

*This report is a draft version. Work is still on progress.
The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.*

Synthesis Report: Forest Genetic Resources in the EU

Conservation of forest genetic resources has always been an integral part of tree improvement programs. However threats to FGR, such as habitat destruction, fragmentation, pollution, the use of low quality forest reproductive material and climate change have increased in the past decades and with it the need to maintain FGR. During the past two decades several policy instruments and initiatives have contributed to the conservation and sustainable use of forest genetic resources at both global and European level. The First Ministerial Conference on the Protection of Forests in Europe (MCPFE), held in Strasbourg in 1990, addressed the importance of conserving forest genetic resources (Strasbourg Resolution 2). EUFORGEN was established in October 1994 as an implementation mechanism of Strasbourg Resolution S2 (Conservation of forest genetic resources) of this first FOREST EUROPE Ministerial Conference. Recently, the FAO Commission on Genetic Resources for Food and Agriculture prepared the First State of the World's Forest Genetic Resources. Based on the findings of this report the Global Plan of Action for the conservation and sustainable use of forest genetic resources (GPA-FGR) was adopted, which identifies strategic priorities for action for conservation and sustainable use of FGR at the international, regional and national level.

Much research in forest genetic resources is devoted to species of economic importance. An array of research topics are contributing to the knowledge of forest genetic resources. Adaptive capacity is an important conservation goal in natural forests. Research on estimation of adaptive potential of natural tree populations, especially in relation to climate change is therefore an important research topic. For planted forests, conservation is motivated by the need to maintain a large genetic variation to be used in genetic improvement programmes. Genetic improvement of forest tree species is recent. With the advent of genomics approaches new opportunities for selecting improved germplasm may come. During the past two decades in the frame work of EUFORGEN gene conservation strategies and improved guidelines for management of gene conservation units were developed for forest trees in Europe. The development and broad use of molecular markers (mainly neutral markers) contributed to the population genetic knowledge of many tree species which is relevant to support the identification of conservation priorities. Genomic research is providing important information on adaptive variation and adaptation processes and is contributing to our understanding of the genetic basis of complex traits. This will be relevant for conservation and sustainable management of natural populations of trees in the near future. Landscape genomics contributes to our understanding of spatial patterns of adaptive genetic diversity of tree species.

Topics and outcome of research areas

Topic	Outcome
Tree breeding and improvement	Information on genetic differentiation of tree populations for quantitative traits (mainly productivity and wood quality), definition of seed zones, recommendations for transfer of forest reproductive material
Strategies for conservation (in situ, ex situ)	Technical guidelines for genetic conservation and use of 32 tree species, species distribution maps, minimum requirements for dynamic conservation units of European tree species
Assisted migration	Hardly implemented as a climate change adaptation strategy due to uncertainties about future climate conditions, associated risks and existing policies
Climate change related research	Predictions of the impact of climate change under various climate scenarios, characterisation of bioclimatic envelopes of tree species based on models
Genetic diversity studies	Knowledge on colonisation history, understanding the distribution of genetic diversity among and within populations, mating systems and gene flow (migration, drift), inbreeding levels, phylogenetic relationships.
Adaptation	Better insight in the adaptive differentiation of populations
Gene flow	Insight in the extend of gene flow in trees and its evolutionary consequences, role of long-distance gene flow
Advances in genomics	Increasing number of neutral markers and SNP markers available, reference genome sequences
Applications of genomics research	Increased understanding of complex traits, e.g. adaptive, related to growth and wood properties, genomic tools not yet applied in conservation
Landscape genomics	New methods available to be used for prioritizing populations and geographic areas for in situ conservation
Seed storage technology	Progress in understanding the mechanism of viability loss on drying, and methods that can help conserve recalcitrant species in ex situ cryobanks
Genetic modification	Much progress on traits for industrial purposes, accelerating breeding strategies through early flowering, better understanding of the metabolic engineering of major wood components, but limited commercial application in tree improvement

Trends in forest genetic resources research

Tree breeding and improvement

Forest trees are long-lived woody perennials. Compared to crops, trees are in an early stage of domestication, with the exception of some tree varieties important for wood, pulp, paper and biofuels (Neale, 2007). Genetic improvement of forest tree species is recent, formally initiated in the USA in the 1950s then spread across European countries (Pacques 2013). Ruotsalainen (2014) gives a good overview of short- and long-term breeding programmes in conifer species and birches in some European (mostly northern) countries. The current status of tree breeding in Europe in nine species and species groups (larches, Norway spruce, sitka spruce, mediterranean pines, Scots pine, douglas fir, sycamore maple, common ash and wild cherry) is given in Pacques (2013). Breeding programs of most trees species are based on simple phenotypic selection at the level of single plants ('plus trees') or populations (Gailing et al, 2009). Even the most advanced breeding programs for pines and eucalyptus did not move much beyond the second or third generation of breeding populations. For

example, pine breeding in the US is just now in its 3rd cycle of breeding and testing (Neale and Kremer, 2011). For forest tree breeding, time is a major constraint. Tree breeding mainly addresses species of economic importance and for which artificial regeneration is used for afforestation and/or for reforestation (Pacques, 2013). For example, breeding is especially intensive in areas where short-rotation plantation forestry is practised, e.g. southern pines in the USA, radiate pine in New Zealand, Eucalypts in South America (Ruotsalainen, 2014).

Based on biological and genetic characteristics of the species, different breeding strategies have been developed taking advantage of species peculiarities such as inter-specific crossing possibilities and hybridization benefits (e.g. poplars and larches), selfing success to remove deleterious genes (e.g. in maritime pine to improve stem form), ease of generative (e.g. breeding programmes based on recurrent selection for Scots pine, Norway spruce and Douglas fir) or vegetative (e.g. clonal programme for wild cherry, partly for Sitka spruce) propagation. For example, in larch hybrid production, the high juvenile growth rate and disease resistance of Japanese larch and good quality and drought resistance of European larch is combined. Breeding objectives in most tree breeding programmes are rather general balancing between increased volume production, better timber quality and wide adaptation of the forest regeneration material. The observed genetic gains in first-generation phenotypic seed orchards vary largely depending on species, breeding goals and intensity of selection. Genetic gains between 3% and 12% have been obtained for height growth and between 0% and 30% for early volume (Pacques, 2013).

Breeding efforts in European countries have resulted in the selection and creation of improved varieties for nearly all forest tree species. Synthetic varieties from first-generation seed orchards are now commercialized in most countries for major conifers but also for some broadleaves. Second-generation or even third generation seed orchards exist in some countries (e.g. maritime pine in France or Scots pine in Scandinavia). Over 1,000 seed orchards are currently available in 24 of the 28 EU countries for 40 different species, mostly Scots pine, Norway spruce and larch (Forest Material, April 2014). Clonal varieties of Sitka spruce obtained from elite families are also traded in some countries (e.g. UK, Ireland). Highly performing cultivars of several broadleaves (poplars, birch, willows, wild cherry, chestnut, black locust, alder, etc.) are also available for agro-forestry and forestry uses (Pacques, 2013).

Strategies for conservation and management

The main goal of genetic conservation is the maintenance of evolutionary processes such as natural selection, genetic drift, gene flow and mutation within tree populations. This so-called dynamic conservation approach can be implemented through in situ conservation (Koskela et al, 2013). In situ conservation of FGR is regarded as the most appropriate, sustainable and cost effective way of conserving FGR (FAO, 2014). According to Forest Europe (2011) a total of 476 385 hectares were managed for in situ gene conservation of forest trees in thirty-eight countries in Europe in 2010. The total area managed for ex situ gene conservation was 7697 hectares in thirty-six countries, and the total area managed for seed production (including seed orchards and seed stands) was 872 077 hectares in thirty-nine countries. A group of five economically important tree species (*Abies alba*, *Fagus sylvatica*, *Picea abies*, *Pinus sylvestris* and *Quercus petraea*) alone account for 74 percent and

66 percent of the total areas managed for in situ gene conservation and seed production, respectively. Koskela et al (2013) mention that these data monitor well the trend in conservation efforts but cannot be used for assessing how well forest genetic resources are conserved in Europe. Although, many countries in Europe developed strategies to maintain maximum forest genetic diversity through in situ and dynamic conservation measures, progress made to implement such programmes or strategies has been slower than expected (Koskela et al, 2013). Geburek and Konrad (2007) discussed a number of problems (methodological and political) related to FGR conservation, mainly in Europe. Among the impediments they mentioned are difficulties in assessing and monitoring genetic erosion and human impacts (e.g. the lack of markers showing adaptive variation and the lack of record keeping on the use and transfer of forest-tree germplasm, complexities of European national structures that make a development of a common strategy problematic, the general unattractiveness of genes as flagships in raising public awareness and the lack of integration of genetic aspects into biodiversity conservation.

Dynamic in situ conservation

Nowadays the Pan-European initiative EUFORGEN (European Forest Genetic Resources programme) has a crucial role in coordination of forest genetic conservation efforts in Europe. Its goal is to promote the conservation and appropriate use of forest genetic resources as an integral part of sustainable forest management in Europe. To date EUFORGEN has developed 34 species distribution maps and developed technical guidelines for conservation and use of 36 species (see www.euforgen.org). It also initiated the establishment of an European information system on forest genetic resources (EUGIS), launched in 2010, which provides georeferenced and harmonized data on dynamic conservation units of forest trees in Europe (www.eufgis.org).

In 2013 Koskela et al. developed pan-European minimum requirements for these genetic conservation units. In line with these minimum requirements, data on these dynamic conservation units were entered into the EUGIS database, which contains currently 3164 units and 100 tree species in 34 countries. Recently an assessment was carried out on the data on dynamic conservation units in the EUGIS database for 33 European countries (Lefevre et al, 2013). The results showed that the pan-European network of the dynamic conservation units covers all environmental zones and that conservation efforts are highly variable among species. For example 7 key target species (*Abies alba*, *Fagus sylvatica*, *Larix dedidua*, *Picea abies*, *Pinus sylvestris*, *Quercus rober* and *Quercus petraea*) are conserved in 60% of the conservation units. A gap analyses carried out for 11 species revealed that there are significant gaps in the genetic conservation efforts in terms of species and the geographical distribution of the units (particular in the marginal distribution of their range). Schueler et al, 2014 argued that beside identifying gaps in this dynamic conservation approach also the vulnerability of this large scale network of conservation units to climate change should be assessed. Therefore they identified the vulnerability of the individual conservation units in the light of the projected impact of climate change. Their results suggest that there is a strong need to intensify monitoring efforts and to develop additional conservation measures for populations in the most vulnerable units. As an example they mentioned establishment of dynamic conservation populations outside the current species distribution ranges within European assisted migration schemes.

EUFORGEN is now using the EUFGIS data to develop a pan-European genetic conservation strategy for forest trees and a genetic monitoring scheme for selected conservation units. (FAO, 2014, p 211). Genetic monitoring is an important measure to assess the genetic variation and to monitor changes over time. However, genetic monitoring of forests is at a very early stage, with so far only a small number of pilot studies. Aravanopoulos (2011) proposed a genetic monitoring approach including three indicators (natural selection, genetic drift and a gene flow-mating system), based on three demographic (age and size class distribution, reproductive fitness and regeneration abundance) and four genetic verifiers (effective population size, allelic richness, latent genetic potential and outcrossing/actual inbreeding rate). An operational genetic monitoring system for the dynamic conservation units may become feasible and cost-effective when more powerful and affordable molecular markers and novel statistical tools become available (Koskela, et al, 2013). Special research attention goes to developing management and conservation strategies for peripheral populations at risk due to climate change. The main question here is 'do these populations have reduced genetic potential to respond adaptively to environmental change?'. A long standing hypothesis is that peripheral populations exhibit low genetic diversity and greater genetic differentiation as a consequence of smaller effective population size and greater geographical isolation relative to geographically central populations. If peripheral populations compared to central populations are genetically depauperate owing to genetic drift and low gene flow, then they are perhaps of little significance in terms of future evolutionary potential (Eckert et al, 2008). On the other hand, Hampe and Petit (2005) argue that peripheral populations maintain substantial genetic variation and therefore deserve conservation effort. High levels of genetic differentiation are observed among such populations, leading to high levels of regional genetic diversity. Selection for local adaptation is expected in these populations and together with reduced gene flow this may result in distinct ecotypes. Review of empirical work on the genetic variation and its partitioning between central and peripheral populations, have produced mixed results so far. The outcome of a review of Eckert et al (2008) found evidence for reduced within genetic diversity and increased differentiation in peripheral populations based on 134 studies (115 species including forest trees).

Ex situ conservation

In situ conservation can be complemented by static conservation. This approach is more often seen as a short-term conservation strategy. For instance in the case of rare species, when in situ conservation is not feasible, or when species are at serious risk of extinction in the wild, or for safety duplication purposes. Ex situ conservation can also be dynamic if planted tree populations are exposed to natural selection and can evolve across generations (Koskela and Lefevre, 2013). Reported ex situ conservation activities by European countries are provenance trials, seed orchards, clonal repositories, botanical gardens and arboreta, and seed and pollen gene banks. 401 species are conserved ex situ in Europe according to FAO (2014). Ex situ conservation activities such as in vitro storage or DNA storage are hardly reported (e.g. *Pinus sylvestris* clones, *Ulmus* dormant buds) (FAO, 2014).

Assisted migration

Assisted migration is a conservation strategy that has been proposed to accelerate adaptation to climate change (McLachlan et al, 2007). Assisted migration refers to the relocation of a species to a location outside its existing or historical range that is predicted to be favourable for persistence under future climate projections (Loss et al, 2011). The capacity of trees to adapt to rapid climate change depends in part on the rate at which tree species are capable of extending their ranges. However simulations studies show that adaptation probably lag behind the moving optimum. (Kremer in Koskela et al, 2007). Therefore, assisted migration is proposed as an alternative climate change adaptation strategy. For efficient use of assisted migration understanding of phenotypic plasticity and knowledge of the adaptive range of tree populations is needed e.g. based on provenance trials. However, assisted migration has been the subject of discussion and criticism within the conservation community. The risk that a species will become invasive is one of the main criticism among uncertainty about future climate conditions and risks such as genetic pollution, hybridisation or introduction of pathogens (Loo et al, 2011, Park and Talbot, 2012). Trials of assisted migration are currently underway in North America. In Canada, the provinces of British Columbia, Alberta and Quebec have altered seed transfer rules to allow assisted migration to proceed (Pedlar et al, 2012). More recently, a genetic test plantation of North American Douglas-fir provenances in Europe was evaluated in order to quantify how tree populations respond when subjected to climate regime shifts. Based on climate model projections they demonstrated that climate trends observed over the last three decades warrant changes to current use of Douglas-fir provenances in plantation forestry throughout Western and Central Europe. Based on these findings they present a strong argument for making changes to the reforestation stock when establishing new Douglas-fir plantations in Europe (Isaac-Renton et al, 2014). Assisted migration may not be appropriate for every species, population or ecosystem and has not been widely implemented yet (Loo et al, 2011). Williams and Drumroese (2013) conclude that implementation of assisted migration will ultimately be limited by politics, costs, location and time.

Climate change related research

Climate change presents a new challenge to conservation of forest genetic resources and therefore gets much attention in forest genetics research. According to the Intergovernmental Panel on Climate Change (IPCC, 2007), climate change could increase average temperatures by 2–4°C in Europe over the next 50 years and cause considerable changes in regional and seasonal patterns of precipitation. This will alter the environmental conditions to which forest trees in Europe are adapted. There is still considerable uncertainty in the future climate projections based on different models. (Koskela et al, 2007). Populations may follow two main strategies to react to climate change i.e. adapt and/or migrate. In Europe, it is expected that most forest tree species will not be able to migrate northwards fast enough to follow the estimated temperature isocline shift given current climate change predictions. Over the next 100 years bioclimatic envelopes of forest tree species in Europe are predicted to be shifted northwards and eastwards from 300 to <800 km with considerable variation across models and species. It is estimated that at maximum, trees would be able to shift their range from 10 to 70 km during the next coming hundred years, not taking into account habitat fragmentation. Hence, natural dispersion would need to be assisted by artificial seed transfer to cope with the shifting climatic envelopes. (Kremer in Koskela et al, 2007, Kremer et al,

2012). There are different viewpoints regarding the impacts of climate change on FGR, which mainly relate to the species concerned and the environment (Loo et al, 2011). Climate change impacts are expected to be severe in dry, high-temperature regions where trees are at their adaptive limits (e.g. Lindner et al, 2010) and on marginal populations of temperate species (Matyas et al, 2009). Latitudinal and altitudinal migration have been already observed at the leading edge. At the rear edge of distribution, an increased mortality due to drought and heat stress is predicted for many species. (Kremer et al, 2014).

Genetic diversity studies

A wide range of molecular markers are used to evaluate the genetic diversity in tree species, to study the variation within and between populations, to study population genetic processes that influence patterns of genetic diversity, in particular gene flow, genetic drift, hybridization and mating systems. Some general results regarding genetic processes derived from marker studies are that forest tree species maintain high levels of diversity within populations and low among populations (Hamrick, 2004, Savolainen et al, 2007), gene flow is more influenced by pollen than by seed (Petit et al., 2005), gene flow covers long distances in northern and temperate forests where most species are wind pollinated (Petit and Hampe, 2006) and that many tree species naturally have large effective population sizes (Petit and Hampe, 2006). Also outcrossing forest tree species have a high level of individual heterozygosity (Gailing et al, 2009). Molecular marker studies have been carried out for most temperate species. Examples of species which are studied for their distribution of genetic diversity range-wide are beech (SSRs and isozymes), ash (cpDNA), oak (cpDNA), *Pinus sylvestris* (mtDNA), *Picea abies* (mtDNA, cpDNA) (for references see Pautasso et al, 2009). This population genetic knowledge is used to guide conservation actions e.g. designing sampling for conservation, prioritize populations for conservation. Especially, phylogeographic studies may help identify key regions deserving priority for conservation (Petit et al, 2003).

Most molecular markers such as SSRs, RAPDS, AFLPs based on genomic DNA are assumed to be selectively neutral. SNPs are a new type of markers and of high interest because of their abundance in the genome and their potential association with disease and adaptive traits. Nowadays SNPs are overtaking SSRs as the marker of choice, mainly because of their amenability for high-throughput genotyping and cost effectiveness. Diversity array Technology (DArT) is an affordable, high-throughput marker technology which has potential for genome-wide diversity analysis in several species of *Eucalyptus*, and phylogenetic analyses (Steane et al, 2011). Molecular-marker studies have contributed little to our understanding of natural selection and adaptation in forest tree populations. Most studies showed different patterns of molecular markers and quantitative variation. In conifers, molecular markers typically show far less variation than adaptive traits when sampled in the same populations or across the same range. (Gonzalez-martinez et al, 2006).

Adaptation

Knowledge of the adaptive genetic potential of forest tree populations is crucial for evaluating their risk to climate change. Adaptation responses include plasticity, epigenetic processes and genetic variation. Trees have high degree of plasticity and epigenetic adaptive responses (Schueler et al,

2014). This might suggest that trees may be more robust to climate change than thought. For studying adaptive genetic diversity in forest-tree populations common garden experiments (provenance tests, progeny and clonal tests) are traditionally used. Such studies have focussed on traits of economic importance such as survival, growth, wood properties, cold-hardiness, drought tolerance, and pest and disease resistance. These tests gave insight in broad patterns of adaptive genetic variation. Most forest tree species show clines of adaptive traits, due to diversifying selection in the past, in spite of a high level of gene flow in large continuous populations (Savolainen et al., 2007, reviewed in Alberto et al, 2013). Clinal adaptive divergence patterns in survival, growth and phenological traits have been found along a latitudinal gradient in several tree species (e.g. *Populus tremula*, *Pinus sylvestris*, *Picea mariana* and *Picea sitchensis*) (Krutovsky et al, 2012). Field experiments can estimate genetic parameters on measurable traits, but the individual genes underlying the traits are not known. (Gonzalez-Martinez et al, 2006). Alternatively to these common garden experiments adaptive genetic potential can be estimated on the basis of information from DNA sequences and genetic marker information. Recent experimental studies (Eckert et al, 2010) suggest that adaptive genetic variation has a pangenomic distribution, so multilocus approaches should be implemented to elucidate its genomic basis (Neale and Kremer, 2011).

Gene flow

Estimation of gene flow is important for monitoring processes that may affect isolation and local adaptation of tree populations, such as habitat loss and fragmentation, increased environmental stress due to climate change, introgression from domesticated trees into their wild relatives, introduction of maladapted germplasm during reforestation. Insight in these processes is important for developing strategies for management and conservation (Krutovsky et al, 2012). Gene flow is common in forest trees (reviewed in Savolainen et al, 2007). Wind-pollinated trees typically exhibit extensive gene flow within and among stands, resulting in broad mating neighborhoods. Average pollen dispersal distances in animal-pollinated trees vary depending on pollinator size, abundance and behaviour. Also animal-pollinated trees can still have an extensive amount of contemporary gene flow (Krutovsky et al, 2012). In Kremer et al, (2012b) a recent literature overview of long-distance pollen and seed dispersal in trees is given. Documented distances of effective pollen dispersal (led to successful mating) are up to 100 km. Wind-driven effective seed dispersal is up to a few kilometres. Animal-mediated seed dispersal can reach a scale of tens of kilometres, but generally, pollen dispersal distances are considerably longer than that of seeds. Estimated tree pollen and seed dispersal curves are very leptokurtic, which should increase the probability of gene flow between distant populations. Flowering tree density and synchrony affect the outcrossing rates and the distances for gene flow, but depends on species incompatibility systems. (Krutovsky et al, 2012).

A growing number of studies to estimate gene flow in trees have shown that isolated or fragmented populations do not become genetically depauperate due to increased selfing, inbreeding, genetic drift and reduced migration. This is in contrary to the main hypothesis that habitat fragmentation will reduce gene exchange among populations and that small, fragmented patches of a once-large continuous forest tree population will display reduced genetic diversity (Bacles and Jump, 2011). Apparently, trees in fragmented populations can have more long-distance pollen flow than trees in

large populations. As a consequences of extensive gene flow to fragmented populations genetic drift may not be so strong as expected. (Krutovsky et al, 2012 and references therein). Gene flow may constrain local adaptation or enhance the response to selection, depending on gene flow intensity, type of selection, demographic structure and genetic architecture of the trait under selection. Both theoretical and empirical data suggest that the positive effects of gene flow on adaptation may often dominate over the negative effects. This differs for leading edge, core and rear edge of forest distributions (Kremer et al, 2012b). At the leading limit gene flow may enhance local adaptation because of pre-adapted alleles coming from more central populations. At the southern edge gene flow may increase maladaptation, because of import of less adapted alleles. However, our understanding of the interaction between gene flow and local adaptation is still limited and mainly based on theoretical models and indirect inferences from the observed difference in population differentiation at neutral vs adaptive traits (Kremer et al, 2012b, 2014).

Advances in genomics

Research in forest tree genomics has lagged behind that of model and agricultural systems. The main reasons for this are the long generation times of forest trees, their large genomes, and the lack of well-characterized mutations for reverse genetic approaches and limited funding. However, progress has been made during the last 20 years owing to the advent of next-generation sequencing (NGS) technologies, but advanced genomic research programs are primarily restricted to a few species in the genera *Pinus*, *Picea*, *Pseudotsuga*, *Populus*, *Eucalyptus*, *Quercus* and *Castanea*, actually all species of significant economic value. There is a growing amount of information and genomic resources of forest trees, see also tree genomics databases such as http://dendrome.ucdavis.edu/treegenes/pubdata/summary_count.php.

Genome sequencing projects

Completed and running genome-sequencing projects in forest trees are limited to about 25 species from mainly four families: Pinaceae (pines, spruces and firs), Saliceae (poplars and willows), Myrtaceae (*Eucalyptus*) and Fagaceae (oaks, beeches and chestnuts). Black cotton wood (*Populus trichocarpa* Torr. and Gray) was the first forest tree genome to be sequenced (Tuskan et al, 2006). This species was chosen because of its relatively small genome (450 Mb) and because it is a target feedstock species for biofuels. With this sequence as a reference, additional *Populus* species are being sequenced using NGS platforms. The second forest tree sequenced was *Eucalyptus grandis* (Myburg et al, 2014). *Eucalyptus* species and their hybrids are important commercial species (Neale et al, 2013 and references therein). As conifers have a large genome size (20-30 Gbp), more than 7 times the size of the human genome and high frequency of repetitive sequences they were lagging behind in genome sequencing. With advances in DNA sequencing technologies sequencing of conifer genomes became possible (Neale et al, 2013). Several conifer species are sequenced now or under way in different projects in Sweden (Norway spruce), USA (loblolly pine, sugar pine, Douglas fir), Canada (*Picea glauca*), Spain (*P. pinaster*) and Europe (Scots and Maritime pines) First assemblies of the white spruce genome (*Picea glauca*) (Birol et al. 2013), *Picea abies* (Norway spruce) (Nystedt et al, 2013) and *Pinus taeda* (Loblolly pine) (Zimin, et al, 2014) are recently released. For a recent

overview see also the NCBI website. There is significant activity in resequencing for polymorphism discovery (Neale et al, 2013) taking advantage of the reference genome sequences.

Genomic resources

Genomic resources such as expressed gene sequences (ESTs), Unigenes, reference genetic maps and SNPs are reasonably well developed for the best studied genera (Pinus, Picea, Pseudotsuga, Populus, Eucalyptus, Quercus and Castanea) (Mackay et al, 2012 for conifer genera and websites therein, see Neal and Kremer, 2011, Neale et al, 2013 for overview).

Comparative genomics

Comparative genomics can be a powerful approach to understanding gene functions in trees. Both the lack of standard reverse-genetic tools in forest trees, such as easy transformation and regeneration and their long generation time makes understanding gene function challenging in trees. There are several databases available to facilitate collaboration on comparative-genomic analyses, of which TreeGenes is the most comprehensive one. However infrastructures and a more open access structure of data could be better developed to facilitate collaboration (Neale et al, 2013). Comparative mapping is by now restricted to comparison between species within a genus of family, e.g. studies reported for Pinus, Eucalyptus, or Fagaceae, but will become possible among distantly related plant groups soon as more genomes will be sequenced (for references see Abril et al, 2011). Comparative mapping has shown that genes are conserved over long chromosomal regions among related species.

Other 'omics'

There are large-scale investigations in gene expression in forest trees. Transcript profiling investigations are focussed on growth (embryogenesis, root formation, bud phenology, wood properties such as xylem, wood-forming tissue, biotic factors such as pathogens, insects, as well as abiotic factors, such as drought, cold acclimation (Mackay et al, 2012 and references therein). Most research is conducted in Populus (reviewed in Gratapaglia et al, 2009), while wood formation has received the most attention. An understanding of transcriptome dynamics and structure in forest trees is still poor (Neale et al, 2013 for overview, Neale and Kremer, 2011). There is only a modest contribution of proteomics to the knowledge of forest trees. Abril et al (2011) gives a nice overview of papers published on forest tree proteomics, limited to a reduced number of genera, including Pinus, Populus, Picea, Eucalyptus, Fagus and Quercus and mainly using first-generation approaches. Most of the forest tree proteomics papers in their overview deal with descriptive proteomics and differential expression and are aimed at studying growth and development, responses to stresses, organogenesis, and embryogenesis. Forest tree proteomics is a challenging research area. Initial proteome reference maps for some key organs and processes have been established (i.e. leaf maturation, rooting, wood formation). The major bottleneck in tree proteomics is the low number of proteins that can be studied. (Abril et al, 2011). The drought proteome has been examined in several tree species including poplar and spruce. Initial drought studies in poplar revealed limited overlap

between drought-related transcripts and proteins, suggesting the need for complementary approaches to unveil the mechanisms and molecular plasticity that control drought responses in trees. Complexities in the proteome among and within individual trees highlight the diversity of drought response (reviewed in Hamanishi et al, 2011). Metabolomics is lagging behind. For example metabolite profiling is an integral component of functional genomics research in a genotype of hybrid poplar (*Populus tremula* x *P. Alba*) (Harfouche et al, 2012).

Genetic mapping

Genetic linkage maps have been developed for the most economically important tree species using a variety of markers (Kremer et al, 2012a, reviewed in Ritland et al, 2011). More recently, high density gene-based linkage maps have been constructed for several conifers (e.g. *P. glauca*, *P. taeda* and *P. pinaster*), due to SNP marker development and high-throughput genotyping methods. Together with the reference map of loblolly pine this provides opportunities for comparative mapping. (Mackay et al, 2012).

Applications of genomic research

Most traits of economic or ecological interest in forest trees are complex quantitative traits, e.g. growth, yield, wood properties, resistance to diseases and insects, resistance or tolerance to abiotic stresses. To study the genes that underlie complex traits genetic approaches such as QTL and association mapping are used in forest trees (Neal and Kremer, 2011).

QTL mapping

QTLs have been mapped in most major forest tree species for a wide range of characters including phenology, wood quality, cold hardiness and drought tolerance (reviewed in Neale and Kremer, 2011; Harfouche et al, 2012 and Grattapaglia et al, 2009). QTL detection has limited application. The use of marker-assisted selection (MAS), following QTL mapping has not been applied in most tree breeding programs (except for within-family selection of elite families), due to low levels of LD, polygenic character of the traits and the variable expression of QTLs across environments (Neale and Kremer, 2011, Grattapaglia 2014). For example, in conifers QTLs explain about 5 – 15% of the total phenotypic variation associated with the trait (Gonzalez-Martinez et al, 2006). The availability of high resolution maps with new genotyping technologies and large pedigrees may renew the interest in QTL mapping (e.g. for growth traits in *Populus*) (Grattapaglia et al, 2009, reviewed in Mackay et al, 2012).

Association genetics

Recently, association genetics had become the preferred method to identify selectable genetic polymorphisms for MAS (Neale and Kremer, 2011). E.g. candidate gene associations' studies have been conducted with respect to wood-quality genes (*P. taeda*, *P. radiata*, *P. glauca*), drought-related traits (*P. taeda*), timing of growth cessation or cold tolerance initiation (*P. menziesii*, *P. sitchensis*) and aridity-related environmental variants (*P. taeda*) (reviewed in Mackay et al., 2012, Khan and Korban, 2012). These studies confirm results of earlier QTL mapping studies, i.e. most marker-trait

associations identified account for only a small amount of the genetic variation (Mackay et al, 2012). For example, in *Pinus taeda* for drought related traits the individual effects of SNPs within a gene were < 4%. In a study in *Eucalyptus* SNP variation in the Cinnamoyl CoA Reductase explained up to 5.9 % of the variation for the wood property trait microfibril angle (Khan and Korban, 2012). Genome wide association mapping (GWAS) is tractable when carried out in populations sufficiently large to detect common and rare alleles and to yield polymorphisms of value for breeders. GWAS is feasible for species where whole genome sequences are available, because of the enormous amount of SNP markers required in outcrossing trees and because of rapid decay of LD. If not, the candidate-gene approach has been preferred (Gailing et al, 2009). Genome-wide association mapping studies have been reported for a few species, for example for wood characteristics in *Populus* (Porth et al, 2013) and in *Cryptomeria japonica* (Uchiyama et al, 2013). However this might change when reference genome sequences become available and high density SNP genotyping or full-genome resequencing becomes cost-effective (Neale and Kremer, 2011).

Genomic selection

Genomic selection (GS) is fundamentally distinct from marker assisted selection in that all available genetic markers are fitted simultaneously to develop a prediction model utilizing phenotypic and genotypic data collected from a training population. These models are then used to predict genomic breeding value of progeny in future generations (Harfouche et al, 2012). Genomic selection is particularly well suited for tree species as they have long generation times, characteristics with low heritability, traits that are expensive to measure and traits that are expressed late in the life cycle. Therefore impact on tree breeding can be greater than in crops. Current studies suggest that by incorporating genomic selection into tree breeding improvement, prediction models will reach accuracies that are equal to or higher than traditional phenotypic selection when using populations with small effective size ($N_e < 30$). (see Grattapaglia, 2014, Harfouche et al, 2012). However, application of genomic selection in forest tree breeding is still in its infancy. There are a few proof-of-concept studies on genomic selection in loblolly pine and eucalyptus published, suggesting that genomic selection could be successful. The main hurdle is the lack of high throughput genotyping platforms for trees, the high genotyping costs. The extent of LD between markers and QTL is an important factor in GS. As genome wide average LD is low and declines rapidly in forest trees, this would mean that more markers and larger training populations are needed than in other species. Another aspect is that tree breeding programs maintain large breeding populations (> 200 individuals) to conserve genetic diversity, for such species more markers would be necessary to establish linkages. Additionally, the first studies show that it is likely that GS models in forest trees will be population specific (Grattapaglia, 2014, Isik 2014 and references therein).

Phenotyping

There is a need for high-throughput phenotyping technologies in forest trees both for traditional breeding and genomic selection. Forest trees are difficult to phenotype, they are big and long-lived and growth and phenotyping of trees in laboratory environments has limited value (Neale and Kremer, 2011). Methods for fast screening of large numbers of trees for breeding of wood characters are under development. For example, methods are developed for measuring modulus of elasticity of

standing trees, heart wood moisture content and concentration of extractives contributing to heart wood decay resistance (Ruotsalainen 2014). Emerging techniques are X-ray micro densitometry and NIRS (near-infrared spectrometry).

Landscape genomics

Landscape genetics is an approach combining genetic markers data and GIS data and may be used to improve in situ conservation of forest genetic resources. For example, in Austria they used a geographic grid based gap analysis for *Picea abies* to identify new genetic conservation units to cover neutral genetic marker diversity as well adaptive genetic diversity in the current network of conservation units (Schueler et al, 2012). Genomic approaches may allow the identification of adaptive genetic variation related to key traits, such as phenology and drought tolerance, so that management and selection of in situ conservation units may focus on maintaining adaptive genetic potential. In this context a landscape genomics approach provide valuable tools for understanding how selection in natural forest tree populations is shaping local adaptation. By using the latest high throughput genome sequencing tools to genotype a large number of SNPs across the genome and associating those markers with climate or other environmental variables, it is possible to begin to sort out adaptive from neutral variation (Sork et al, 2013, Manel and Holdegger, 2013, Allendorf, et al, 2010). Most forest trees show continuous variation of phenological traits along geographical patterns and investigations have been aimed at conducting association testing and clinal signature detection at the genomic level (Savolainen and Pyhajarvi, 2007). According to Neal and Kremer (2011) comprehensive databases of the standing adaptive genetic variation in forest tree populations and corresponding ecological data are needed so that this information can be combined with climate models to assess a species' ability to respond by either migration or adaption to predicted change of the environment. This application is expected to be a major focus of forest tree genomics research in the near future (see also Manel and Holdegger, 2013, Holdegger et al, 2010). So far, most landscape genomic studies have used tools such a as outlier locus detection. Other studies directly correlate allele frequencies with environmental factors. An example is the study of the landscape patterns of adaptive variation in *P. taeda* in the US. In this study they used environmental association analysis to search for correlations between climate variables and single nucleotide polymorphisms (SNPs) genotyped across a range-wide sample of loblolly pine populations while accounting for neutral levels of population structure (Eckert et al, 2010). There are still many challenges in assessing adaptive variation in empirical studies and to move the field of adaptive landscape genetics beyond association studies (Parisod and Holdegger, 2012). The approach has not been applied often yet. In a recent study of De Kort et al. (2014), they provided empirical evidence of local adaptation to temperature in black alder, *Alnus glutinosa* L. Gaertn, by using a landscape genomics analysis of *A. glutinosa* populations sampled across a regional climate gradient with phenotypic trait assessment in a common garden experiment. Examination of hundreds to thousands of loci in hundreds of individuals across landscapes will also improve the assessments of the interactions of gene flow, genetic drift influencing the evolution and persistence of populations, therefore both neutral and adaptive variation should be included (Allendorf et al, 2010).

Seed storage technology

Seed storage is not widely adopted as an effective long term ex situ conservation strategy. Many tree species produce seeds that cannot be stored in conventional seed banks due to sensitivity to desiccation. Over the last 20 years considerable progress has been made in understanding the mechanism of viability loss on drying, diagnosis which flora produce such seeds and developing methods that can help conserve such species in ex situ cryo-banks (Pritchard et al, 2014). Woody species are often difficult to establish in vitro, due to problems at the different stages of shoot culture establishment (contamination, polyphenols and tannins, growth rejuvenation, multiplication and rooting). In vitro storage includes restricted or minimal growth conditions and cryopreservation. Minimal growth storage has been reported for several tree species (e.g. Eucalyptus and Poplars) and is generally considered as a short- to medium-term conservation approach. Improved techniques for cryopreservation is an important innovation for recalcitrant seed. Woody species can be cryopreserved as shoot-tips, seeds or isolated embryo axes and embryogenic calluses. Most of the research has been on species that are of interest to commercial forestry. Cryopreservation of hardwood trees has become increasingly successful since the introduction the vitrification/one-step freezing protocol using PVS2. Survival rates higher than 50 percent are reported for e.g. Prunus, Populus species. Cryopreservation of embryogenic cultures of conifers is well advanced (e.g. in Abies, Larix, Picea, Pinus and Pseudotsuga). Cryopreservation of embryogenic calluses and somatic embryos from hardwood trees has been limited. Despite the progress made with cryopreservation, only a limited number of truly recalcitrant tree species have been successfully cryopreserved. This might have several reasons such as large seeds with high moisture content, viable tissue culture protocols are lacking, seed lots with high variation in moisture content and maturity stage. (FAO, 2014 and references therein).

Genetic modification

There is a diversity of species being pursued in genetic engineering research. About 33 forest tree species have been successfully transformed and regenerated (Straus et al, 2009). Traits under investigation include wood chemistry, herbicide resistance, insect resistance, disease resistance, growth, stature, salt tolerance, dormancy induction, nutritional conditions, onset of flowering, sterility, phytoremediation, cold tolerance, gene induction systems and rootability (reviewed in Abril et al, 2011). Genetic engineering is mainly used for industrial purposes. For example, the development of cold-tolerant hybrid Eucalyptus trees and lignin-modified hybrid poplar trees for biomass production under environmental stress and biofuel production. In addition, genetic engineering is used to accelerate breeding strategies in trees through early flowering. In addition to industrial purposes, research is conducted to use genetic engineering as a tool for conservation. For instance in the case of elm and chestnut in the USA it is used to rescue major tree species that have devastated by exotic diseases (Strauss et al, 2009). Future needs are improvement of more complex traits such as nutrient uptake and nutrient use efficiency. Another future area for improvement is engineering of belowground trait (Harfouche et al, 2011). Substantial progress has been made in understanding the metabolic engineering of lignin, a major wood component, because the genes that encode enzymes involved in lignin biosynthesis have been almost identified in poplar. However, the metabolic engineering of cellulose and hemicelluloses in forest trees awaits further development. (See Suzuki and Suzuki, 2014, for references and species). The next step in genetic improvement of

trees is cisgenesis and intragenesis. In this regard, activation tagging, a forward-genetic approach, is being used to generate mutant lines and identify endogens controlling key traits (tree stature, phloem formation) in poplar. Now it is mainly used to provide fundamental insight into poorly understood processes in trees. Other approaches such as RNA interference (RNAi) and the use of microRNAs (miRNAs) have been used in poplar (Harfouche et al, 2011). Although, genetic engineering of forest trees is studied intensively especially in the USA, commercial applications are currently limited to small-scale plantations of insect resistant poplars in China (Ruotsalainen, 2014). The majority of field trials have occurred in poplar (*Populus*) because of its status as a model organism for tree genomics and biotech, and most have occurred in the USA. Also, field trials have been conducted in a number of other tree species, e.g. pines and eucalypts (Strauss et al, 2009). Genetic engineering poses challenges in terms of public acceptance and policy. Recent increasingly stringent recommendations adopted by the CBD may impede much of the field research (Strauss et al, 2009).

Implications for the conservation and use of genetic resources in the EU and globally

Advances in genomic research including marker applications, gene identification and gene expression, understanding tree adaptation mechanisms may contribute to the conservation and sustainable use of forest genetic resources in Europe. Molecular data is becoming available rapidly nowadays, but has hardly been applied in conservation and tree improvement. Genomic research has most focussed on tree species of economic importance. Most studies are related to traits of commercial value such as growth and wood properties and to a lesser extent to biotic or abiotic stresses (e.g. poplar). Gene expression studies, QTL detection and association genetics will contribute to our knowledge on candidate genes. These studies will allow to estimate the diversity of these genes in natural populations to assess patterns of adaptive genetic variation. This genomic information may in the future support management plans and conservation of forest genetic resources (Neale and Kremer, 2011). For instance, information on the distribution patterns of adaptive diversity may contribute to locating areas suitable for the establishment of in situ gene conservation units. Currently, prioritising in situ conservation units (EUFGIS database) is based on information of the species' distributions, environmental conditions and, when available range-wide distribution on genetic diversity (Lefevre et al, 2013). Furthermore, genomic information may assist in selecting planting stock with high adaptive potential, designing programs for reforestation and ecosystem restoration. In the future understanding adaptive genetic variation will help to predict the response of tree populations to climate change, which may help developing seed transfer guidelines and guiding the use of forest reproductive material in the face of climate change. Landscape genomics approaches are just beginning to be applied to forest tree populations. Together with advances in bio-informatics and geographically based information, it will be possible to contribute to the management and conservation of diversity at the landscape level. A more immediately contribution of genomics to conservation is the increasing number of neutral markers available due to NGS technology. A genome-wide screening of genetic variation and exclusion of loci under selection (outlier loci) might improve the accuracy of estimation of parameters that require neutral

markers (Allendorf et al, 2010, Gonzalez-martinez et al, 2006). Both types of variation neutral and adaptive are important to gain knowledge of genetic variation patterns and processes and should be combined to make optimal management decisions (Funk et al, 2012). The epigenome may play an important role in phenotypic plasticity and adaptive potential of forest trees. However, knowledge of epigenetic effects in trees should be further developed.

Recent advances in forest genomics have increased the understanding of the genetic basis of different traits. Although genetic markers have found many applications, marker assisted selection (MAS) based on QTL mapping or association mapping has not been applied in practical breeding. The most promising strategy for using genomic information in applied tree breeding is now genomic selection, due to recent advances in sequencing technology and analytical tools (Isik, 2014).

Projected changes in climate are a serious threat to forestry and forest genetic resources. However, breeding programs hardly include traits relevant for adaptation to climate change such as pest and disease resistance, drought resistance, phenotypic plasticity. Tree improvement offers potential for adaptation to changing climate together with the need to conserve a large genetic base for an unpredictable future. Incorporation of climate change considerations into conservation strategies and measures is needed. EUFORGEN established a pan-European network of genetic conservation units for forest tree species according to pan-European minimum requirements (Koskela et al. 2013). With the predicted climate change it is expected that complementary measures to this network of in situ conservation units are needed, especially for populations within high risk. There is a lack of knowledge of assisted migration as a potential climate change adaptation strategy. Lefevre et al (2013) detected a poor coincidence between FGR conservation and other biodiversity conservation objectives within the analysed dynamic conservation units' network, which can be improved. Integration of conservation and use of FGR as part of sustainable forest management into national policies and strategies is needed.