due to habitat fragmentation (and the decline in N_e), for which conservation strategies are now invoked. Isozyme markers (15 loci) were used to characterize the genetic diversity of *Carapa procera*, which occurs in low density within a tropical rain forest [15]. Its characteristics were high within-population

diversity (comparable to temperate gymnosperms), high heterozygosity and a lack of spatial structure consistent with the highly outcrossing nature of the species, leading to extensive pollen-mediated gene flow that prevented local genetic differentiation. When 63 SNP polymorphisms (surveyed by eco-tilling) in nine different genes with broad functional properties were targeted as a feature for understanding DNA variation in 41 wild populations of a small western black cottonwood (*P. trichocarpa*) sample panel [40], it was found that heterozygosity was high ($H_O = 0.47$) and that overall nucleotide diversity at the gene level ($\pi = 0.0018$) among populations was low. Similarly, low average π values of the segregating sites were obtained for other forest tree species, such as *P. nigra* ($\pi = 0.0024$; [84]) and *Pinus sylvestris* ($\pi = 0.0025$; [85]). Much higher overall nucleotide diversity levels in a conifer were uncovered for *P. taeda* ($\pi = 0.00398$; [86]). Among the studied poplars, interestingly, the European species, *P. tremula*, showed the highest nucleotide diversity ($\pi = 0.007$ [87] or even $\pi = 0.0111$ [88], dependent on the surveyed genes), but differences in diversity were also consistent with its different and complex demographic history. However, nucleotide diversity is best interpreted on a gene-by-gene basis, as population history and selection affect these mutation rates more specifically [40,89]. In a similar context, assessing the adaptive genetic diversity in forest trees is important to harness this adaptive potential for future forest management and conservation purposes [90]. Candidate genes underlying a specific trait of interest are typically selected (*cf.* nine candidate genes for bud burst in *Quercus petraea*: $\pi = 0.00615$ [91]; 121 candidate genes for cold hardiness in *Pseudotsuga menziesii* var. *menziesii*: $\pi = 0.004$ [92]; 13 candidate genes for drought stress in *Pinus pinaster* $\pi = 0.00548$ [64]). While most of this detected variation was largely attributed to purifying selection (an excess of nucleotide diversity at synonymous vs. non-synonymous sites), as commonly observed in forest trees, patterns of strong diversifying selection in candidate genes were also uncovered [64].

5. Conclusions

This review summarizes the major molecular marker types that have been developed to replace the more problematic phenotypic markers in plants used at the infancy of genetic diversity studies. While we touched on their most important applications and showed how broad such applications have been, here we focused specifically on forest genetic diversity studies. We emphasize that integrative approaches using future climate modeling have been very successful in uncovering potential threats of declines of the genetic diversity and the distribution of forest tree species, so that timely precautions to preserve the species can be undertaken. Associated with the substantial drop in whole genome sequencing costs making the sequencing of genetically complex organisms more affordable, inventorying the complete portfolio of genetic resources has become feasible. This will also open new avenues for the conservation of previously marginalized and undervalued forest tree species that are considered of less economic value, but nevertheless represent value to the local ecosystems. While the present review focused primarily on the genetic diversity assessed for pure species, we also stress the importance of investigating natural species hybrid zones as important sources of population genetic diversity in forest tree management [93,94]. Finally, we stress the value of integrating knowledge on adaptive complex traits as a companion to molecular markers for making informative management and conservation decisions.

Acknowledgments

Funds from the Johnson's Family Forest Biotechnology Endowment, The British Columbia Forest Investment Account Forest Genetics Conservation and Management Program, the National Science and Engineering Research Council of Canada Discovery and Industrial Research Chair (Yousry A. El-Kassaby) are greatly appreciated.

Author Contributions

Ilga Porth and Yousry A. El-Kassaby both contributed substantially to the conception and drafting of the review. Both authors approved the final version.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Schoen, D.J.; Brown, A.H.D. Intraspecific variation in population gene diversity and effective population-size correlates with the mating system in plants. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 4494–4497.
2. Jansson, S.; Ingvarsson, P.K. Cohort-structured tree populations. *Heredity* **2010**, *105*, 331–332.
3. Food and Agriculture Organization of the United Nations. Report of the 14th Regular Session of the Commission on Genetic Resources for Food and Agriculture. Available online: <http://www.fao.org/docrep/meeting/028/mg468e.pdf> (accessed on 16 February 2014).
4. Haines, R. *Biotechnology in Forest Tree Improvement: With Special Reference to Developing Countries*, 1st ed.; Food and Agriculture Organization of the United Nations: Rome, Italy, 1994; p. 230.
5. Lapinjoki, S.; Tammisola, J.; Kauppinen, V.; Wright, A.; Vihera-Aarnie, A.; Hagqvist, R.; Sundquist, J.; Martimo, T. The Finnish project for genetic mapping and its applications in breeding, biotechnology and research of white birch. In Proceedings of the Sixth Meeting of the International Conifer Biotechnology Working Group, Raleigh, NC, USA, 23 April 1992.
6. Strauss, S.H.; Lande, R.; Namkoong, G. Limitations of molecular-marker-aided selection in forest tree breeding. *Can. J. Forest Res.* **1992**, *22*, 1050–1061.
7. Grattapaglia, D.; Resende, M.D.V. Genomic selection in forest tree breeding. *Tree Genet. Genomes* **2011**, *7*, 241–255.
8. Resende, M.D.V.; Resende, M.F.R., Jr.; Sansaloni, C.P.; Petrolí, C.D.; Missiaggia, A.A.; Aguiar, A.M.; Abad, J.M.; Takahashi, E.K.; Rosado, A.M.; Faria, D.A.; *et al.* Genomic selection for growth and wood quality in Eucalyptus: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytol.* **2012**, *194*, 116–128.
9. Jain, S.K.; Allard, R.W. Population studies in predominantly self-pollinated species, 1. Evidence for heterozygote advantage in a closed population of barley. *Proc. Natl. Acad. Sci. USA* **1960**, *46*, 1371–1377.
10. Brown, A.H.D. Isozymes, plant population genetic structure and genetic conservation. *Theor. Appl. Genet.* **1978**, *52*, 145–157.

11. Andersen, J.R.; Lubberstedt, T. Functional markers in plants. *Trends Plant Sci.* **2003**, *8*, 554–560.
12. Brown, A.H.D. Enzyme polymorphism in plant-populations. *Theor. Popul. Biol.* **1979**, *15*, 1–42.
13. Ritland, K.; Jain, S. A Comparative-study of floral and electrophoretic variation with life-history variation in *Limnanthes-alba* (Limnanthaceae). *Oecologia* **1984**, *63*, 243–251.
14. Ritland, K.; Meagher, L.D.; Edwards, D.G.W.; El-Kassaby, Y.A. Isozyme variation and the conservation genetics of Garry oak. *Can. J. Bot.* **2005**, *83*, 1478–1487.
15. Doligez, A.; Joly, H.I. Genetic diversity and spatial structure within a natural stand of a tropical forest tree species, *Carapa procera* (Meliaceae), in French Guiana. *Heredity* **1997**, *79*, 72–82.
16. El-Kassaby, Y.; Ritland, K. Impact of selection and breeding on the genetic diversity in Douglas-fir. *Biodivers. Conserv.* **1996**, *5*, 795–813.
17. El-Kassaby, Y.A. Domestication and genetic diversity—Should we be concerned? *Forest. Chron.* **1992**, *68*, 687–700.
18. Chaisurisri, K.; El-Kassaby, Y.A. Genetic diversity in a seed production population vs. natural-populations of sitka spruce. *Biodivers. Conserv.* **1994**, *3*, 512–523.
19. El-Kassaby, Y.A. Genetic variation within and among conifer populations: Review and evaluation of methods. In *Biochemical Markers in the Population Genetics of Forest Trees*; Fineschi, S., Malvolti, M.E., Cannata, F., Hattemer, H.H., Eds.; SPB Academic Publishing: The Hague, The Netherlands, 1991; pp. 61–76.
20. Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. Construction of a genetic-linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **1980**, *32*, 314–331.
21. Szmidt, A.E.; El-Kassaby, Y.A.; Sigurgeirsson, A.; Alden, T.; Lindgren, D.; Hallgren, J.E. Classifying seedlots of *Picea-sitchensis* and *Picea-glauca* in zones of introgression using restriction analysis of chloroplast DNA. *Theor. Appl. Genet.* **1988**, *76*, 841–845.
22. Glaubitz, J.C.; El-Kassaby, Y.A.; Carlson, J.E. Nuclear restriction fragment length polymorphism analysis of genetic diversity in western redcedar. *Can. J. Forest Res.* **2000**, *30*, 379–389.
23. Mullis, K.; Faloona, F.; Scharf, S.; Saiki, R.; Horn, G.; Erlich, H. Specific enzymatic amplification of DNA *in vitro*—The polymerase chain-reaction. *Cold Spring Harb. Symp. Quant. Biol.* **1986**, *51*, 263–273.
24. Williams, J.G.K.; Kubelik, A.R.; Livak, K.J.; Rafalski, J.A.; Tingey, S.V. DNA polymorphisms amplified by arbitrary primers are useful as genetic-markers. *Nucleic Acids Res.* **1990**, *18*, 6531–6535.
25. Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; Vandelee, T.; Hornes, M.; Frijters, A.; Pot, J.; Peleman, J.; Kuiper, M. AFLP—A new technique for DNA-fingerprinting. *Nucleic Acids Res.* **1995**, *23*, 4407–4414.
26. Zietkiewicz, E.; Rafalski, A.; Labuda, D. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain-reaction amplification. *Genomics* **1994**, *20*, 176–183.
27. Rafalski, A. Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* **2002**, *5*, 94–100.
28. Buckler, E.S.; Thornsberry, J.M. Plant molecular diversity and applications to genomics. *Curr. Opin. Plant Biol.* **2002**, *5*, 107–111.
29. Mondini, L.; Noorani, A.; Pagnotta, M. Assessing plant genetic diversity by molecular tools. *Diversity* **2009**, *1*, 19–35.

30. Nybom, H.; Weising, K.; Rotter, B. DNA fingerprinting in botany: Past, present, future. *Investig. Genet.* **2014**, *5*, 1–35.
31. Perez, T.; Albornoz, J.; Dominguez, A. An evaluation of RAPD fragment reproducibility and nature. *Mol. Ecol.* **1998**, *7*, 1347–1357.
32. Callen, D.F.; Thompson, A.D.; Shen, Y.; Phillips, H.A.; Richards, R.I.; Mulley, J.C.; Sutherland, G.R. Incidence and origin of “null” alleles in the (AC)_n microsatellite markers. *Am. J. Hum. Genet.* **1993**, *52*, 922–927.
33. Sutton, B.C.S.; Flanagan, D.J.; El-Kassaby, Y.A. A simple and rapid method for estimating representation of species in spruce seedlots using chloroplast DNA restriction-fragment-length-polymorphism. *Silvae Genet.* **1991**, *40*, 119–123.
34. Sutton, B.C.S.; Flanagan, D.J.; Gawley, J.R.; Newton, C.H.; Lester, D.T.; El-Kassaby, Y.A. Inheritance of chloroplast and mitochondrial-dna in *Picea* and composition of hybrids from introgression zones. *Theor. Appl. Genet.* **1991**, *82*, 242–248.
35. Demesure, B.; Sodzi, N.; Petit, R.J. A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast dna in plants. *Mol. Ecol.* **1995**, *4*, 129–131.
36. Petit, R.J.; Csaikl, U.M.; Bordacs, S.; Burg, K.; Coart, E.; Cottrell, J.; van Dam, B.; Deans, J.D.; Dumolin-Lapègue, S.; Fineschi, S.; *et al.* Chloroplast DNA variation in European white oaks—Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecol. Manag.* **2002**, *156*, 5–26.
37. Gamache, I.; Jaramillo-Correa, J.P.; Payette, S.; Bousquet, J. Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: Imprint of a founder effect associated with postglacial colonization. *Mol. Ecol.* **2003**, *12*, 891–901.
38. Comai, L.; Young, K.; Till, B.J.; Reynolds, S.H.; Greene, E.A.; Codomo, C.A.; Enns, L.C.; Johnson, J.E.; Burtner, C.; Odden, A.R.; *et al.* Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J.* **2004**, *37*, 778–786.
39. Parida, S.K.; Kumar, K.A.R.; Dalal, V.; Singh, N.K.; Mohapatra, T. Unigene derived microsatellite markers for the cereal genomes. *Theor. Appl. Genet.* **2006**, *112*, 808–817.
40. Gilchrist, E.J.; Haughn, G.W.; Ying, C.C.; Otto, S.P.; Zhuang, J.; Cheung, D.; Hamberger, B.; Aboutorabi, F.; Kalynyak, T.; Johnson, L.; *et al.* Use of Ecotilling as an efficient SNP discovery tool to survey genetic variation in wild populations of *Populus trichocarpa*. *Mol. Ecol.* **2006**, *15*, 1367–1378.
41. Liewlaksaneeyanawin, C.; Zhuang, J.; Tang, M.; Farzaneh, N.; Lueng, G.; Cullis, C.; Findlay, S.; Ritland, C.E.; Bohlmann, J.; Ritland, K. Identification of COS markers in the Pinaceae. *Tree Genet. Genomes* **2009**, *5*, 247–255.
42. Bodénès, C.; Chancerel, E.; Gailing, O.; Vendramin, G.G.; Bagnoli, F.; Durand, J.; Goicoechea, P.G.; Soliani, C.; Villani, F.; Mattioni, C.; *et al.* Comparative mapping in the Fagaceae and beyond with EST-SSRs. *BMC Plant Biol.* **2012**, *12*, 153–170.
43. Poland, J.A.; Rife, T.W. Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* **2012**, *5*, 92–102.
44. Davey, J.W.; Hohenlohe, P.A.; Etter, P.D.; Boone, J.Q.; Catchen, J.M.; Blaxter, M.L. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **2011**, *12*, 499–510.

45. Zalapa, J.E.; Cuevas, H.; Zhu, H.Y.; Steffan, S.; Senalik, D.; Zeldin, E.; McCown, B.; Harbut, R.; Simon, P. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *Am. J. Bot.* **2012**, *99*, 193–208.
46. Geraldes, A.; Pang, J.; Thiessen, N.; Cezard, T.; Moore, R.; Zhao, Y.; Tam, A.; Wang, S.; Friedmann, M.; Birol, I.; *et al.* SNP discovery in black cottonwood (*Populus trichocarpa*) by population transcriptome resequencing. *Mol. Ecol. Resour.* **2011**, *11*, 81–92.
47. Chen, C.; Mitchell, S.; Elshire, R.; Buckler, E.; El-Kassaby, Y. Mining conifers' mega-genome using rapid and efficient multiplexed high-throughput genotyping-by-sequencing (GBS) SNP discovery platform. *Tree Genet. Genomes* **2013**, *9*, 1537–1544.
48. Tsai, H.; Howell, T.; Nitcher, R.; Missirian, V.; Watson, B.; Ngo, K.J.; Lieberman, M.; Fass, J.; Uauy, C.; Tran, R.K.; *et al.* Discovery of rare mutations in populations: TILLING by sequencing. *Plant Physiol.* **2011**, *156*, 1257–1268.
49. Marroni, F.; Pinosio, S.; di Centa, E.; Jurman, I.; Boerjan, W.; Felice, N.; Cattonaro, F.; Morgante, M. Large-scale detection of rare variants via pooled multiplexed next-generation sequencing: Towards next-generation Ecotilling. *Plant J.* **2011**, *67*, 736–745.
50. Hartl, D.L.; Clark, A.G. *Principles of Population Genetics*, 4th ed.; Sinauer Associates Inc.: Sunderland, MA, USA, 2007; p. 652.
51. Wright, S. *Evolution and the Genetics of Populations, Volume 2: The Theory of Gene Frequencies*; University of Chicago Press: Chicago, IL, USA, 1984; p. 519.
52. Raymond, M.; Rousset, F. An exact test for population differentiation. *Evolution* **1995**, *49*, 1280–1283.
53. Weir, B.S.; Cockerham, C.C. Estimating f-statistics for the analysis of population-structure. *Evolution* **1984**, *38*, 1358–1370.
54. Slatkin, M. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **1995**, *139*, 1463–1463.
55. Excoffier, L.; Smouse, P.E.; Quattro, J.M. Analysis of molecular variance inferred from metric distances among DNA haplotypes—Application to human mitochondrial-DNA restriction data. *Genetics* **1992**, *131*, 479–491.
56. Meirmans, P.G. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* **2006**, *60*, 2399–2402.
57. Nei, M. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 3321–3323.
58. Hedrick, P.W. A standardized genetic differentiation measure. *Evolution* **2005**, *59*, 1633–1638.
59. Lewontin, R.C. The apportionment of human diversity. In *Evolutionary Biology, Volume 6*; Dobzhansky, T., Hecht, M.K., Steere, W.C., Eds.; Appleton-Century-Crofts: New York, NY, USA, 1972; pp. 381–398.
60. Jost, L. G(ST) and its relatives do not measure differentiation. *Mol. Ecol.* **2008**, *17*, 4015–4026.
61. Whitlock, M.C. G'(ST) and D do not replace F(ST). *Mol. Ecol.* **2011**, *20*, 1083–1091.
62. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008. Available online: <http://www.R-project.org> (accessed on 16 February 2014).
63. Beaumont, M.A.; Nichols, R.A. Evaluating loci for use in the genetic analysis of population structure. *Prod. R. Soc. B* **1996**, *263*, 1619–1626.

64. Eveno, E.; Collada, C.; Guevara, M.A.; Leger, V.; Soto, A.; Diaz, L.; Léger, P.; González-Martínez, S.C.; Cervera, M.T.; Plomion, C.; *et al.* Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. *Mol. Biol. Evol.* **2008**, *25*, 417–437.
65. Kimura, M. Number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* **1969**, *61*, 893–903.
66. Ohta, T. Mutational pressure as main cause of molecular evolution and polymorphism. *Nature* **1974**, *252*, 351–354.
67. Tajima, F. Evolutionary relationship of DNA-sequences in finite populations. *Genetics* **1983**, *105*, 437–460.
68. Kreitman, M.E.; Aguade, M. Excess polymorphism at the *Adh* locus in *Drosophila-melanogaster*. *Genetics* **1986**, *114*, 93–110.
69. Tajima, F. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **1989**, *123*, 585–595.
70. Fu, Y.X.; Li, W.H. Statistical tests of neutrality of mutations. *Genetics* **1993**, *133*, 693–709.
71. Fay, J.C.; Wu, C.I. Hitchhiking under positive Darwinian selection. *Genetics* **2000**, *155*, 1405–1413.
72. Zeng, K.; Fu, Y.X.; Shi, S.; Wu, C.-I. Statistical tests for detecting positive selection by utilizing high-frequency variants. *Genetics* **2006**, *174*, 1431–1439.
73. McDonald, J.H.; Kreitman, M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **1991**, *351*, 652–654.
74. Hudson, R.R.; Kreitman, M.; Aguade, M. A test of neutral molecular evolution based on nucleotide data. *Genetics* **1987**, *116*, 153–159.
75. Rozas, J.; Sanchez-DelBarrio, J.C.; Messeguer, X.; Rozas, R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **2003**, *19*, 2496–2497.
76. Hamrick, J.L.; Godt, M.J.W.; Sherman-Broyles, S.L. Factors influencing levels of genetic diversity in woody plant species. *New Forest.* **1992**, *6*, 95–124.
77. Mueller-Starck, G.; Baradat, P.; Bergmann, F. Genetic variation within European tree species. *New Forest.* **1992**, *6*, 23–47.
78. Neale, D.B.; Williams, C.G. Restriction-fragment-length-polymorphism mapping in conifers and applications to forest genetics and tree improvement. *Can. J. Forest Res.* **1991**, *21*, 545–554.
79. Buschiazzo, E.; Ritland, C.; Bohlmann, J.; Ritland, K. Slow but not low: Genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. *BMC Evol. Biol.* **2012**, *12*, 8–21.
80. Strauss, S.H.; Hong, Y.P.; Hipkins, V.D. High-levels of population differentiation for mitochondrial-DNA haplotypes in *Pinus-radiata*, *muricata*, and *attenuata*. *Theor. Appl. Genet.* **1993**, *86*, 605–611.
81. Kremer, A.; Petit, R.J.; Zanetto, A.; Fougere, V.; Ducouso, A.; Wagner, D.B.; Chauvin, C. Nuclear and organelle gene diversity in *Quercus robur* and *Q. petraea*. In *Genetic Variation in European Populations of Forest Trees*; Müller-Starck, G., Ziehe, M., Eds.; Sauerländer's Verlag: Frankfurt am Main, Germany, 1991; pp. 141–166.

82. Petit, R.J.; El Mousadik, A.; Pons, O. Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol.* **1998**, *12*, 844–855.
83. Mellick, R.; Rossetto, M.; Allen, C.; Wilson, P.D.; Hill, R.S.; Lowe, A. Intraspecific divergence associated with a biogeographic barrier and climatic models show future threats and long-term decline of a rainforest conifer. *The Open Conservat. Biol. J.* **2013**, *7*, 1–10.
84. Zaina, G.; Morgante, M. Nucleotide Diversity and Linkage Disequilibrium in *Populus nigra*. In Proceedings of the XLVIII Italian Society of Agricultural Genetics—SIFV-SIGA Joint Meeting, Lecce, Italy, 15–18 September 2004.
85. Garcia-Gil, M.R.; Mikkonen, M.; Savolainen, O. Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. *Mol. Ecol.* **2003**, *12*, 1195–1206.
86. Brown, G.R.; Gill, G.P.; Kuntz, R.J.; Langley, C.H.; Neale, D.B. Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15255–15260.
87. Ismail, M.; Soolanayakanahally, R.Y.; Ingvarsson, P.K.; Guy, R.D.; Jansson, S.; Silim, S.N.; El-Kassaby, Y.A. Comparative nucleotide diversity across North American and European populus species. *J. Mol. Evol.* **2012**, *74*, 257–272.
88. Ingvarsson, P.K. Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European Aspen (*Populus tremula* L., Salicaceae). *Genetics* **2005**, *169*, 945–953.
89. Mosca, E.; Eckert, A.J.; Liechty, J.D.; Wegrzyn, J.L.; la Porta, N.; Vendramin, G.G.; Neale, D.B. Contrasting patterns of nucleotide diversity for four conifers of Alpine European forests. *Evol. Appl.* **2012**, *5*, 762–775.
90. Gailing, O.; Vornam, B.; Leinemann, L.; Finkeldey, R. Genetic and genomic approaches to assess adaptive genetic variation in plants: Forest trees as a model. *Physiol. Plant.* **2009**, *137*, 509–519.
91. Derory, J.; Scotti-Saintagne, C.; Bertocchi, E.; le Dantec, L.; Graignic, N.; Jauffres, A.; Casasoli, M.; Chancerel, E.; Bodénès, C.; Alberto, F.; *et al.* Contrasting relations between diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. *Heredity* **2010**, *105*, 401–411.
92. Eckert, A.J.; Wegrzyn, J.L.; Pande, B.; Jermstad, K.D.; Lee, J.M.; Liechty, J.D.; Tearse, B.R.; Krutovsky, K.V.; Neale, D.B. Multilocus patterns of nucleotide diversity and divergence reveal positive selection at candidate genes related to cold hardiness in coastal Douglas Fir (*Pseudotsuga menziesii* var. *menziesii*). *Genetics* **2009**, *183*, 289–298.
93. Hamilton, J.A.; Lexer, C.; Aitken, S.N. Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* × *P. glauca*). *Mol. Ecol.* **2013**, *22*, 827–841.
94. De la Torre, A.R.; Wang, T.; Jaquish, B.; Aitken, S.N. Adaptation and exogenous selection in a *Picea glauca* × *Picea engelmannii* hybrid zone: Implications for forest management under climate change. *New Phytol.* **2014**, *201*, 687–699.