Forest Ecology and Management 299 (2013) 6–13

Contents lists available at SciVerse ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

# Molecular tools and aspen management: A primer and prospectus

Karen E. Mock [a,*], Bryce A. Richardson [b], Paul G. Wolf [c]

[a] Ecology Center and Department of Wildland Resources, Utah State University, Logan, UT 84322-5230, United States
[b] USDA Forest Service, Rocky Mountain Research Station, Provo, UT 84606-1856, United States
[c] Ecology Center and Department of Biology, Utah State University, Logan, UT 84322-5305, United States

ARTICLE INFO

*Article history:*
Available online 29 November 2012

*Keywords:*
Aspen
*Populus tremuloides*
Molecular
Genomic
Ecology

ABSTRACT

Aspen (*Populus tremuloides*) is an iconic species in North American landscapes, highly valued for recreation, fiber, wildlife and livestock forage, carbon sequestration, biodiversity, and as a fuelbreak. However, there are rising concerns about the ability of aspen to persist in portions of its range, based on bioclimatic modeling, physiological thresholds and mortality surveys. Our ability to predict and mitigate aspen decline will depend on our understanding of the factors influencing aspen establishment and persistence. Genetic techniques are providing important insights about reproductive strategies, evolutionary and demographic histories, and adaptive capacity in western aspen, often with important and novel management implications. A suite of new genetic tools is also becoming available as a result of innovations in genomic sequencing technology, along with the availability of an annotated reference genome for *Populus* species. Here we present a synthesis of how these genetic tools have been, and could be, used to answer questions directly relevant to aspen ecology and management. We also present a brief description of the molecular tools themselves. Our goal is to provide an informational resource to forest managers about the utility of traditional and emerging genetic technologies, with specific relevance to aspen ecology and management.



## 1. Introduction

Aspen (*Populus tremuloides*) is a species of enormous ecological, economic, and social value. Ecologically, aspen has the largest distribution of any North American tree (Little, 1971), is a foundation species associated with high levels of biodiversity, and responsible for a tremendous amount of carbon flux at the continental scale (Hogg et al., 2002). Economically, aspen is an important source of fiber, wildlife and livestock forage. Socially, aspen is a particularly beloved species, gracing the names of municipalities, apparel, and institutions, and appearing frequently in art, jewelry, and even on a US postage stamp. Aspen is especially valued in western landscapes, where, as the dominant deciduous forest tree, it is associated with high levels of plant and wildlife diversity (DeByle and Winokur, 1985; Stohlgren et al., 1997; Mills et al., 2000; Simonson et al., 2001).

Like many temperate plants, aspen has undergone a dramatic and relatively recent range expansion since the last glacial maximum (ca. 18,000 year bp), and changing climates, management practices and management legacies continue to influence aspen distribution (Frey et al., 2004; Rehfeldt et al., 2009; Landhausser et al., 2010). At the continental scale, aspen distribution is likely limited by low temperatures at high latitudes and elevations (Turner et al., 2003) and by low soil moisture at lower latitudes and elevations (Hogg and Hurdle, 1995; Hogg et al., 2008; Anderegg et al., 2012) but pathogens, fire suppression, and ungulate herbivores can also influence patterns of regional and local persistence (Frey et al., 2004). Recent concerns about the sustainability of aspen in western and southern portions of the species range have emerged based on reports of large-scale mortality and climate projections (Worrall et al., 2008; Rehfeldt et al., 2009). Unfortunately, these are the same areas where aspen has some of the highest ecological and social values (DeByle and Winokur, 1985). As a result, forest managers are faced with a pressing need for effective conservation, restoration and mitigation strategies. Such strategies can be informed by the long and rich history of research in aspen regeneration ecology and physiology (Peterson and Peterson, 1993), but our understanding of these topics is rapidly changing and necessarily incomplete.

Scientific advancement typically occurs in conjunction with episodes of rapid technological progress, interspersed with periods of reflection and integration across disciplines. The advent and rapid evolution of molecular tools in recent decades has enabled new lines of scientific inquiry in many fields, including forest ecology. In many cases, these emerging tools now allow us to ask direct questions about molecular function, variability, and adaptation – research which previously relied on untested assumptions about

* Corresponding author. Tel.: +1 435 797 7870.
*E-mail address:* karen.mock@usu.edu (K.E. Mock).

0378-1127/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.foreco.2012.11.004

organismal processes and histories. A barrier to the application of these molecular tools to meaningful management questions is the limited communication between molecular scientists and managers and the highly technical nature of the genetic data. Our objective here is to provide an overview of current and emerging molecular tools in a manner accessible to scientists and managers from a broad range of backgrounds; particularly those involved in the study and management of aspen. It is our hope that such a synthesis will improve the communication between molecular researchers and applied scientists and managers, facilitating novel questions and integration across both fields of study.

## 2. Questions about aspen that can be addressed using molecular tools

We have organized our synthesis around four general types of analyses using molecular tools that can be used to understand aspen ecology, each with specific management applications: identification of individuals, assessment of landscape-scale gene flow patterns, assessment of adaptive variation, and identification of cytotype. For each general type of analysis, we explain why the information is relevant to aspen management, which molecular tools are appropriate, and how the results of such analyses are changing or expected to change our perspectives in aspen management.

### 2.1. Identification of individuals

#### 2.1.1. Clonality in aspen

Aspen is renowned for its ability to form clones (genets), which are clusters of stems (ramets) originating from the same seed. Stems in a clone are genetically identical, with the rare exception of mutations that occur within a clone during vegetative growth (i.e., somatic mutations) and which can accumulate with clonal age (Ally et al., 2010). Aspen produces viable seed, but seedling establishment in more xeric landscapes (e.g. the Intermountain West) is relatively rare and tends to follow fire (Jelinski and Cheliak, 1992; Kay, 1993; Romme et al., 1997; Tuskan et al., 2006). Thus, although aspen clones in eastern North America are typically quite small (Wyman et al., 2003), vegetative growth predominates in western landscapes and clones can become quite large (Kemperman and Barnes, 1976).

#### 2.1.2. Delineating clones

Aspen clones can easily be differentiated using molecular tools (see Microsatellites, below). In the absence of molecular tools, investigators have used morphological and phenological traits to distinguish, enumerate, and delineate individual clones (Kemperman and Barnes, 1976). These approaches can result in overestimation of clonal numbers when traits vary within a clone (e.g. due to stem age and environmental conditions), or underestimation when large numbers of small clones are present or when traits lack variability among clones. Phenotypic similarities are particularly problematic when adjacent clones are half-siblings, as might be expected around a fecund female clone. By contrast, appropriate molecular fingerprinting techniques can distinguish even closely related individuals. However, many genetic markers may be necessary to distinguish full siblings, which can be identical at many loci. Further, somatic mutations (discussed above) do occur in aspen, particularly in older clones, and care must be taken to avoid overestimating clonal richness (see Microsatellites, below).

#### 2.1.3. Management importance of clonal delineation

When individual clones can be accurately identified, a range of questions relevant to management becomes more tractable. For example, once clonal boundaries and sizes are described in an area, stand composition and age structure can be described in the context of clonal richness, local and regional genetic diversity can be accurately measured, and differences between clones (e.g. mortality, chemistry, physiology, and susceptibility to disease or herbivory) can be characterized. The importance of clonal boundaries in driving other boundaries (e.g. soil microbial and plant understory diversity) can also be assessed, and by looking at clonal sizes and juxtaposition, it may even be possible to use aspen genetic data to corroborate fire histories.

#### 2.1.4. How the molecular identity of clones has already changed perceptions of aspen ecology/management

In western landscapes, the obvious phenotypic (morphologic and phenologic) boundaries between large clones led to a general perception among foresters that aspen stands generally consisted of only one to a few clones, that these clones were quite ancient, and that sexual regeneration almost never occurred. This view underlies the primary regeneration strategy for western aspen: coppicing to stimulate suckering. Genetic work, initially using isozymes (Cheliak and Pitel, 1984; Hyun et al., 1987; Hipkins and Kitzmiller, 2004) and more recently using microsatellites (Mock et al., 2008; De Woody et al., 2009) suggests that clonal diversity within western aspen stands is far greater than previously recognized. These studies have also shown that clonal richness in western landscapes tends to be highly patchy, with most individual clones detected only once on a 50 m grid (Mock et al., 2008; De Woody et al., 2009). These findings, together with recent observations of seeding events following fires (Kay, 1993; Tuskan et al., 1996; Stevens et al., 1999; Turner et al., 2003; Elliott and Baker, 2004; Romme et al., 2005), suggest that establishment by recent seeding events may be more common than previously supposed. Elevated clonal richness and small clone size are not *de facto* proof of a recent seeding event; this would require clonal aging or information about landscape history. Nevertheless, management strategies for western aspen should consider genetic diversity and episodes of seedling recruitment (Long and Mock, in press).

#### 2.1.5. Future individual-level possibilities with molecular tools: clonal aging

Using microsatellites and DNA sequencing data (see DNA sequences, below), it should be possible to estimate clonal age based on the accumulation of mutations within clones. This approach has been developed in aspen by Ally et al. (2008, 2010), and although the method currently requires numerous simplifying assumptions and yields large confidence intervals, improvements could come with the availability of large-scale sequencing data. The ability to age individual clones could provide remarkable insights about clonal establishment, persistence, and resilience.

### 2.2. Assessment of landscape-scale gene flow patterns

#### 2.2.1. Gene flow in aspen

Because aspen has wind-dispersed pollen and seeds, movement of genes from one region to another is expected to be extensive. However, seedling recruitment (and hence gene flow) may be limited by biological, climatic or physical features at a landscape scale, and also by low fitness of propagules from source populations in novel environments. Geographic barriers in western North America may be particularly important to gene flow patterns in aspen, given the varied topography and steep ecological gradients present.

#### 2.2.2. The relevance of contemporary and historical gene flow patterns to aspen ecology and management

Information about contemporary gene flow patterns in aspen can be used to understand natural seed sources and sinks as well

as to identify areas where hybridization with other species (e.g. *P. grandidentata, P. alba*) occurs (e.g., Pauley, 1956; Spies and Barnes, 1982; Burczyk et al., 2006; Grivet et al., 2009). This information, along with information about clonal diversity, can help provide a context for regional and local patterns of mortality, succession, and phenotypic variation.

Information about historical gene flow patterns can be used to identify distinct lineages that have been isolated for long periods of time and may merit separate management (Moritz, 1994). Historical gene flow patterns can also be used to reconstruct range shifts following past climate changes (e.g. post-glacial expansion), which can provide a valuable context for current climate change modeling and future predictions about species distributions.

A recent study using microsatellites has described genetic structuring across the entire range of aspen (Callahan, 2012). In this study, two distinct genetic groups were discovered: one in the southwestern portion of the species range, south of the last glacial maximum, and one in the northern portion of the species range. Populations in the southwestern portion of the range showed less genetic diversity and more isolation than populations in the northern portion of the range, and apparently did not contribute to the northward range expansion following glacial retreat. These findings provide a framework for a variety of ecological questions in aspen, including the degree and mechanisms of reproductive isolation, the possibility of parallel genetic structuring in sympatric species, and the possibility of ecotypic variation.

#### 2.2.3. Using molecular tools to reconstruct demographic histories

Microsatellites and DNA sequencing data (see following sections) have been used in many other tree species to describe family structure and parentage (Dow and Ashley, 1996) and the geography of population expansions (Ingvarsson, 2008) and bottlenecks (Boys et al., 2005), which we will refer to collectively as "demographic histories". Although these approaches have not yet been used in North American aspen, such information could be valuable in understanding (1) the movement of genes via pollen vs. seed and (2) how short- and long-term climate histories are influencing aspen population expansion and contraction in particular regions. This information could be useful to forest managers in optimizing stand size, distribution, and genetic diversity for maximum resilience. This information could also be valuable in modeling aspen responses to past (and future) climate changes.

With the emergence of single nucleotide polymorphisms (SNPs; see below) as a common and accessible tool for *Populus* species (Keller et al., 2010; Geraldes et al., 2011; Harrison and Kidner, 2011; Marmiroli et al., 2011), and the development of new statistical approaches for estimating demographic parameters (Theunert et al., 2012), such analyses in aspen are becoming much more accessible. Demographic histories are particularly informative when overlain with adaptive traits such as phenology and herbivore defense, as has been done in European aspen (*P. tremula*) (Ingvarsson et al., 2006; Hall et al., 2007; Bernhardsson and Ingvarsson, 2012).

### 2.3. Assessment of adaptive variation

#### 2.3.1. Common garden studies

Adaptive variation is of particular interest to resource managers, and is a fundamental assumption underlying seed transfer zones and conservation guidelines (e.g., Randall and Berrang, 2002). Traditionally, local adaptation in plants has been assessed using common garden studies, reciprocal translocations, or provenance plantings in a range of ecological conditions to determine plant fitness and evolutionary potential in particular geographic zones or environments. Such common garden studies are difficult and expensive for long-lived tree species, although once established, they can yield valuable information in aspen. For example, Gray et al. (2011) have recently used reciprocal translocations to demonstrate adaptational lag in Canadian aspen, to develop a risk model linked to climate change projections, and to make spatially explicit recommendations about assisted migration in anticipation of future climate change patterns.

#### 2.3.2. Genomic technologies and adaptive variation

Although most population-level differences in genomes are neutral (of no adaptive significance), there is a small subset of such differences that do underlie adaptive divergence. Until very recently, the identification of such differences has been difficult, so inferences about adaptive divergence had to be made based on a combination of common garden studies, neutral markers (used to identify evolutionarily independent lineages or gene flow barriers), and habitat differences. However, improved technologies can now assay hundreds or thousands of molecular markers at a time in an individual or population (DNA sequencing and SNPs, below), providing the means to identify genomic sites which may have adaptive variation (Keller et al., 2011; Marroni et al., 2011). Using large-scale sequencing technologies and novel statistical approaches, it is possible to scan genomes for regions that show a statistical "signature" of being under natural selection (e.g. those that show ecological correlations not present in neutral markers) (Nielsen, 2005; Joost et al., 2007; Strasburg et al., 2012) or an association with a particular trait. Once these genomic regions are identified, they can be linked to specific genes or molecular pathways, depending on the genomic resources available for the species.

Another approach to studying adaptive variation and gene function is to survey the subset of genes that are expressed in an individual. This pool of expressed genes is called the transcriptome, and it can vary by individual, specific tissue, and environmental condition (see cDNA, below). For example, in response to low soil moisture, a particular aspen tree might up- or down-regulate the expression of a suite of genes, and the fitness of the individual in this environment may be a function of sequence variation and/or regulation of these genes. The identification of genes involved in responses to particular conditions can be a powerful way to understand biochemical pathways, ecological tradeoffs, and adaptation. One of the limiting steps in the analysis of transcriptome variance is the establishment of a "library" of genes that might be expressed in a particular species (i.e. a complete reference transcriptome). An effort is currently underway to characterize transcriptomes in several North American forest trees, including aspen (USFS Western Forest Transcriptome Survey) (Rai et al., submitted for publication). This will become one of the many publicly-available resources (see DNA sequences, below) that will enable genomics-based research in forest trees.

Of more interest to managers, once genes of potential adaptive significance are identified, alleles or combinations of alleles at multiple loci can be linked to specific types of ecological gradients (Guillot et al., 2009; Richardson et al., 2009; Holderegger et al., 2010). This type of information could revolutionize the way seed zones are delineated, could lead to the identification of specific genes involved in adaptation to specific conditions (e.g. drought) (Eveno et al., 2008; Eckert et al., 2010; Keller et al., 2012) and could be used to guide breeding programs for particular types of tree improvement.

### 2.4. Identification of cytotypes

#### 2.4.1. Polyploidy and adaptation

Polyploidy, or the existence of variants in a population with three or more complete copies of the genome per cell, is extremely common in plants (Wood et al., 2009). Cytotype, or the level of polyploidy, can be determined by counting chromosomes

(chromosome staining and microscopy) or by estimating the size and density of nuclei (e.g. Flow Cytometry, below). Different cytotypes (e.g. diploids, triploids, tetraploids) frequently have different morphological and physiological traits, and these may be favored in certain environments (Hegarty and Hiscock, 2007). Odd-numbered cytotypes (e.g. triploids; 3×) have greatly reduced fertility, so are unlikely to make a significant contribution to the local gene pool. However, these cytotypes may persist vegetatively, particularly if they have selective advantages in certain environments, and they may be a "bridge" to the formation of sexual tetraploid lineages (Ramsey and Schemske, 1998). Physiological studies have shown that polyploids can have greater xylem conduit size (Pockman and Sperry, 1997), and in some species may be more drought tolerant (Maherali et al., 2004). However, polyploids also have larger cell sizes, which may lead to structural vulnerabilities (Stebbins, 1938; Otto and Whitton, 2000). Understanding the frequency and geographic distribution of cytotypes can provide important management insights about optimal regeneration strategies, and cytotype information may be an important characteristic to control in ecological studies or provenance planting in aspen.

#### 2.4.2. Triploidy in Aspen

Using microsatellites and flow cytometry (see sections below), a recent study has documented that a large proportion of clones in western North America are triploid and that triploidy corresponds with climate variables (Mock et al., 2012). Furthermore, there is a strong tendency for the larger clones in some western study sites to be triploids (Mock et al., 2008). These patterns have a number of implications. First, both triploidy and clonality are elevated in western landscapes, and these traits may be synergistic and important in extending the geographic range of aspen. Second, triploids are expected to have reduced fertility, potentially reducing the proportion of viable seeds in areas dominated by large triploid clones. Finally, triploids may have different susceptibilities to climate change than diploids, resulting in mortality patterns that have a strong bias toward either triploids or diploids. The physiological implications of triploidy in aspen (growth rate, suckering rate, pathogen defense, drought tolerance) have yet to be characterized.

## 3. Molecular tools

### 3.1. Overview and history

People have been assessing and documenting genetic variation for centuries. Early approaches used visible characteristics (such as flower color or leaf shape). As soon as Mendel's work was widely known at the turn of the twentieth century, many researchers sought to uncover the inheritance patterns for these visible characters. By the 1950s, protein variation and secondary compounds could be examined in plants, and by the 1970s it was possible to explore DNA itself; the raw material of genes. With DNA-based analyses, it became possible to assess differences in many more portions of the genome, not just those genes coding for specific proteins and traits. Most of these non-coding regions were presumed to be selectively neutral (i.e. of no adaptive value), and allowed the direct assessment of relatedness, gene flow patterns, and phylogenetic relationships without the influence of selection. Recent technological developments make it possible to target specific genes over large numbers of individuals, to randomly sample large numbers of genes at the individual level, to sample multiple whole genomes within a species, or even to genetically characterize entire biological communities. Increasingly, particular questions can be answered using a variety of different tools, i.e. there is less "optimality" as more techniques emerge. Below, we outline some of the tools that have been and are being applied in the field, with the goal of providing an introduction to interested scientists and managers with limited backgrounds in genetics.

### 3.2. Isozymes

One of the first approaches applied to natural populations involved extracting metabolically active enzymes from living tissue, then separating the enzymes using an electric field in a gel. The gel can then be treated and stained so that different forms of an enzyme (isozymes) can be seen as dark bands. This enables variation in the underlying genes to be measured. Isozymes have been used on many hundreds of species of plants and animals, including aspen (Cheliak and Pitel, 1984; Gallo and Geburek, 1991; Lund et al., 1992; Liu and Furnier, 1993a,b; Hipkins and Kitzmiller, 2004; De Woody et al., 2009). Thus, isozyme analysis revealed underlying genetic variation that was hidden when viewing an organism with the naked eye.

### 3.3. DNA sequences

By the 1970s, techniques for examining genetic variation had moved from the biochemists' workbench to molecular biology laboratories. Geneticists were now examining variation at the DNA level. One of the first approaches to capturing variation in the sequence of nucleotides in a DNA chain was by cutting the DNA into pieces and looking for variation in the size and number of pieces. Most naturally-occurring bacteria produce enzymes to restrict the entry of foreign DNA by cutting at characteristic recognition sequences. The same enzyme will always cut DNA when it encounters that specific sequence. Variation in the DNA sequence therefore results in different cut patterns, or restriction fragment length polymorphisms (RFLPs). The restriction enzymes (endonucleases) can be purchased commercially. However, this technique only captures variation in the regions where DNA is cutting (the recognition sequence) and still much of the variation can be missed. By the end of the 1970s, techniques had been established for sequencing entire pieces of DNA. Now, rather than data consisting of complex lists of DNA fragments of different lengths, the string of nucleotides can be described simply by the sequence of the four bases. Different individuals are then compared by examining the base at a particular position in the sequence. The simplicity of DNA sequence data belies, and actually enables, some very complex and powerful analyses.

One of the initial breakthroughs in sequencing was polymerase chain reaction (PCR), which allows the synthesis of many thousands of copies of the same genomic region across many individuals, facilitating comparisons at the individual, population, and species levels. More recently, additional technological developments have made it possible to obtain many millions of short sequences from each individual. Until about 2007, DNA sequencing technology was progressing at about the same rate that computers get faster, doubling in speed about every 2 years. Then something drastic happened: a range of so-called next generation (or more appropriately "second generation") technologies became widely available. Sequence costs plummeted and the rate of data accumulation began overloading many databases. By 2012 the array of methods for generating DNA sequence data has become awe-inspiring. Several detailed reviews of these techniques are emerging (e.g. Niedringhaus et al., 2011; Liu et al., 2012). The scale and efficiency of data acquisition has also changed such that many molecular biology labs will now outsource work to specialist facilities. Researchers are no longer limited by the rate at which DNA sequences can be acquired; rather, progress is now limited by the ability to filter out relevant information from massive datasets.

The first tree genome to be fully sequenced was *Populus trichocarpa* (black cottonwood) (Tuskan et al., 2006). Because this tree is

a close relative of aspen, its genome can be used as an important reference for aspen genetic studies using a variety of different types of markers. For instance, if sequence variation is detected along a latitudinal gradient in aspen, the *P. trichocarpa* genome can be used to determine the genomic location and identity of that particular genetic site. Alternatively, sequence variation in known genes from *P. trichocarpa* can be characterized in aspen with respect to particular ecological gradients, as has been done for day length perception genes in *P. tremula* (European aspen) (Ingvarsson et al., 2008).

Once DNA sequence data (of any type) are available, they can easily be used by other researchers as a reference or in data mining studies. This is facilitated by the availability of public databases, and the requirements of some funding agencies and many journals that such data be made available. One of the most widely used depositories for DNA sequence data is the National Center for Biotechnology information (NCBI; http://www.ncbi.nlm.nih.gov/), which also provides an extensive range of tools for data mining and analysis. Many of the databases on this site are mirrored at the DNA Data Bank of Japan (DDBJ; www.ddbj.nig.ac.jp) and the European Molecular Biology Laboratory (EMBL; www.embl.org). The Dendrome Project (dendrome.ucdavis.edu) focuses specifically on genomic data for forest trees (Neale and Kremer, 2011), and includes links to databases on several species including aspen (http://aspendb.uga.edu/). These databases provide a wide range of information that can be easily accessed and then used on related species, or to find specific genes or genomic regions for new studies. In many ways the availability of public data and free analysis software is the most powerful set of tools in the molecular biology laboratory.

### 3.4. Amplified Fragment Length Polymorphisms (AFLPs)

For management purposes, one of the most useful applications of genetic tools is the comparison of populations. However, populations generally differ with respect to *frequencies* of certain alleles (variants at a locus) rather than being invariant (or "fixed") for alternative alleles. In order to compare the overall genetic divergence between populations, then, many different loci should be queried and allele frequency patterns compared. When populations show allele frequency differences at many different loci, this provides confidence that the differentiation is due to population phenomena (i.e. restricted gene flow and independent demographic or evolutionary histories), and is not an artifact of a single locus. Although this concept applies to any of the tools discussed here, a variety of specific marker systems have been developed to query many different loci simultaneously; one of the earliest was a technique called "amplified fragment length polymorphism" analysis, or AFLPs.

In AFLP analysis, DNA-cutting enzymes are used to cut the genomic DNA at specific places, leaving a range of fragment sizes (Vos et al., 1995; Mueller and Wolfenbarger, 1999). The lengths of these fragments will vary among individuals and populations, but there are too many fragments to analyze efficiently. The fragment pools for each individual are then subsampled by selectively amplifying only a set of particular fragments using PCR. These fragments can then be visualized, and the pattern of presence/absence is scored across each individual. The advantages of AFLP analysis are (a) that it can be applied to any organism with very little investment in marker development and (b) that it can provide information on hundreds of loci simultaneously. The disadvantage is that much of the sequence variation can be missed, since only fragment sizes, and not sequence data, are obtained. While AFLP analysis is still being used in many plant and animal systems, this technique will gradually be replaced by other marker systems and large-scale sequencing approaches.

### 3.5. Microsatellites

A common problem with comparing DNA sequences is that often there is insufficient variation for comparing individuals at the DNA sequence level. One solution is to focus on the most variable regions of genomes. One of the most variable types of regions in the genome is 'satellite DNA'; regions of highly repetitive nucleotide sequences. This DNA tends to be furthest from the regions that encode proteins, and it generally does not get expressed. Thus, differences do not carry any selective disadvantages and variation can build up faster. An additional reason for more variation in satellite DNA is that the mechanisms of DNA change (mutation) can be orders of magnitude faster than for protein-coding regions. Micosatellites are regions made up of short nucleotide repeats, with a repeating 2–5 nucleotides long, which can be repeated up to a 100 times within that microsatellite (e.g., GATGATGATGATGAT...). Variation is encountered in the number of repeats at that region. One copy of the region might have 20 copies of 'GAT', whereas a different individual (or chromosome) may have 35 copies. Populations can possess many alternative copy numbers, which are the microsatellite version of alleles of a gene. Analysis of microsatellites reveals some of the highest levels of variation of any part of the genome and this seems to be the case across plants, animals, and fungi. Most species possess thousands of microsatellite loci. A small subset of these regions is used to establish DNA fingerprints in humans for forensic and paternity cases. Microsatellites have also been used to examine genetic variation in aspen (Dayanandan et al., 1998; Wyman et al., 2003). Because each microsatellite contains so much variation, data from a small number of microsatellite loci (6–10) can be sufficient to distinguish clones. Scoring errors, the presence of closely related individuals, and mutations within clones present risks in using microsatellites for clonal identity in aspen, although risks can be minimized through the use of appropriate quality control measures in the laboratory, the careful selection of loci, and various statistical approaches (Parks and Werth, 1993; Arnaud-Haond et al., 2005; Hoffman and Amos, 2005; Schnittler and Eusemann, 2010).

### 3.6. Single nucleotide polymorphisms (SNPs)

A specific region of sequenced DNA contains many nucleotide positions that are identical across all samples, and only a few positions that vary. More recent approaches can target only those nucleotide positions that vary: single nucleotide polymorphisms (SNPs). These data can be sampled in a variety of ways. If your organism has been well-studied (such as humans or laboratory model organisms) then tools are available that simply collect data at the variable positions. For example, in humans well over 10 million SNPs have been characterized (www.ncbi.nlm.nih.gov/SNP/). On average, 1 position in 1000 is variable. This provides outstanding power for comparing individuals and populations. In organisms for which no complete genome sequence is available, SNPs can also be targeted by screening from vast genome-scale sequencing projects that are now affordable and feasible. Examples of such approaches include reduced representation sequencing approaches such as Genotyping by Sequencing (GBS) (Van Tassell et al., 2008; Elshire et al., 2011) and Restriction Site Associated DNA markers (RAD tags)(Miller et al., 2007; Hohenlohe et al., 2010; Davey et al., 2011). In these techniques, DNA sequences are obtained from a subset of nucleotide sites across the genome using sequencing tools that can produce many (tens of millions) of sequence reads. Similar regions from different individuals are then compared, and the variable regions are extracted by computer analysis. This approach can be used to isolate tens of thousands of informative SNPs in a very fast and cost-effective way for an organism for which genomic resources are previously unavailable. These

techniques are revolutionizing population genetic studies, since large numbers of individuals can be assayed and allele frequency data for large numbers of genomic sites can be obtained.

### 3.7. cDNA analysis and transcriptome surveys

Only a small portion of most genomes is transcribed into RNA and only a fraction of that is translated into proteins. That portion transcribed is referred to as the transcriptome, and fractions of this can now be easily captured. First, it is necessary to extract RNA from the tissues being studied. The transcribed portion of the genome varies by individual, specific tissue, and environmental conditions or history. Although this is a small portion of the genome, information about genes transcribed under certain conditions can lead to the discovery of ecologically important genes. In *Populus*, transcriptome analysis has been used to characterize genetic responses to drought (Wilkins et al., 2009; Raj et al., 2011), day length (Hoffman et al., 2010), elevated $CO_2$ (Tallis et al., 2010), embolism response (Secchi et al., 2011) and other conditions. RNA is generally unstable and difficult to analyze, but it can be converted to DNA by the enzyme reverse transcriptase. This DNA copy of the RNA is cDNA, which can then be sequenced by a range of different strategies. Often, this is more efficient (and cheaper) than attempting to sequence and assemble an entire genome.

### 3.8. Flow cytometry and cytology

Individuals within a species can vary at specific genes or genomic regions. In plants it is also possible for the numbers of genomes present in cells to vary among individuals (i.e. polyploidy; see above). There are several indirect ways that polyploidy might be detected (e.g. physical traits or genotypes), but obtaining direct evidence requires some extra approaches. One approach is simply to count the chromosomes. Such cytological techniques for staining and visualizing chromosomes have been used for about a century (Belling, 1926), but they tend to require specific tissues such as root tips or flower buds (in which cells are actively dividing) and they are time-consuming. An alternative is flow cytometry, whereby the distribution of DNA content in individual cells is examined in an automated machine-driven system. This requires careful tissue homogenization, appropriate DNA staining, and the use of individuals of known chromosome number as reference points (Suda and Travnicek, 2006; Mock et al., 2012). With these tools in place it is possible to screen large numbers of individuals for chromosome number.

## 4. Future directions

Here we have provided an overview of some of the currently available genetic tools and how they might provide useful information relative to the ecology and management of aspen forests. We also thought it useful to mention two of the emerging genetic approaches, epigenetics and metagenomics, which will be more broadly available in the coming years and which may fundamentally change our perceptions of genetics and ecology.

### 4.1. Epigenetics

Until recently, the finest scale at which the genes of an organism were assessed was the sequence of the four bases in DNA. Interestingly, within most cells of most organisms, the bases in DNA can become chemically modified during the lifetime of the individual. The most common such modification is the addition of a methyl group to certain nucleotides; another is modification of proteins involved DNA packaging in the nucleus. These modifications can affect the degree to which specific genes are expressed (see Section 3.7). Collectively the modifications are termed epigenetic effects. Thus, individuals can have identical DNA sequences, but can have different phenotypes based on epigenetic effects (Boyko and Kovalchuk, 2008; Golldack et al., 2011; Lauria and Rossi, 2011). In some cases, epigenetic effects are the result of ecological conditions experienced by the individual, and in some cases these changes can be heritable across generations. The emerging sequencing technologies enable analysis of epigenetic variation, but there are so far relatively few reports of epigenetic variation in natural populations of plants (Bossdorf et al., 2008; Bossdorf and Zhang, 2011; Richards, 2011; Abratowska et al., 2012). In aspen, given the ability of individual clones to persist for perhaps many thousands of years, epigenetic effects could be a significant component of genetic variation. Discovering epigenetic effects and associated ecological conditions (e.g. herbivory or frost damage) may be an exciting area of study in the future, and may eventually lead to some novel management recommendations in aspen.

### 4.2. Metagenomics

With current sequencing technologies, it is possible not only to obtain genomic sequences of individuals, but also genomic sequences of environmental samples (with all the component organisms). This approach is becoming more tractable as genomic databases expand (Tyson et al., 2004; Eisen, 2007). For instance, the soils associated with aspen clones could be subject to metagenomic sequencing, yielding DNA from plant roots, mycorrhizae, other fungi, and soil bacteria. This scale of analysis allows assessment of whole communities and how they change along clonal boundaries, with clonal diversity, or across different environments. This approach is a particularly powerful way to assess microbial communities, since most of these organisms cannot be cultured in a laboratory setting. In aspen, studies are already showing clone-specific differences in soil chemical processes in common garden studies (Madritch et al., 2006); metagenomic studies may eventually help us understand the effects of clonal diversity and stand structure on ecologically important soil processes (including nutrient cycling, carbon sequestration, and water retention) in natural populations.

## 5. Conclusions

Here we have provided an overview of some of the current and emerging genetic/genomic tools that can be used to answer important questions in natural resource management. We focus specifically on the way that these technologies can be used to address questions important to forest managers and ecologists, and we use aspen as our example species throughout. Specifically, we explain how molecular tools are already being used to understand the frequency and geography of seedling recruitment vs. suckering in aspen, how genetic variation is distributed across the enormous range of the species, and how the species has responded to past climate changes. Long and Mock (in press) have recently published a complementary review on how information from several recent studies (including genetic studies) are fundamentally changing management perspectives on aspen in the western US.

We also suggest that an increasing range of ecological questions in aspen are becoming tractable as molecular tools evolve, including clonal aging, the influence of genet diversity on soil processes, and a molecular understanding of adaptation to particular environments. It is our hope that the information presented here will help foster linkages between emerging molecular approaches and applied research, potentially improving management and conservation of this important forest species.

## References


Abratowska, A., Wasowicz, P., Bednarek, P.T., Telka, J., Wierzbicka, M., 2012. Morphological and genetic distinctiveness of metallicolous and non-metallicolous populations of *Armeria maritima s.l.* (Plumbaginaceae) in Poland. Plant Biol. 14, 586–595.

Ally, D., Ritland, K., Otto, S.P., 2008. Can clone size serve as a proxy for clone age? An exploration using microsatellite divergence in *Populus tremuloides*. Mol. Ecol. 17, 4897–4911.

Ally, D., Ritland, K., Otto, S.P., 2010. Aging in a long-lived clonal tree. PLoS Biol. 8, e1000454.

Anderegg, W.R.L., Berry, J.A., Smith, D.D., Sperry, J.S., Anderegg, L.D.L., Field, C.B., 2012. The roles of hydraulic and carbon stress in a widespread climate-induced forest die-off. Proc. Natl. Acad. Sci. USA 109, 233–237.

Arnaud-Haond, S., Alberto, F., Teixeira, S., Procaccini, G., Serrao, E.A., Duarte, C.M., 2005. Assessing genetic diversity in clonal organisms: low diversity or low resolution? Combining power and cost efficiency in selecting markers. J. Hered. 96, 434–440.

Belling, J., 1926. The iron-acetocarmine method of fixing and staining chromosomes. Biol. Bull. 50, 160–162.

Bernhardsson, C., Ingvarsson, P.K., 2012. Geographical structure and adaptive population differentiation in herbivore defence genes in European aspen (*Populus tremula* L., Salicaceae). Mol. Ecol. 21, 2197–2207.

Bossdorf, O., Zhang, Y.Y., 2011. A truly ecological epigenetics study. Mol. Ecol. 20, 1572–1574.

Bossdorf, O., Richards, C.L., Pigliucci, M., 2008. Epigenetics for ecologists. Ecol. Lett. 11, 106–115.

Boyko, A., Kovalchuk, I., 2008. Epigenetic control of plant stress response. Environ. Mol. Mutagen. 49, 61–72.

Boys, J., Cherry, M., Dayanandan, S., 2005. Microsatellite analysis reveals genetically distinct populations of red pine (*Pinus resinosa*, Pinaceae). Am. J. Bot. 92, 833–841.

Burczyk, J., Adams, W.T., Birkes, D.S., Chybicki, I.J., 2006. Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. Genetics 173, 363–372.

Callahan, C.M., 2012. Continental-scale characterization of molecular variation in quaking aspen. M.S. Thesis, Utah State University.

Cheliak, W.M., Pitel, J.A., 1984. Electrophoretic identification of clones in trembling aspen. Can. J. For. Res. 14, 740–743.

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., Blaxter, M.L., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Gen. 12, 499–510.

Dayanandan, S., Rajora, O.P., Bawa, K.S., 1998. Isolation and characterization of microsatellites in trembling aspen (*Populus tremuloides*). Theor. Appl. Genet. 96, 950–956.

De Woody, J., Rickman, T.H., Jones, B.E., Hipkins, V.D., 2009. Allozyme and microsatellite data reveal small clone size and high genetic diversity in aspen in the southern Cascade Mountains. For. Ecol. Manage. 258, 687–696.

DeByle, N.V., Winokur, R.P., 1985. Aspen: ecology and management in the western United States. In: US Department of Agriculture Forest Service, Rocky Mountain Forest and Range Experiment Station. Fort Collins, CO, USA, p. 283.

Dow, B.D., Ashley, M.V., 1996. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. Mol. Ecol. 5, 615–627.

Eckert, A.J., van Heerwaarden, J., Wegrzyn, J.L., Nelson, C.D., Ross-Ibarra, J., Gonzalez-Martinez, S.C., Neale, D.B., 2010. Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). Genetics 185, 969–982.

Eisen, J.A., 2007. Environmental shotgun sequencing: its potential and challenges for studying the hidden world of microbes. PLoS Biol. 5, 384–388.

Elliott, G.P., Baker, W.L., 2004. Quaking aspen (*Populus tremuloides* Michx.) at treeline: a century of change in the San Juan Mountains, Colorado, USA. J. Biogeogr. 31, 733–745.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6, e19379.

Eveno, E., Collada, C., Guevara, M.A., Leger, V., Soto, A., Diaz, L., Leger, P., Gonzalez-Martinez, S.C., Cervera, M.T., Plomion, C., Garnier-Gere, P.H., 2008. Contrasting patterns of selection at *Pinus pinaster* Ait. Drought stress candidate genes as revealed by genetic differentiation analyses. Mol. Biol. Evol. 25, 417–437.

Frey, B.R., Lieffers, V.J., Hogg, E.H., Landhausser, S.M., 2004. Predicting landscape patterns of aspen dieback: mechanisms and knowledge gaps. Can. J. For. Res. 34, 1379–1390.

Gallo, L.A., Geburek, T., 1991. Genetics of isozyme variants in *Populus tremula, P. tremuloides* and their hybrids. Euphytica 53, 225–233.

Geraldes, A., Pang, J., Thiessen, N., Cezard, T., Moore, R., Zhao, Y., Tam, A., Wang, S., Friedmann, M., Birol, I., Jones, S.J.M., Cronk, Q.C.B., Douglas, C.J., 2011. SNP discovery in black cottonwood (*Populus trichocarpa*) by population transcriptome resequencing. Mol. Ecol. Resour. 11, 81–92.

Golldack, D., Luking, I., Yang, O., 2011. Plant tolerance to drought and salinity: stress regulating transcription factors and their functional significance in the cellular transcriptional network. Plant Cell Rep. 30, 1383–1391.

Gray, L.K., Gylander, T., Mbogga, M.S., Chen, P.-y., Hamann, A., 2011. Assisted migration to address climate change: recommendations for aspen reforestation in western Canada. Ecol. Appl. 21, 1591–1603.

Grivet, D., Sebastiani, F., Gonzalez-Martinez, S.C., Vendramin, G.G., 2009. Patterns of polymorphism resulting from long-range colonization in the Mediterranean conifer Aleppo pine. New Phytol. 184, 1016–1028.

Guillot, G., Leblois, R., Coulon, A., Frantz, A.C., 2009. Statistical methods in spatial genetics. Mol. Ecol. 18, 4734–4756.

Hall, D., Luquez, V., Garcia, V.M., St Onge, K.R., Jansson, S., Ingvarsson, P.K., 2007. Adaptive population differentiation in phenology across a latitudinal gradient in European Aspen (*Populus tremula*, L.): a comparison of neutral markers, candidate genes and phenotypic traits. Evolution 61, 2849–2860.

Harrison, N., Kidner, C.A., 2011. Next-generation sequencing and systematics: what can a billion base pairs of DNA sequence data do for you? Taxon 60, 1552–1566.

Hegarty, M., Hiscock, S., 2007. Polyploidy: doubling up for evolutionary success. Curr. Biol. 17, R927–R929.

Hipkins, V.D., Kitzmiller, J.H., 2004. Genetic variation and clonal distribution of quaking aspen in the central Sierra Nevada. Trans. West. Section Wildlife Soc. 40, 32–44.

Hoffman, J.I., Amos, W., 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Mol. Ecol. 14, 599–612.

Hoffman, D.E., Jonsson, P., Bylesjo, M., Trygg, J., Antti, H., Eriksson, M.E., Moritz, T., 2010. Changes in diurnal patterns within the *Populus* transcriptome and metabolome in response to photoperiod variation. Plant Cell Environ. 33, 1298–1313.

Hogg, E.H., Hurdle, P.A., 1995. The aspen parkland in western Canada – a dry-climate analog for the future boreal forest. Water Air Soil Pollut. 82, 391–400.

Hogg, E.H., Brandt, J.P., Kochtubajda, B., 2002. Growth and dieback of aspen forests in northwestern Alberta, Canada, in relation to climate and insects. Can. J. For. Res. 32, 823–832.

Hogg, E.H., Brandt, J.P., Michaellian, M., 2008. Impacts of a regional drought on the productivity, dieback, and biomass of western Canadian aspen forests. Can. J. For. Res. 38, 1373–1384.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A., Cresko, W.A., 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet. 6, e1000862. http://dx.doi.org/10.1371/journal.pgen.1000862.

Holderegger, R., Buehler, D., Gugerli, F., Manel, S., 2010. Landscape genetics of plants. Trends Plant Sci. 15, 675–683.

Hyun, J.O., Rajora, O.P., Zsuffa, L., 1987. Genetic variation in trembling aspen in Ontario based on isozyme studies. Can. J. For. Res. 17, 1134–1138.

Ingvarsson, P.K., 2008. Multilocus patterns of nucleotide polymorphism and the demographic history of *Populus tremula*. Genetics 180, 329–340.

Ingvarsson, P.K., Garcia, M.V., Hall, D., Luquez, V., Jansson, S., 2006. Clinal variation in phyB2, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). Genetics 172, 1845–1853.

Ingvarsson, P.K., Garcia, M.V., Luquez, V., Hall, D., Jansson, S., 2008. Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae). Genetics 178, 2217–2226.

Jelinski, D.E., Cheliak, W.M., 1992. Genetic diversity and spatial subdivision of *Populus tremuloides* (Salicaceae) in a heterogeneous landscape. Am. J. Bot. 79, 728–736.

Joost, S., Bonin, A., Bruford, M.W., Despres, L., Conord, C., Erhardt, G., Taberlet, P., 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Mol. Ecol. 16, 3955–3969.

Kay, C.E., 1993. Aspen seedlings in recently burned areas of Grant Teton and Yellowstone National Parks. Northwest Sci. 67, 94–104.

Keller, S.R., Olson, M.S., Silim, S., Schroeder, W., Tiffin, P., 2010. Genomic diversity, population structure, and migration following rapid range expansion in the Balsam Poplar, *Populus balsamifera*. Mol. Ecol. 19, 1212–1226.

Keller, S.R., Soolanayakanahally, R.Y., Guy, R.D., Silim, S.N., Olson, M.S., Tiffin, P., 2011. Climate-driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus balsasmifera* L. (Salicaceae). Am. J. Bot. 98, 99–108.

Keller, S.R., Levsen, N., Olson, M.S., Tiffin, P., 2012. Local Adaptation in the flowering-time gene network of balsam poplar, *Populus balsamifera* L. Mol. Biol. Evol. 29, 3143–3152.

Kemperman, J.A., Barnes, B.V., 1976. Clone size in American aspens. Can. J. Bot. 54, 2603–2607.

Landhausser, S.M., Deshaies, D., Lieffers, V.J., 2010. Disturbance facilitates rapid range expansion of aspen into higher elevations of the Rocky Mountains under a warming climate. J. Biogeogr. 37, 68–76.

Lauria, M., Rossi, V., 2011. Epigenetic control of gene regulation in plants. Biochim. Biophys. Acta Gene Regul. Mech. 1809, 369–378.

Little, E.L.J., 1971. Atlas of United States trees, conifers and important hardwoods, vol. 1. US Department of Agriculture Miscellaneous Publication 1146, Washington, DC.

Liu, Z., Furnier, G.R., 1993a. Comparison of allozyme, RFLP, and RAPD markers for revealing genetic-variation within and between trembling aspen and bigtooth aspen. Theor. Appl. Genet. 87, 97–105.

Liu, Z., Furnier, G.R., 1993b. Inheritance and linkage of allozymes and restriction-fragment-length-polymorphisms in trembling aspen. J. Hered. 84, 419–424.

Liu, L., Li, Y.H., Li, S.L., Hu, N., He, Y.M., Pong, R., Lin, D.N., Lu, L.H., Law, M., 2012. Comparison of next-generation sequencing systems. J. Biomed. Biotechnol. http://dx.doi.org/10.1155/2012/251364 (Article ID 251364).

Long, J.N., Mock, K.E., in press. Changing perspectives on regeneration ecology and genetic diversity in western aspen: implications for silviculture. Can. J. For. Res.
Lund, S.T., Furnier, G.R., Mohn, C.A., 1992. Isozyme variation in quaking aspen in Minnesota. Can. J. For. Res. 22, 521–524.
Madritch, M., Donaldson, J.R., Lindroth, R.L., 2006. Genetic identity of *Populus tremuloides* litter influences decomposition and nutrient release in a mixed forest stand. Ecosystems 9, 528–537.
Maherali, H., Pockman, W.T., Jackson, R.B., 2004. Adaptive variation in the vulnerability of woody plants to xylem cavitation. Ecology 85, 2184–2199.
Marmiroli, M., Visioli, G., Maestri, E., Marmiroli, N., 2011. Correlating SNP genotype with the phenotypic response to exposure to cadmium in *Populus* spp.. Environ. Sci. Technol. 45, 4497–4505.
Marroni, F., Pinosio, S., Di Centa, E., Jurman, I., Boerjan, W., Felice, N., Cattonaro, F., Morgante, M., 2011. Large-scale detection of rare variants via pooled multiplexed next-generation sequencing: towards next-generation ecotilling. Plant J. 67, 736–745.
Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., Johnson, E.A., 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res. 17, 240–248.
Mills, T.R., Rumble, M.A., Flake, L.D., 2000. Habitat of birds in ponderosa pine and aspen/birch forest in the Black Hills, South Dakota. J. Field Ornithol. 71, 187–206.
Mock, K.E., Rowe, C.A., Hooten, M.B., Dewoody, J., Hipkins, V.D., 2008. Clonal dynamics in western North American aspen (*Populus tremuloides*). Mol. Ecol. 17, 4827–4844.
Mock, K.E., Callahan, C.M., Islam-Faridi, M.N., Shaw, J.D., Rai, H.S., Sanderson, S.C., Rowe, C.A., Ryel, R.J., Madritch, M.D., Gardner, R.S., Wolf, P.G., 2012. Widespread triploidy in western North American aspen (*Populus tremuloides*). PLoS ONE 7 (10), e48406. http://dx.doi.org/10.1371/journal.pone.0048406.
Moritz, C., 1994. Defining evolutionarily significant units for conservation. Trends Ecol. Evol. 9, 373–375.
Mueller, U.G., Wolfenbarger, L.L., 1999. AFLP genotyping and fingerprinting. Trends Ecol. Evol. 14, 389–394.
Neale, D.B., Kremer, A., 2011. Forest tree genomics: growing resources and applications. Nat. Rev. Genet. 12, 111–122.
Niedringhaus, T.P., Milanova, D., Kerby, M.B., Snyder, M.P., Barron, A.E., 2011. Landscape of next-generation sequencing technologies. Anal. Chem. 83, 4327–4341.
Nielsen, R., 2005. Molecular signatures of natural selection. Annu. Rev. Genet. 39, 197–218.
Otto, S.P., Whitton, J., 2000. Polyploid incidence and evolution. Annu. Rev. Genet. 34, 401–437.
Parks, J.C., Werth, C.R., 1993. A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). Am. J. Bot. 80, 537–544.
Pauley, S.S., 1956. Natural hybridization of the aspens. In: Minnesota Forestry Notes, p. 2.
Peterson, E.B., Peterson, N.M., 1993. Ecology and silviculture of trembling aspen, in: Comeau, P.G., Harper, G.J., Blache, M.E., Boateng, J.O., Thomas, K.D. (Eds.). Forestry Canada, Pacific Forestry Centre, and B.C. Ministry of Forests, pp. 31–51.
Pockman, W.T., Sperry, J.S., 1997. Freezing-induced xylem cavitation and the northern limit of *Larrea tridentata*. Oecologia 109, 19–27.
Rai, H.S., Mock, K.E., Richardson, B.A., Cronn, R.C., Hayden K. J., Wright, J.W., Knaus, B.J., Wolf P.G., submitted for publication. Transcriptome characterization and detection of gene expression differences in aspen (*Populus tremuloides*).
Raj, S., Brautigam, K., Hamanishi, E.T., Wilkins, O., Thomas, B.R., Schroeder, W., Mansfield, S.D., Plant, A.L., Campbell, M.M., 2011. Clone history shapes *Populus* drought responses. Proc. Natl. Acad. Sci. USA 108, 12521–12526.
Ramsey, J., Schemske, D.W., 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annu. Rev. Ecol. Evol. Syst. 29, 467–501.
Randall, W.K., Berrang, P., 2002. Washington Tree Seed Transfer Zones. Washington State Department of Natural Resources. <www.dnr.wa.gov/Publications/lm_wfn_seedzone_book.pdf> (accessed 30.09.12.).
Rehfeldt, G.E., Ferguson, D.E., Crookston, N.L., 2009. Aspen, climate, and sudden decline in western USA. For. Ecol. Manage. 258, 2353–2364.
Richards, E.J., 2011. Natural epigenetic variation in plant species: a view from the field. Curr. Opin. Plant Biol. 14, 204–209.
Richardson, B.A., Rehfeldt, G.E., Kim, M.S., 2009. Congruent climate-related genecological responses from molecular markers and quantitative traits for western white pine (*Pinus monticola*). Int. J. Plant Sci. 170, 1120–1131.
Romme, W.H., Turner, M.G., Gardner, R.H., Hargrove, W.W., Tuskan, G.A., Despain, D.G., Renkin, R.A., 1997. A rare episode of sexual reproduction in Aspen (*Populus tremuloides* Michx) following the 1988 Yellowstone fires. Nat. Areas J. 17, 17–25.
Romme, W.H., Turner, M.G., Tuskan, G.A., Reed, R.A., 2005. Establishment, persistence, and growth of aspen (*Populus tremuloides*) seedlings in Yellowstone National Park. Ecology 86, 404–418.
Schnittler, M., Eusemann, P., 2010. Consequences of genotyping errors for estimation of clonality: a case study on *Populus euphratica* Oliv. (Salicaceae). Evol. Ecol. 24, 1417–1432.
Secchi, F., Gilbert, M.E., Zwieniecki, M.A., 2011. Transcriptome response to embolism formation in stems of *Populus trichocarpa* provides insight into signaling and the biology of refilling. Plant Physiol. 157, 1419–1429.
Simonson, S.E., Opler, P.A., Stohlgran, T.J., Chong, G.W., 2001. Rapid assessment of butterfly diversity in a montane landscape. Biodivers. Conserv. 10, 1369–1386.
Spies, T.A., Barnes, B.V., 1982. Natural hybridization between *Populus alba* L. and the native aspens in southeastern Michigan. Can. J. For. Res. 12, 653–660.
Stebbins, G.L., 1938. Cytological characteristics associated with the different growth habits in the dicotyledons. Am. J. Bot. 25, 189–198.
Stevens, M.T., Turner, M.G., Tuskan, G.A., Romme, W.H., Gunter, L.E., Waller, D.M., 1999. Genetic variation in postfire aspen seedlings in Yellowstone National Park. Mol. Ecol. 8, 1769–1780.
Stohlgren, T.J., Coughenour, M.B., Chong, G.W., et al., 1997. Landscape analysis of plant diversity. Landsc. Ecol. 12, 155–170.
Strasburg, J.L., Sherman, N.A., Wright, K.M., Moyle, L.C., Willis, J.H., Rieseberg, L.H., 2012. What can patterns of differentiation across plant genomes tell us about adaptation and speciation? Philos. Trans. R. Soc. B-Biol. Sci. 367, 364–373.
Suda, J., Travnicek, P., 2006. Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry – new prospects for plant research. Cytometry A 69, 273–280.
Tallis, M.J., Lin, Y., Rogers, A., Zhang, J., Street, N.R., Miglietta, F., Karnosky, D.F., De Angelis, P., Calfapietra, C., Taylor, G., 2010. The transcriptome of *Populus* in elevated $CO_2$ reveals increased anthocyanin biosynthesis during delayed autumnal senescence. New Phytol. 186, 415–428.
Theunert, C., Tang, K., Lachmann, M., Hu, S., Stoneking, M., 2012. Inferring the history of population size change from genome-wide SNP data. Mol. Biol. Evol. 29, 3653–3667.
Turner, M.G., Romme, W.H., Reed, R.A., Tuskan, G.A., 2003. Post-fire aspen seedling recruitment across the Yellowstone (USA) Landscape. Landsc. Ecol. 18, 127–140.
Tuskan, G.A., Francis, K.E., Russ, S.L., Romme, W.H., Turner, M.G., 1996. RAPD markers reveal diversity within and among clonal and seedling stands of aspen in Yellowstone National Park, USA. Can. J. For. Res. 26, 2088–2098.
Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., dePamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y., Rokhsar, D., 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313, 1596–1604.
Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M., Solovyev, V.V., Rubin, E.M., Rokhsar, D.S., Banfield, J.F., 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature 428, 37–43.
Van Tassell, C.P., Smith, T.P.L., Matukumalli, L.K., Taylor, J.F., Schnabel, R.D., Lawley, C.T., Haudenschild, C.D., Moore, S.S., Warren, W.C., Sonstegard, T.S., 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nat. Methods 5, 247–252.
Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., 1995. AFLP – a new technique for DNA fingerprinting. Nucleic Acids Res. 23, 4407–4414.
Wilkins, O., Waldron, L., Nahal, H., Provart, N.J., Campbell, M.M., 2009. Genotype and time of day shape the *Populus* drought response. Plant J. 60, 703–715.
Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B., Rieseberg, L.H., 2009. The frequency of polyploid speciation in vascular plants. Proc. Natl. Acad. Sci. USA 106, 13875–13879.
Worrall, J.J., Egeland, L., Eager, T., Mask, R.A., Johnson, E.W., Kemp, P.A., Shepperd, W.D., 2008. Rapid mortality of *Populus tremuloides* in southwestern Colorado, USA. For. Ecol. Manage. 255, 686–696.
Wyman, J., Bruneau, A., Tremblay, M.F., 2003. Microsatellite analysis of genetic diversity in four populations of *Populus tremuloides* in Quebec. Can. J. Bot. 81, 360–367.