The effects of clonal forestry on genetic diversity in wild and domesticated stands of forest trees

Pär K. Ingvarsson a, b and Helena Dahlberg a

aUmeå Plant Science Centre, Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden; bDepartment of Plant Biology, Swedish University of Agricultural Sciences, Uppsala BioCentre, Uppsala, Sweden

ABSTRACT

The level of genetic diversity maintained in a population is determined by the combined action of mutation, gene flow, genetic drift and selection. Forest tree breeding is a relatively recent phenomenon compared to most crop species and the material that is being deployed is, genetically, often very similar to wild-growing populations. The introduction of vegetative propagation has been hailed as a more efficient and flexible method than seed orchards to rapidly realize breeding progress and to adapt material to future climate change. What remains unclear is how a large deployment of vegetatively propagated material may affect the patterns of genetic diversity within and among forest stands. Here we review what is currently known about genetic diversity in managed and natural forest stands and specifically address the impacts of clonal forestry. To assess this we develop a quantitative model to describe the consequences of clone deployment on genetic and genotypic diversity in Swedish forests. We conclude with some remarks specific to Swedish conditions, likely scenarios for clonal deployment and finally some suggestions for future research priorities.

Introduction

The level of genetic diversity that is maintained in a population is determined by a complex interplay between different evolutionary forces. Maintaining sufficient levels of genetic diversity is crucial for the persistence of most species, for instance to preserve the ability to adapt to changing environmental conditions (Gregorius 1991). Levels of heterozygosity are directly linked to fitness at an individual and this leads to the expectation that at the population level, mean fitness and level of heterozygosity are correlated (Reed and Frankham 2003). Because the maintenance of genetic diversity is inversely related to population size, small populations are more likely to experience negative consequences of reduced genetic diversity. For instance, deleterious, recessive alleles are more likely to becoming common enough to be expressed in homozygous form in small populations that ultimately leads to a reduction in viability and/or fecundity (Charlesworth and Willis 2009). However, gene flow among populations counters the loss of variation due to genetic drift by (re)introducing novel genetic variation. By introducing vegetative propagation as a way of efficiently and rapidly introduce new variants or adapt material for different climates, the natural processes affecting genetic diversity are to a large degree bypassed. It remains unclear how the deployment of vegetatively propagated material may affect genetic diversity within and among forest stands and to assess the possible consequences of clonal deployment we first need understand current patterns of genetic diversity in natural and managed stands as well as how these might change with the introduction of clonal forestry.

Genetic drift and gene flow

Many forest trees have very wide geographic distributions and large population sizes; they are also often highly outcrossing, have long generation times and, in many cases, have extensive gene flow over large geographic distances through both pollen and seeds. Taken together, these life-history traits promote the maintenance of abundant levels of genetic variation in most forest tree populations (Hamrick and Godt 1996). Data compiled from a large number of plant species also confirm these expectations; long-lived, woody perennial plants with wide geographic ranges, particularly those with a boreal temperate distribution, generally have high levels of genetic diversity (heterozygosity) and a large fraction of polymorphic loci with many alleles per locus, i.e. allelic richness (Hamrick and Godt 1996).

Although genetic diversity is ultimately determined by a balance of mutation, gene flow and genetic drift, these processes take a very long time to reach equilibrium. Historical effects, such as past changes in population sizes and levels of gene flow, can have long-standing effects on present-day genetic diversity. For instance, population size fluctuations...
associated with the last (and even preceding) glaciations have been shown to influence current day genetic diversity and have also been shown to have more subtle effects, such as changing the expected frequency distribution of segregating genetic variants (i.e. where the two alleles at a locus are different). For instance, in a rapidly expanding population, the loss of alleles by genetic drift will be reduced and many more novel mutations are expected to survive elimination due to stochastic reasons. This will result in an excess of rare alleles compared to equilibrium expectations and can easily be detected by tallying the frequencies of all segregating alleles in a population (Keinan and Clark 2012).

Finally, it is important to distinguish between putatively neutral genetic variation, which is not affected by natural selection, and genetic variation in traits that confer adaptations of some sort. The theory discussed above largely assumes that the effects of natural selection are negligible and thus only relates to neutral genetic variation. Genetic structuring of variation underlying adaptive traits can be substantially different from neutral genetic variation, and care must be taken not to confuse the two (Savolainen et al. 2013). Nevertheless, understanding how neutral genetic variation is structured is important as it provides important information on the (combined) effects of genetic drift and gene flow and shapes the background genetic variation upon which natural selection act.

**Natural selection and local adaptation**

It is well known that natural selection can act to differentiate populations for adaptive traits, resulting in local adaptation, even in the face of substantial migration. A classic example in forest trees is the adaptive response to the steep latitudinal gradient in the length of the growing season that characterizes northern environments. Forest trees often show latitudinal clines for important phenological traits, such as the initiation or cessation of growth, onset of flowering or the development of frost tolerance. This adaptation to local growing conditions occurs despite on-going gene flow that continuously introduces potentially maladaptive genetic variation (Kremer et al. 2012). High levels of gene flow among populations are demonstrated by levels of genetic differentiation at neutral molecular markers, which is often low in forest trees. Since populations are nevertheless able to adapt to local environments in the face of extensive gene flow, natural selection acting on traits that confer local adaptation must be strong. In fact, these observations can be used to devise a method for identifying traits that are under diversifying selection. This method works by first estimating genetic differentiation at neutral loci, measured as $F_{ST}$, which is a function of background levels of gene flow and genetic drift (as described in the preceding section). These estimates are then compared with genetic differentiation at quantitative traits affected by natural selection. Genetic differentiation at quantitative traits is measured as $Q_{ST}$, which is simply the fraction of total variation in a trait that is distributed among populations (Wright 1951; Spitze 1993). In the absence of natural selection, $F_{ST} = Q_{ST}$. Deviations from this expectation can therefore be used to test for the action of natural selection on a quantitative trait (Leinonen et al. 2013; Savolainen et al. 2013).

Even if neutral markers display low genetic differentiation in most forest trees, consistent with extensive gene flow among populations, less is known about patterns of genetic differentiation at loci directly involved in controlling quantitative traits of adaptive value (quantitative trait loci or QTLs) that are affected by natural selection. Recent theoretical models suggest that genetic differentiation at QTLs is better predicted by $F_{ST}$ than $Q_{ST}$, even in cases where quantitative traits display large differences among populations. The reason for low genetic differentiation at the underlying QTLs (similar to $F_{ST}$ at neutral loci), even for traits that show high genetic differentiation (high $Q_{ST}$) is that selection generates covariances between individual QTLs that act to reinforce the total effect on a quantitative trait. This means that even if individual QTLs display relatively low genetic differentiation the combined effect of multiple QTLs changing in parallel leads to strong genetic differentiation at the phenotypic level (Latta 2013).
So, even if adaptive traits display large differences among populations, genetic differentiation at the underlying loci is likely much more similar to what we expect for neutral genetic variation. This observation only strengthens the value of quantifying genetic structure at neutral loci, especially as we in most cases have very poor knowledge of the actual genetic variants that are involved in shaping adaptive traits (i.e. the actual QTLs).

Causes and consequences of genetic diversity

Genetic diversity in forest trees

Forest trees, as a group, share many life-history characteristics that influence how genetic variation is structured within and among populations. As reviewed by Hamrick and Godt (1996), long-lived woody perennials maintain higher levels of genetic diversity within populations compared to other plant types (Figure 1(a)). This is perhaps not all that surprising, given that many forest tree species have large population sizes and cover extensive geographic areas. Also, forest trees are long lived, meaning that genetic variation will be maintained for decades or centuries just by the longevity of individuals, thus slowing the erosion of genetic variation by genetic drift. While most forest trees have large population sizes with largely outcrossing mating systems, they also tend to maintain a large genetic load. Genetic load here refers to the existence of partially or fully recessive deleterious alleles in a population. These harmful alleles are more likely to be exposed to selection in crosses between closely related individuals, increasing the probability to form homozygotes (both alleles at a locus are the same recessive), resulting in a phenomenon known as inbreeding depression. Thus, one consequence of a reduction in population size is the increased likelihood of accumulation of deleterious alleles in a population due to genetic drift. Also, small population size and the concomitant reduction in genetic variation increase the possibility for consanguineous matings that result in the manifestation of inbreeding depression. Furthermore, data suggest an overall positive relationship between population size, genetic diversity and fitness in most plant species (Leimu et al. 2006). This serves as an important reminder that genetic variation not only may act as a buffer to changing environmental conditions, but also plays an important role in determining fitness in natural populations.

Forest trees are often wind pollinated and many have adaptations that allow seed dispersal to occur over great distances. This can result in lower levels of genetic differentiation among populations compared to other plant species, as seen in long-lived perennials (Figure 1(b)). On average, up to (and sometimes exceeding) 90% of all genetic variation in widespread forest trees can be found within local populations, with the remaining variation distributed among populations across its range (Figure 1). As described above, the observation of low genetic differentiation among populations is largely based on data for neutral markers. Most forest tree species display extensive local adaptation despite the recent introduction of maladapted alleles through long-distance gene flow from surrounding areas, and it remains to be determined how genetic variation is structured at underlying loci controlling adaptive variation. The extensive gene flow seen in most forest trees works to reintroduce genetic variation that would potentially be lost through genetic drift and other processes, thus ensuring that levels of standing variation in natural populations is usually high. The high levels of genetic variation would work as a buffer to both abiotic or biotic environmental changes, and can perhaps be one reason for the success (in evolutionary terms) of many forest trees.

Effects of breeding and modern forestry practices on genetic diversity

There are several steps in the breeding process that can act to reduce genetic diversity. Phenotypic selection in the breeding cycle and selection during seed and seedling production can all have effects on the genetic variation present in a species. Breeding by necessity entails selecting superior individuals for future reproduction and it is well known from many crop species that such selection have the possibility to induce genetic bottlenecks, both in the traits that are direct targets of selection and across the entire genome of the organisms undergoing domestication (Doebly 1989). In crops, that in many cases have been under domestication for up to 10,000 years, the most important factors determining the severity of loss of genetic diversity during domestication is the duration of the domestication period and how large the population is during this period, collectively known as the domestication bottleneck (Eyre-Walker et al. 1998). An exceptionally strong domestication bottleneck is expected to leave little variation remaining in the genome, making it difficult to distinguish neutral loci from loci that have been the target of selection during the domestication process (Wright et al. 2005; Doebly et al. 2006).

Most forest tree species are either undomesticated or in the very early stages of domestication (Neale 2007; Harfouch et al. 2012). Studies in crop species have, as outlined above, highlighted the possibilities for loss of genetic variation during domestication and in most forest tree breeding operations care has consequently been taken to ensure that breeding populations are large enough to minimize loss of genetic diversity (Libby 1973; Harfouch e al. 2012).

In current forestry practices, the most prominent drawback highlighted by Hallsby (2013) is an increased cost of planting, compared to natural reforestation or sowing. However, further objections to planting, and modern forestry practices at large, is a fear of reduced genetic diversity which in turn can result in reduced biodiversity and lower abilities for populations to respond to natural selection (Black-Samuelsen 2012; Black-Samuelsen et al. 2017). For instance, if genetic variation is reduced stands could suffer reduced abilities to adapt to novel pests or pathogens or future climate change (Aitken et al. 2008). Hallsby (2013) argued that genetic variation in planted stands are likely kept at comparable levels to naturally regenerated stands since the plus trees (phenotypically superior trees chosen for breeding) used in modern seed orchards are derived from a wide geographic range.
However, a recent report from Skogsstyrelsen have suggested that effects of tree breeding on genetic diversity in is an area where more research is needed (Black-Samuelsson et al. 2017). Consequently, if clonal forestry should be introduced in Swedish forestry at larger scales than today, managing genetic diversity will be an even more important topic to address.

As noted above, high levels of pollen gene flow characterize most forest tree species and this is often manifests itself as high levels of pollen contamination in seed orchard. A review study done on six different conifer species gave an average contamination rate in seed orchards of 45% (Adams and Burczyk 2000). Pollen contamination in seed orchards thus results in an uneven quality of seeds between sites and years, since it is expected to lower the genetic gain by introducing alleles from surrounding forests (Adams and Burczyk 2000; Kang et al. 2001). If, on the other hand, gene flow from seed orchards into natural stands is large, it can reduce genetic diversity of surrounding populations and the offspring might be less adapted to local environments since genotypes from seed orchards are often transplanted from elsewhere (Adams and Burczyk 2000).

Williams and Savolainen (1996) argue that few organisms carry as many lethal alleles as conifers do and because of this, conifers are strongly affected by inbreeding depression and selling or strong inbreeding would lead to large losses of offspring and to low adult fecundity. There have been reports of total seed crop failure after selfing in both Norway spruce (Picea abies) and Scots pine (Pinus sylvestris) (Tigerstedt 1973). Eriksson et al. (1973) compared selfed and open pollinated progenies from four different mother trees in an experimental plantation of Norway spruce and found lower survival rates among the selfed offspring during germination, their first summer and when moved to the plantation site. There was also a reduction in average tree height, average trunk diameter and average trunk volume in selfed progeny, compared to open pollinated offspring, in adult trees (61 years of age). A nursery and field trial spanning 10 years monitored both outcrossed and selfed families from three natural Norway spruce populations. The selfed families suffered higher mortality, slower germination, earlier bud set and a shorter shoot growth period. Shoot growth rate, height and diameter growth were all reduced and inbreeding depression was significantly different from zero in all traits measured (Skroppa 1996). Kärkkäinen et al. (1996) showed that the number of lethal equivalents in Scots pine (P. sylvestris) varies between trees originating from a northern and a southern population (4.5 vs 6.9, respectively). These differences likely reflected historical differences in selfing rates in the two populations, where the northern population displayed higher selfing rates and thus possibly partial purging of deleterious alleles due to higher levels of inbreeding (Kärkkäinen et al. 1996). A more recent study was done in two Douglas-fir (Pseudotsuga menziesii) test-plantations. From a nine-parent founder population multiple families were created, among them nine selfed families. When the populations reached age 26, the survival of the inbred families was estimated to be 10% and 16% relative to outbred families in the two locations. Inbreeding depression for survival was thus calculated (as in Burdon and Russell 1998) to be over 80% at age 26 in selfed trees (Stoehr et al. 2015).

Direct and indirect effects of genetic diversity

The direct effects of genetic diversity on a population are comparably well understood. As noted above, one consequence of a reduction in population size and a concomitant reduction in genetic diversity is that it can trigger the accumulation of deleterious alleles, inbreeding and inbreeding depression. Since both the accumulation of deleterious alleles and inbreeding depression have detrimental effects on survival and reproduction, such effects could lead to a negative feedback loop where loss of genetic variation leads to a reduction in survival or reproductive potential, which further reduces population size. This negative feedback loop is known as a “mutational meltdown” (Lynch et al. 1995) and results in progressive deterioration of fitness and the eventual extinction of the population unless genetic variation can be replenished. Another possible consequence of low levels of genetic variation is that it may reduce a population’s ability to respond to changes in the environment. Such environmental changes can include either abiotic factors, such as changes in climatic conditions, or biotic factors, for instance the introduction of novel pests or pathogens (Reusch et al. 2005). The rather unpredictable nature of genotype-environment interactions makes it hard to predict how populations will respond to future environmental change, as even genetically uniform populations may yield variable phenotypic end products. These problems are compounded in many commercial forest tree species, where long rotation times make predicting effects of future environmental changes very uncertain. Finally, low genetic diversity can also result in reduced productivity, as genetically homogeneous groups of individuals are expected to suffer from more intense competitive interactions than do genetically diverse groups (Boyden et al. 2008). While mutualistic melt-downs are unlikely in forest trees, due to their large and stable population sizes, a reduced ability to adapt to environmental change could become an issue if the rate of change of the environment is rapid. This is an area that has only recently received attention from forest geneticists (Aitken and Whitlock 2013).

Moreover, effects of genetic diversity can have far-reaching impacts beyond the population of the species in question. Studies over the last decade have shown that genetic variation within dominant keystone species can have consequences not only for the species itself but also for both associated communities and even for whole ecosystems (Fritz and Price 1988; Aguilar and Boecklen 1992; Messina et al. 1996; Whitham et al. 2003). This effect is known as an “extended phenotype” which suggest that effects of phenotypes are not limited to the biological processes of the organisms expressing the phenotype but instead include all effects it may have on the environment. A nice illustration of these points is provided by an experiment conducted by Johnson and Agrawal (2005) that manipulated patches of Evening Primrose (Oenothera biennis) to contain one, four or eight
unique plant genotypes. They hypothesized that high genetic diversity within a plant species offers a greater number of different niches than a population with low diversity, leading to an altered composition and greater abundance of arthropods with increasing diversity. Colonizing arthropods were surveyed on the plants and for each plant the arthropods were identified and counted. The results showed a linear increase of cumulative arthropod richness with increasing numbers of plant genotypes from 1 to 4 to 8. There were 18% more arthropods on patches with eight genotypes compared to patches with only one genotype and the pattern was consistent across the whole growing season (Johnson and Agrawal 2005). When looking at the total abundance of arthropods, genotype diversity had no effect but when divided into herbivores, omnivores and predators both omnivores and predators showed an increase in abundance with an increase in genotype diversity. The fitness of plants also increased with genotype diversity, with mean genotype fitness being 27% higher in patches with more than one genotype (Johnson and Agrawal 2005). Similarly, a study performed in natural hybrid zone of two cottonwoods species (Salicaceae: Populus) found that almost 60% of the variation in arthropod diversity among different trees could be explained by plant genetic diversity (Wimp et al. 2004). In a two-year survey of European aspen (Populus tremula), natural variation in functional leaf traits and herbivory was studied. The results showed that arthropod species richness was moderately inheritable among different clones and genetic variation structured arthropod communities in young trees. Leaf rust fungus on the other hand showed a strong heritability at both the population and clone level (Robinson et al. 2012).

**Consequences of clone deployment on genetic and genotypic diversity**

The probability of loss or establishment of genetic variants is determined by the effective size of the population, \( N_e \). The effective population size is a core concept in population genetics that aims to capture the effects of random sampling of allele frequencies in a finite population and relate these to evolutionary change in the population. These random sampling effects are usually referred to as “genetic drift”. Together with mutation rate, effective population size determines the level of genetic diversity (largely neutral diversity that is under little or no influence of natural selection) maintained in a population at equilibrium. The effective population size is, in turn, affected by a number of processes, such as variation in sex ratio or offspring numbers, fluctuations in population size over time or inbreeding (Charlesworth 2009). The concept of effective population size is useful because it captures complex demographic patterns and reduces them to a standard population genetic model that is more amenable for analyses, thereby highlighting how these processes affect rates of genetic drift. In most cases, these processes act to reduce the effective population size below the actual census size of the population, so that levels of genetic diversity maintained are usually far less than what could be expected based on census size. Even if the expected \( N_e \) is substantially lower than the census size, it nevertheless scales well with actual population size, so comparisons can be made among populations of different census sizes.

When talking about genetic variation in populations it is also important to distinguish between genetic diversity and genotypic diversity. Genetic diversity refers to the amount of genetic variation present in a population and depends on the number and frequency of alleles that are segregating. Genotypic diversity, however, refers to the number of unique genotypes that are present. As an example, ten unique clones harbour the same amount of genetic diversity as all potential offspring generated when using these trees as parents. However, genotypic diversity will be lower if these ten trees are deployed as clones as opposed to planting offspring generated by randomly mating the ten clones. In the latter scenario, recombination associated with meiosis ensures that virtually all offspring will have unique genotypes, even when generated from such a limited pool of parents.

**The effects of clonal deployment on genetic diversity**

To get a more quantitative picture of how the deployment of clonal material affects levels of genetic diversity in a population, it is instructive to consider the following population genetic model. Consider a population consisting of \( N \) diploid individuals, consequently harbouring \( 2N \) alleles at a locus of interest. Together with some additional assumptions, such as constant population size, neutral genetic variation and discrete generations, this constitutes the standard Wright-Fisher model of population genetics (Hein et al. 2005). A key parameter in the Wright-Fisher model is the effective population size (\( N_e \)), which determines the rate of genetic drift and ultimately the amount of genetic variation that can be maintained. There are many different effective population sizes that can be defined, depending on whether we are interested in, for instance, the variance of allele frequency change, change in inbreeding or decline in heterozygosity (Ewens 1982). When the interest is the amount of genetic diversity that is maintained in a population, a suitable effective population size is the “coalescent effective population size” (Sjödin et al. 2005), defined as:

\[
N_e = \frac{1}{2} \overline{\tau}
\]

where \( \overline{\tau} \) is the expected time it takes for two randomly sampled alleles to coalesce to a common ancestor. For a Wright-Fisher population, \( \overline{\tau} = 2N_e \) and the effective population size reduces to the census size of diploid individuals in the population (Balloux et al. 2003). The coalescent effective population size is also naturally related to the amount of genetic diversity that is maintained at mutation-drift equilibrium, such that,

\[
H_e = \frac{\theta}{1 + \theta}
\]

where \( \theta = 4N_e\mu \) and \( \mu \) is the mutation rate. There are several ways to estimate genetic diversity in natural populations, but a convenient measure of genetic variation at the nucleotide
level is nucleotide diversity (usually denoted $\pi$). The expected nucleotide diversity in a population is proportional to the effective population size ($E[\pi] = \theta = 4N_e \mu$) and independent of sample size (Li 1997). It is thus straightforward to assess how clonal deployment affects genetic diversity in a population by working out how the introduction of clones changes the average coalescence times of individuals. Balloux et al. (2003) showed how the effects on genetic diversity of replacing a fraction of individuals in a population by clones can be assessed by the effects on the coalescent effective population size. Assuming a large, randomly mating population (essentially a Wright-Fisher population as defined above), the coalescent time of two randomly sampled individuals is, as stated above, $2N_e$. Assume a fraction $c$ of these individuals are replaced by clonally replicated individuals, derived from $n$ unique progenitors that are equally represented in the population (so $n$ is the number of clones used, which for now are assumed to be sampled at random from the base population). The effect on genetic diversity is then simply the weighted average of the pairwise coalescence times of randomly sampled individuals. Sampling two clonally replicated individuals occurs with a probability $c^2$. With probability $1/n$ these individuals replicate of the same clone and coalesce occurs in the previous generation. With probability $(1-1/n)$ the clones are different and have a coalescence time of $2N_e$.

Two individuals regular individuals are sampled with probability $(1-c)^2$ and also have a coalescence time of $2N_e$. Likewise, a regular individual and a are selected with probability $2c(1-c)$ and also have a coalescence time of $2N_e$. These approximations are probably realistic if the trees used for clonal propagation are selected from a base population harboring most of the genetic diversity present in local populations, although this assumption can easily be relaxed if necessary. Putting all this together, we can derive the average coalescence time as a function of the number ($n$) and proportion ($c$) of clones used:

$$\bar{t}_c = 2N_e \left(1 - \frac{c^2}{n}\right) + \frac{c^2}{n}$$

(3)

Equation (3) can be used to calculate the number of clones needed to capture a given fraction of the original genetic diversity, by simply letting $c = 1$. As shown in Figure 2(a), 20 unique clonal progenitors are needed to effectively capture 95% of the genetic diversity present in the base population (see also Wakeley and Sargsyan 2009). To get an idea of the genetic diversity that is maintained (relative to a population without any planting of clones), we can evaluate Equation (3) for a range of values for $c$ and $n$. A few selected scenarios are depicted in Figure 2(B), where it is assumed that sites harbour 2000 trees. Figure 2(B) demonstrates that once a sufficient number of clones are used per site, most of the genetic diversity can be maintained, even if a large fraction of clonally replicated individuals is deployed. However, even if a relatively small sample (~20 individuals) is sufficient to capture most of the genetic diversity present in a population, the probability of losing an allele when sampling is very much dependent on the allele frequency in the source population (Figure 3). With modest sample sizes, most common alleles will be be included but rare alleles have appreciable probability of being lost even with quite large sample sizes (Gregorius 1980). These effects may seem somewhat contradictory but can be understood by the fact that genetic diversity is proportional to $2p(1-p)$, where $p$ is the allele frequency at locus $i$. So while rare alleles have appreciable probability of being lost when sampling a individuals (as shown in Figure 3), they will actually make rather small contributions to overall genetic diversity in the population.

**The effects of clonal deployment on genotypic diversity**

Deployment of clonal material is expected to affect not only genetic diversity but also genotypic diversity, that is the number of unique genotypes present. To assess the impact of clonal deployment on genotypic diversity we can apply a similar reasoning as for genetic diversity. The genotypic effective population size is slightly more complicated to derive, since we need to keep track of not only identity by decent but also genotypic diversity, that is the number of complete genotypes at a locus. Fortunately, Balloux et al. (2003) showed that in a randomly mating population the mean coalescence time of genotypes at a locus is approximately given by,

$$\bar{t}_g \approx 3N_e c$$

(4)

For a completely clonal population, the mean coalescence time at a locus is then simply a function of number of unique clones used (Balloux et al. 2003). With these approximations, together with the assumptions above of a clonal fraction $c$, a unique genotypes. If these 2000 individuals are replaced by unique clones used from $n$ unique clonal progenitors we get:

$$\bar{t}_g \approx 3N_e (1 - c^2) + \frac{c^2}{n} (1 - 3n + 3n^2)$$

(5)

The number of clones needed to capture even a fraction of the original genotypic diversity at a locus is substantially greater that what is needed to ensure an adequate coverage of genetic diversity (Figure 4(A)). Figure 4(B) depicts Equation (5) evaluated across a range of clonal proportions and numbers of clones used per site. It is immediately apparent that clonal deployment has a substantially greater effect on genotypic diversity than on genetic diversity (compare Figures 2(B) and 4(D)). As an example, assume locus that harbours 10 segregating single nucleotide polymorphisms (SNPs) that can combine to form $3^{10} = 59049$ possible multilocus genotypes. The maximum genotypic diversity in the population would consequently occur if all 2000 trees carried unique genotypes. If these 2000 individuals are replaced with clones derived from 20 or 200 unique trees, we could at best capture 1% or 10% of the original genotypic diversity, respectively. However, if the same site was replaced with sexually produced offspring from the same 20 individuals we would contain the same genotypic diversity present in the original population, if recombination rates are high enough within the locus to ensure that alleles are segregating independently. If recombination rates are more limited, the number of multilocus genotypes will be low than the maximum but nevertheless greater than if clonally
propagated trees were used. Bulked seed propagation from a limited set of selected clones will thus provide a middle ground that should ensure moderate levels of both genetic and genotypic diversity, while still providing benefits in terms of the increased productivity associated with selection of clonal material. Table 1 contrasts genetic and genotypic diversity for a number of different origins of newly planted material both on a local scale (the level of a single stand) and on a regional level.

**Outlines of Swedish scenarios for clonal forestry**

To evaluate the genetic consequences of introducing different forms of clonal forestry, a number of scenarios have to be evaluated. In principle, there are three factors that determine the effect of clone deployment on genetic diversity. First and foremost is how well the parents from which any clones are derived capture the genetic diversity present in Swedish forests. The second factor is how many clones and how many ramets of these clones will ultimately be deployed. Finally, the total area planted with clone material is of importance.

The sample size of trees needed to effectively capture more than 95% of the genetic variation in a source population depends on the measure of genetic diversity used. For instance, for measures like the number of alleles or allelic richness the sample size needed to ensure a 95% or greater capture of genetic diversity is on the order of a few hundred trees (Gapare et al. 2008; Bashalkhanov et al. 2009). However, for nucleotide diversity the same number is only on the order of 20 individuals (Figure 2(a)). The discrepancy between the different measurements of genetic diversity stems from the fact that most genetic variants are rare. Thus, large sample sizes are required to capture most rare variants. However, rare variants contribute little to heterozygosity or nucleotide diversity where common alleles are much more important and smaller sample sizes are consequently needed to capture these more common alleles. The core breeding population for Norway spruce used in Swedish tree breeding consists of more than a thousand plus trees, and care has been taken when these were sampled to cover most areas to increase the likelihood of capturing any locally abundant, but globally rare variation. Based on the sampling scheme and size of the base population, these trees likely capture the majority of genetic variation that is present in Swedish spruce forests (Androsiuk et al. 2013).

Once a sufficient number of clones are used to re-plant sites, loss of genetic diversity is going to be minimal (Figure 2(b)). However, genotypic diversity will be substantially reduced even if large numbers of unique clones are used. Since it is the genetic diversity in the breeding populations from which clones are derived that establishes the upper limit of the diversity that will be present in clonal material, care must be taken to ensure that this population conserves genetic diversity over time. In the Swedish breeding programme, genetic diversity is conserved by structuring the population in subpopulations and using a strategy for selecting and mating, providing almost equal contributions to all new generations from the original founder trees (Rosvall 2011). Although the breeding population provides the parents for all clones used, the actual amount of genetic variation captured by clone plantings will be less than what is present in the source (breeding) population (Figure 2(a,b)), modest numbers of clones are needed to ensure that most genetic diversity is preserved. However, even if effects of clone plantings on genetic diversity are modest or even negligible, the effects on genotypic diversity are expected to be substantial (Figure 2(c)).
The analyses above have considered a natural population of spruce on a specific land area planted with various fractions and numbers of clones originating from the breeding population, itself being a sample of the natural population. Under Swedish conditions, there is in addition a large land area with naturally regenerating spruce that will maintain the genetic and genotypic diversity of the Swedish spruce population, even with extensive application of clonal forestry. With this in mind, what is most important is to find out appropriate genotypic diversity to minimize risks, i.e. number of clones per stand to maintain adaptability to environmental changes, as well as resistance to pests and deceases.

Figure 3. The probability of losing an allele as a function of sample size for different allele frequencies in the source population. Note that the x-axis is shown in log-scale.

Figure 4. (A) Fraction of genotypic diversity maintained as a function of number of clones deployed per site; and (B) Genotypic diversity maintained relative to what would be maintained in a situation with no clonal plantings. Other parameters as in Figure 2.
Conclusions and future research priorities

In order to fully understand how deployment of clone material will affect the genetic diversity and structure of Swedish forests, it is important to establish “baseline” levels that represent naturally occurring levels of genetic diversity. Such baseline levels of genetic diversity in older forests that pre-date the deployment of improved material derived from the Swedish breeding programmes are crucial when assessing if and how current forest practices alter genetic variation both on a local and regional scale. Such studies should ideally assess a large number of populations across as many breeding zones as possible, to provide an unbiased view of the genetic diversity present in Swedish forests. Also, such studies should monitor both genetic and genotypic diversity since reductions in either genetic or genotypic diversity are known to have detrimental effects, such as increased competition, increased susceptibility to pests and pathogens and lower productivity (reviewed in Jump et al. 2009).

Acknowledgements

Part of this research has been supported through the project “Genomic breeding of Norway spruce for new bio-products” funded by the Swedish Foundation for Strategic Research (SSF).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Stiftelsen för Strategisk Forsknings [Genomic breeding of Norway spruce for new bio-prod].

References

Adams WT, Burczyk J. 2000. Magnitude and implications of gene flow in gene conservation reserves. In Forest Conservation Genetics: Principles and Practices. Young, A., Boshier, D. and Boyle, T. (eds). CSIRO Publishing. Collingwood, Victoria, Australia; pp. 215–224.
Aguilar JM, Boecklen WJ. 1992. Patterns of herbivory in the Quercus grisea × Quercus gambelii species complex. Oikos. 64:498–504.
Aitken SN, Whitlock MC. 2013. Assisted gene flow to facilitate local adaptation to climate change. Annu Rev Ecol. 44:367–388.
Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S. 2008. Adaptation, migration or extinction: climate change outcomes for tree populations. Evol Appl. 1:95–111.
Androsiuk P, Shimoa A, Westin J, Lindgren D, Fries A, Wang X-R. 2013. Genetic status of Norway spruce (Picea abies) breeding populations for Northern Sweden. Silvae Genet. 62:127–136.
Balloux F, Lehmann L, de Meuûs T. 2003. The population genetics of clonal and partially clonal diploids. Genetics. 164:1635–1644.
Bashalkhanov S, Pandey M, Rajora OP. 2009. A simple method for estimating genetic diversity in large populations from finite sample sizes. BMC Genet. 10:84.
Black-Samuelsson S. 2012. The state of forest genetic resources in Sweden - report to FAO. Skogssstyrelsen, rapport 12/2012.
Black-Samuelsson S, Bergquist J, Ugglia C. 2017. Skogsträdens genetisk mångfald: status och åtgärdsbehov. Skogssstyrelsen, rapport 7/2107.
Boydén S, Binkley D, Stape JL. 2008. Competition among eucalyptus trees depends on genetic variation and resource supply. Ecolgy. 89:2850–2859.
Burdon RD, Russell JH. 1998. Inbreeding depression in selfing experiments: statistical issues. For Genet. 5:179–189.
Charlesworth B. 2009. Effective population size and patterns of molecular evolution and variation. Nat Genet. 10:195–205.
Charlesworth D, Willis JH. 2009. The genetics of inbreeding depression. Nat Rev Genet. 10:783–796.
Doebely J. 1989. Isozymic evidence and the evolution of crop plants. In: Isozymes in plant biology. Dordrecht: Springer Netherlands; p. 165–191.
Doebely JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. Cell. 127:1309–1321.
Eriksson G, Schelander B, Åkebrand V. 1973. Inbreeding depression in an old experimental plantation of Picea abies. Hereditas. 73:185–193.
Ewens WJ. 1982. On the concept of the effective population size. Theor Popul Biol. 21:373–378.
Eyre-Walker A, Gaut RL, Feldman ML, Gaut BS. 1998. Investigation of the bottleneck leading to the domestication of maize. Proc Natl Acad Sci USA. 95:4441–4446.
Fritz RS, Price PW. 1988. Genetic variation among plants and insect community structure: willows and sawflies. Ecology. 69:845–856.
Gapare WJ, Yanchuk AD, Aitken SN. 2008. Optimal sampling strategies for capture of genetic diversity differ between core and peripheral populations of Picea sitchensis (Bong.) Carr. Conserv Genet. 9:411–418.
Gregorius HR. 1980. The probability of losing an allele when diploid genotypes are sampled. Biometrics. 36:643–652.
Gregorius HR. 1991. Gene conservation and the preservation of adaptability. In Species conservation: a population-biological approach. Basel: Birkhäuser; p. 31–47.
Hallsby G. 2013. Planting av barrträd. Skogsskötselserien nr 3, Skogssstyrelsen.
Hamrick JL, Godt MJW. 1996. Effects of life history traits on genetic diversity in plant species. Philos Trans Roy Soc B Biol Sci. 351:1291–1298.
Harfouche A, Meilan R, Kirst M, Morgante M, Boerjan W, Sabatti M, Scarascia Mugnozza G. 2012. Accelerating the domestication of forest trees in a changing world. Trends Plant Sci. 17:64–72.
Hein J, Schierup MH, Wiuf C. 2005. Gene genealogies, variation and evolution: a primer in coalescent theory. Oxford: Oxford University Press.
Johnson MT, Agrawal AA. 2005. Plant genotype and environment interact to shape a diverse arthropod community on evening primrose (Oenothera biennis). Ecology. 86:874–885.
Jump AS, Marchant R, Penuelas J. 2009. Environmental change and the option value of genetic diversity. Trends Plant Sci. 14:51–58.
Kang KS, Lindgren D, Mullin TJ. 2001. Prediction of genetic gain and gene diversity in seed orchard crops under alternative management strategies. Theor Appl Genet. 103:1099–1107.
Kåreklånsen K, Koski V, Savolainen O. 1996. Geographical variation in the breeding depression of Scots pine. Evolution. 50:111–119.
Keinan A, Clark AG. 2012. Recent explosive human population growth has resulted in an excess of rare genetic variants. Science. 336:740–743.
Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R, Brindel JR, Gomulkiewicz R, Klein EK, Ritland K, et al. 2012. Long-distance gene flow and adaptation of forest trees to rapid climate change. Ecol Lett. 15:378–392.
Latta RG. 1998. Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. Am Nat. 151:283–292.

Table 1. Effects of new material from different origins on genetic and genotypic diversity.

| Origin of new material | Local | Regional | Local |
|------------------------|-------|----------|-------|
| Natural regeneration   | low   | high     | high  |
| Seed orchards          | high  | high     | medium|
| Few unique clones      | low-medium | low | low |
| Many unique clones     | medium-high | medium-high | low-medium |

ORCID

Pär K. Ingvarsson http://orcid.org/0000-0001-9225-7521
Le Corre V, Kremer A. 2012. The genetic differentiation at quantitative trait loci under local adaptation. Mol Ecol. 21:1548–1566.
Leimu R, Mutikainen P, Koricheva J, Fischer M. 2006. How general are positive relationships between plant population size, fitness and genetic variation? J Ecol. 94:942–952.
Leinonen T, Mccairns RS, O'Hara RB, Merilä J. 2013. QST–FST comparisons: evolutionary and ecological insights from genomic heterogeneity. Nat Rev Genet. 14:179–190.
Li W-H. 1997. Molecular evolution. Sunderland (MA): Sinauer.
Libby WJ. 1973. Domestication strategies for forest trees. Can J For Res. 3:265–276.
Lynch M, Conery J, Burger R. 1995. Mutational meltdowns in sexual populations. Evolution. 49:1067–1080.
Messina FJ, Richards JH, McArthur ED. 1996. Variable responses of insects to hybrid versus parental sagebrush in common gardens. Oecologia. 107:513–521.
Neale DB. 2007. Genomics to tree breeding and forest health. Curr Opin Genet Dev. 17:539–544.
Reed DH, Frankham R. 2003. Correlation between fitness and genetic diversity. Conserv Biol. 17:230–237.
Reusch TBH, Ehlers A, Hämerli A, Worm B. 2005. Ecosystem recovery after climatic extremes enhanced by genotypic diversity. Proc Natl Acad Sci USA. 102:2826–2831.
Robinson KM, Ingvarsson PK, Jansson S, Albrechtsen BR, Kliebenstein DJ. 2012. Genetic variation in functional traits influences arthropod community composition in aspen (Populus tremula L.). PLoS One. 7:e37679.
Rosvall O, editor. 2011. Review of the Swedish tree breeding programme. Uppsala (Sweden): Skogforsk.
Savolainen O, Lascoux M, Merilä J. 2013. Ecological genomics of local adaptation. Nat Rev Genet. 14:807–820.
Sjödin P, Kaj I, Krone S, Lascoux M, Nordborg M. 2005. On the meaning and existence of an effective population size. Genetics. 169:1061–1070.
Skreppa T. 1996. Diallel crosses in Picea abies II: performance and inbreeding depression of selfed families. For Genet. 3:69–79.
Spitze K. 1993. Population structure in Daphnia obtusa: quantitative genetic and allozymic variation. Genetics. 135:367–374.
Stoehr M, Ott P, Woods J. 2015. Inbreeding in mid-rotation coastal Douglas-fir: implications for breeding. Ann For Sci. 72:195–204.
Tigerstedt PMA. 1973. Studies on isozyme variation in marginal and central populations of Picea abies. Hereditas. 75(1):47–60.
Wakeley J, Sargsyan O. 2009. Extensions of the coalescent effective population size. Genetics. 181:341–345.
Whitham TG, Young WP, Martinsen GD, Gehring CA, Schweitzer JA, Shuster SM, Wimp GM, Fischer DG, Bailey JK, Woolbright S, et al. 2003. Community and ecosystem genetics: a consequence of the extended phenotype. Ecology. 84:559–573.
Williams CG, Savolainen O. 1996. Inbreeding depression in conifers: implications for breeding strategy. For Sci. 42:102–117.
Wimp GM, Young WP, Woolbright SA, Martinsen GD, Keim P, Whitham TG. 2004. Conserving plant genetic diversity for dependent animal communities. Ecol Lett. 7:776–780.
Wright S. 1951. The genetical structure of populations. Ann Eugen. 15:323–354.
Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS. 2005. The effects of artificial selection on the maize genome. Science. 308:1310–1314.