



**Directorate General for  
Agriculture and Rural  
Development**

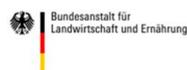
## **Preparatory action on EU plant and animal genetic resources**

### **Literature review**

*This report is a draft version. Work is still on progress.*

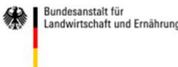
*The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf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

Preparatory Action on EU Plant and Animal Genetic Resources



# CONTENTS

- 1 Introduction.....3
- 2 Main topics and outcomes of past research in genetic resources relevant to the EU .....4
  - 2.1 Origin of diversity ..... 5
  - 2.2 Structure of diversity ..... 5
  - 2.3 Conservation technology ..... 5
  - 2.4 Coordination of conservation ..... 6
  - 2.5 Use of diversity ..... 6
- 3 Trends in genetic resources research.....6
  - 3.1 Origin of diversity ..... 7
  - 3.2 Structure of diversity ..... 8
  - 3.3 Conservation technology ..... 10
  - 3.4 Coordination of conservation ..... 12
  - 3.5 Use of diversity ..... 14
- 4 Implications for the conservation and use of genetic resources in the EU and globally. ... 19
  - 4.1 Increasing importance of genetic resources..... 19
  - 4.2 Improved conservation methods and access to genetic resources ..... 19
  - 4.3 Support for non-genomic sciences ..... 20
  - 4.4 Support for SMEs and NGOs ..... 20
  - 4.5 Scientific developments and societal acceptance ..... 20



# 1 Introduction

Science, and especially genetics, has never been as exciting as today. Developments in genomics have allowed the scientific community to start to truly understand the biological processes that are at the basis of life as we know it. Although the impact on agriculture has, so far, been relatively moderate, it can be expected that in the coming years, genomic science will change agriculture fundamentally. The first signs are already visible in the products of modern plant and animal breeding, new varieties and breeds, more efficient, meeting the changing consumer demands and better adapted to a changing environment. Climate change will require completely new varieties and robust breeds, to be developed quickly with traits so far largely unknown. Whether or not this will also involve genetically modified organisms (GMOs) or gene editing will depend on the political decisions in that regard. Forestry will have to go through similar changes to adapt to the changing environment, and it can be expected that the microbial life associated with agriculture and forestry, will follow suit.

The basis of changes in varieties, breeds, forest trees and microbes is genetic diversity. These changes can involve migration, such as proposed in the forestry domain, introduction of new crops and breeds or adaptation of existing crops and breeds to new circumstances and demands. In all cases GR are involved in two roles: as source of the diversity to be introduced or used in further improvement, and as resource that is threatened by the changes.

At this point science comes in again. Science can provide the answers to questions such as *'what diversity can best be conserved, and how can it best be conserved?'*, *'how can material for use be selected and how can it best be used to allow for those quick changes that are asked for?'*, and *'how can conservation of the diversity be assured to accommodate the demands of the future?'* All these questions surfaced in the review, some in some domains, and some in all.

Comparison of the domains gave a good impression of the 'state of knowledge' in these domains. In the Plant Genetic Resources (PGR) domain, *ex situ* collections and conservation methodologies are fairly well established, and the focus is largely, beyond molecular markers, on full re-sequencing approaches to increase knowledge of the sequence-trait associations, identifying new alleles and quantitative trait loci (QTLs) in germplasm collections for direct use in breeding. In the Animal Genetic Resources domain, methodologies for (*ex situ*) conservation require further development, making use of innovative technologies. Methodologies for maintaining within breed genetic diversity in breeding programs are fairly well established in the animal domain but genomic information will provide more detailed and better insights in genetic diversity. In the other domains, the methodology of conservation is more of an issue, and utilisation less. In the AnGR domain issues regarding genomic selection and genomic prediction of breeding values occur prominently, influencing the PGR and Forest Genetic Resources (FGR) domains, where the lessons learned regarding these genome-wide approaches in animal breeding are slowly being adopted.

Other new themes, such as epigenetics or landscape genomics, also appear in most domains to some extent. Only the Microbial Genetic Resources (MiGR) domain remains rather inward looking, concentrating on the organisation and methodology of conservation.

In this review all these elements will be discussed. Not all literature consulted by the domain experts will be referenced, their reports were the main sources for this overview. Only if statements refer directly to a reference this will be given, for all other sources and a more detailed discussion of the separate domains, please refer to the domain reports.

## 2 Main topics and outcomes of past research in genetic resources relevant to the EU

During the last half century, the awareness about the value and need to conserve genetic diversity of PGR, AnGR, FGR and MiGR has been well established. The activities of the FAO Commission on Genetic Resources for Food and Agriculture (CGRFA) illustrate the global awareness of this issue, the willingness to give it priority and take action. It resulted in valuable descriptions of the '*State of the World's Genetic Resources for Plants*' (FAO 1997, 2009), Animals (FAO 2007) and Forestry (FAO 2014), and corresponding '*Global Plans of Action*'. European countries have been leading in the process of establishing these documents and the following agreements and action plans. A large number of guidelines, standards and other useful documents have been produced to support the national activities to implement the action plans (e.g. FAO 2011, 2012, 2013). The activities of the CGRFA are ongoing and valuable both as a basis for national policies and as source of standards and reference documents.

At the basis of this awareness are the changes in agriculture that occurred during the green revolution in the sixties of the previous century. High input, high yielding varieties were produced that replaced landraces that had been grown since the domestication of the crops, and had slowly adapted to new environments and changed in these environments ever since. Thus, the homogeneous modern varieties replaced the heterogeneous landraces. A similar process is happening in the livestock sector. Local breeds have gradually been replaced by specialized, high productive breeds.

From these changes questions arose that are still determining the current genetic resources (GR) research. Questions that were shaped by developments in society and technology, but have remained basically the same:

- What is the origin of the diversity as we know and use it?
- What is the structure of diversity, how is it distributed over gene pools and within genomes and how to sample for conservation?
- How to technically conserve the gene pool (*in situ*, *ex situ*, on-farm methodologies)?
- How to organise and coordinate conservation (inter-)nationally?

- How to use diversity in breeding and in other services?

These five interconnected questions have formed the basis of the GR research agendas so far and will probably do so in the future. But although the questions have remained the same, the GR research agendas have changed dramatically over the last five decades.

## 2.1 Origin of diversity

Where it concerns research on the origin of diversity, research started with taxonomy, looking at relations between and within species. However this morphology based classical taxonomy has lost its prominence in the last two decades, and research on origin has been taken over by geneticists and bioinformaticians looking at selective sweeps and domestication traits and building phylogenetic trees.

## 2.2 Structure of diversity

The research on the structure of diversity saw a similar shift from agronomical evaluation of diversity in collections or breeding populations, through characterisation with a small number genetic markers towards high density single nucleotide polymorphism (SNP) marker panels, now followed by (re)sequencing of the genomes of germplasm accessions or breeding animals. Where it concerns sampling techniques, the popularity in PGR of quantitative genetic approaches such as core collections has decreased but is still present.

## 2.3 Conservation technology

With respect to practical conservation technologies there is a large difference between the domains. In PGR there was a typical dialectic process. As a response to the initial formal *ex situ* approaches, more informal on-farm conservation methods became popular, followed by a slow return to integrated strategies based on combinations of *ex situ*, on-farm and *in-situ* approaches. In this domain the technology for physical storage of the *ex situ* material did not change much: where possible in the form of frozen dried seeds, where not possible *in vitro* (in test tubes on slow growth conditions) or growing plants. In contrast, in the AnGR domain, the strategies did not change much; the basis was and is conservation of populations or breeds *in vivo*, often with a more or less coordinated breeding strategy on top and an *ex situ* back-up behind. However, the technology for *in vitro* storing (back-ups of) germplasm *ex situ* has developed in the recent decade from purely sperm for an increasing number of species towards other types of cells including embryos and different types of tissues for regenerating individuals always *in cryo* (liquid nitrogen). For forest trees the conservation method has basically always been the same and did not change: *in situ* (semi-)natural stands. However, the threat of climate change has urged the formulation of a new concept: dynamic conservation units, which take into account that assisted migration, might be needed. For the MIGR domain, conservation technologies, all *ex situ*, did not change much either. Depending on the type of organism, approaches based on growth media *in vitro*, in soil or even in host plants are used. These approaches have been improved continuously, but have never changed dramatically.

## 2.4 Coordination of conservation

The coordination of conservation activities, not research *per se* but the topic of many papers, has a similar structure in the PGR, AnGR and FRG domains in Europe, although the level of coordination differs. FRG, where national boundaries have the least significance, sees the best coordination, followed by PGR and AnGR. These three domains have some properly coordinated activities, especially in the field of documentation and inventories, but actual conservation activities are still very much based on the national autonomy. In the FRG domain, EUFORGEN (European Forest Genetic Resources Programme) is a relevant network and played an important role in the establishment of the European Information System on Forest Genetic Resources (EUFGIS), which provides georeferenced and harmonized data on dynamic conservation units of forest trees in Europe. In PGR an interesting initiative, AEGIS, has surfaced but it still has very little impact. A similar initiative started for AnGR developing a European Gene Bank Network (EUGENA). Coordination in the MIGR domain is in its infancy, based on time-bound projects. All four domains are subjects of the international legal developments regarding access and benefit sharing, as a result of the Convention on Biological Diversity (CBD) and the resulting Nagoya protocol. In all domains, this convention seems to have negative effects on access and benefit sharing. In the PGR domain, the International Treaty on PGR for Food and Agriculture is recognized as an alternative mechanism for ABS, trying to create a Multi-Lateral System for exchange of selected crops. For AnGR and FRG the FAO Global Plans of Actions are the most relevant policy frameworks and mechanism to support the sustainable use and conservation including ABS arrangements.

## 2.5 Use of diversity

Finally, in the last decade, the (diversification and) use of GR has come more into focus in the PGR and AnGR domains. The use of FRG as part of sustainable forest management is also getting more attention especially in relation to adaptation of forests and forest management to climate change. In the MIGR domain, the use of the collections has been a less important issue. The concept of conservation through utilisation has been fundamental in AnGR, and in fact also in the FRG community, but less in PGR due to the much higher number of distinct populations to be conserved. It is absent in MIGR. Molecular biology has created new options for use via marker assisted selection (mainly PGR domain) and genomic selection (mainly AnGR domain), and diverse options for gene discovery and mapping in all four domains. Information technology has created options for much better access to information about collections in all four domains.

## 3 Trends in genetic resources research

The trends in GR research will be grouped in the five main research topics, using quotes and segments from the 'domain reports' that are put in each other's perspective.

## 3.1 Origin of diversity

### 3.1.1 Taxonomy and domestication

Genetic / genomic data provide a powerful resource for answering questions on the origin of diversity, in all four domains. Standard research questions concern the geographical origins of crops and breeds, the number of independent domestication events, the molecular changes underlying domestication traits, and the nature of artificial selection during domestication and subsequent crop or breed improvement (Gross & Olsen, 2010). Molecular analyses have shown that convergent phenotypic evolution (i.e. with independent emergence of the same phenotypic trait in distinct lineages) is often based on molecular changes in orthologous genes (i.e. genes that arose from a common ancestor sequence via speciation) or pathways. Studies have revealed that during plant domestication the causal mutations for convergent changes in key traits are likely to be located in particular genes (Lensner & Theissen, 2013). See for examples the evolutionary and domestication studies in cucumber (Qi et al., 2013) and cabbage (*Brassica oleracea*) / rape seed (*B. rapa*) (Liu et al., 2014).

Similarly, in the AnGR domain, a large number of diversity studies has allowed the reconstruction of the domestication, migration, selection and adaptation history of most farm animal species (e.g. Ajmone-Marsan et al., 2010; Groeneveld et al., 2010; Miao et al., 2013). In particular for animals, studies of mitochondrial DNA genomes, which are maternally inherited, have helped to gain more insight in the origin and domestication of species. The availability of high density SNP chips and full sequence information has further increased the understanding of domestication processes, breed characteristics, introgression, signatures of selection and adaptation mechanisms (e.g. Bosse et al., 2014a; Bosse et al., 2014b).

In the past, effective population sizes and inbreeding levels of AnGR could be estimated only on the basis of pedigree data and/or using a limited number of markers. With the availability of sets of high density SNP markers or even whole genome sequence data, the variety of alleles, haplotypes and genotypes can be assessed more precisely (Allendorf et al., 2010; de Cara et al., 2011). Analysis of dense markers will give information about the level of heterozygosity, genetic diversity and signatures of selection.

In FGR, range-wide genetic surveys (cpDNA and mtDNA markers) contributed a lot to the reconstructions of refugia and of the spread of tree populations during the postglacial.

In the MIGR domain sequences of the internal transcribed spacer (ITS) region play an important role in taxonomic classification as phenotypically many species cannot be differentiated from each other (e.g. Schoch et al., 2012).

### 3.1.2 Epigenetics

'Epigenetics' is the study of mitotically and/or meiotically heritable changes in phenotype that arise independent of genetic variation, i.e. the study of heritable information that is not fully explained by DNA sequence variation (Schmitz & Ecker, 2012; Springer, 2013). Epigenetic information plays a role in developmental gene regulation, response to the environment, and in natural variation of gene expression levels (Springer, 2013). The extent to which epigenetic variation contributes to phenotypes remains to be determined. Several examples of phenotypic variation that can arise through epigenetic variation are known from the PGR

and FGR domain. Studies in the model species *Arabidopsis thaliana* have provided the richest source of known epialleles (Schmitz & Ecker, 2012; Becker et al., 2011). Natural variation in DNA methylation has been detected at thousands of loci among strains of *A. thaliana*, maize and rice. However, the ratio between the total number of changes and those resulting in developmental or physiological phenotypes suggests that most of these epigenetically differentially marked regions have little or no effects on genome and gene function. Hence, natural epigenetic variants affecting gene function are at least an order of magnitude less frequent than genetic mutations (Pecinka et al., 2013). Also in the FGR domain, epigenetics is studied as an explanation of the high degree of plasticity observed in forest tree species (Schueler et al., 2014). As a result, trees may be more robust to climate change than thought. 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).

It is now possible to profile the epigenome at high resolution. However, to determine if epigenetic variation can effectively be captured in breeding and GR programs, it is critical to understand its stability and heritability. If epigenetic variation is highly stable and heritable then it may behave in a fashion similar to genetic variation and be easily captured in breeding schemes, but, if epigenetic variation is relatively unstable, and can be influenced by environmental conditions, then it may be important to develop alternative strategies to capture or avoid this variation (Springer, 2013).

## 3.2 Structure of diversity

### 3.2.1 Molecular markers

Since the late 1960s, it has been possible to survey molecular markers across a genome, but for decades the number of markers that could be readily assayed placed limits on the genetic resolution that could be achieved using either experimental or comparative genetic approaches. Only a few years ago, the highest-density genetic maps contained thousand markers (Morrell et al., 2012). The use of molecular markers has led to clarification of the structure of genetic diversity in a wide range of crop species in the PGR domain (Glaszman et al., 2010).

Similar studies were done in the FGR domain, and have shown for, e.g., 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). Examples of species which are studied for their range-wide distribution of genetic diversity are beech (SSRs and isozymes), ash (cpDNA), oak (cpDNA), *Pinus sylvestris* (mtDNA), *Picea abies* (mtDNA, cpDNA) (Pautasso, 2009).

As far as molecular markers are still used, SNPs are overtaking microsatellite SSRs as the marker of choice, mainly because of their amenability for high-throughput genotyping and cost effectiveness. Also, alternative techniques are still having potential such as the Diversity array Technology (DArT) (Steane et al., 2014).

Additional markers systems that are still used in the MIGR domain for population studies and mating type distribution within populations include RFLPs, SSRs, RAPDs and AFLPs. Many of these are used to study natural hybridization processes between closely related species and also to look at the mating type distribution in a population that gives an indication of whether the population is monoclonal (Groenewald, 2007).

### 3.2.2 Next generation sequencing

Due to growing interest in human genome resequencing, a new generation of sequencing technologies has emerged. These next-generation sequencing (NGS) technologies are able to generate DNA sequence data at low cost and at a rate much faster than that of traditional technologies. With NGS technologies, it is possible to resequence entire genomes or sample entire transcriptomes more efficiently and economically and in greater depth than ever before (Furbank & Tester, 2011). This makes it possible to sequence hundreds or even thousands of related genomes, and to determine the genetic basis of trait variation and adaptation (Glaszman et al., 2010; Furbank & Tester, 2011; Varshney et al., 2009).

In the PGR domain, reference genomes have been published for a range of crops, including rice, maize, sorghum, soya bean, cucumber, tomato and even the large genome of bread wheat (Furbank & Tester, 2011; Morrell et al., 2012; Varshney et al., 2010; Varshney et al., 2011, Brenchley et al. 2012). Also in the AnGR sector genomes have been sequenced and annotated for many species during the past decade, including for chicken, dog, cattle, horse, pig, sheep and rabbit (e.g. Groenen et al., 2012; Wade et al., 2009). Research in forest tree genomics has lagged. Black cotton wood (*Populus trichocarpa*) was the first forest tree genome to be sequenced (Tuskan et al., 2006), followed by *Eucalyptus grandis* (Myburg et al., 2014). With advances in DNA sequencing technologies, sequencing of conifer genomes also became possible (Neale et al., 2013) and Norway spruce, loblolly pine, sugar pine, Douglas fir, *Picea glauca*, *P. pinaster*, Scots and Maritime pines all have been sequenced or are underway (Birol et al. 2013; Nystedt et al., 2013; Zimin et al., 2014).

In regards the MIGR domain, it can be observed that the number of ongoing fungal genome sequencing projects is almost tenfold fewer than those of bacterial and archaeal genome projects. The fungi chosen for sequencing represent narrow kingdom diversity; most are pathogens or models. However, there is now a trend in forward-looking multi-year fungal genome sequencing programs designed to capture fungal metabolic and evolutionary diversity. A prominent example is the 1000 fungal genome project that aims to sequence fungal genomes from across the Fungal Tree of Life in order to fill in the gaps. This project has the core goal of providing reference information to inform research on plant-microbe interactions, microbial emission and capture of greenhouse gasses, and environmental metagenomic sequencing (Spatafora, 2011).

In rice, the advent of next-generation high-throughput DNA sequencing technologies and the completion of high-quality reference genome sequences have enabled the development of sequencing-based genotyping and genome-wide association studies (GWAS) that have significantly advanced rice genetics research. (Huang et al., 2013). Many NGS initiatives are on their way, and it can be expected that within the next years thousands of PGR lines will have been resequenced. NGS also entered the livestock sector but has not fully exploded yet. For some species many individuals have been fully sequenced already (Daetwyler et al., 2014). Most of the sequence data seems to be generated in cattle so far (Cantet et al., 2014). Each AnGR livestock sector has to deal with very specific issues to apply and to benefit from genomic selection. When costs of genotyping will drop further, this will result in more marker based genetic evaluations. Possibly, future focus will be more on inclusion of causal variants in the genetic model, instead of genome wide genomic selection only (Cantet et al., 2014).

Schneeberger & Weigel (2011) indicated that in the PGR domain genome sequencing projects are under way for all major crops, but that not all alleles and genes of interest to breeders (e.g. with respect to valuable metabolites or disease resistance) can be found in the gene pool of domesticated species. Therefore, they argue that the largest impact of new sequencing methods might come from the use of these methods to extract genes from non-model, non-crop plants exhibiting heritable variation in important traits. Deployment of such genes to improve crops or engineer microbes that produce valuable compounds may cause a potential paradigm shift for plant biology.

### 3.3 Conservation technology

*Ex situ* seed storage can be used for the conservation of many plant species, but not all species have seeds that will store easily at conventional seed bank conditions (pre-drying seeds to low moisture content with circa 15% relative humidity air and storage at circa – 20°C). Orthodox seeds, such as those of the main crops in the PGR domain used in food and feed production, can be stored for long after dehydration and cooling, but recalcitrant seeds, such as seeds of most FGR species, are sensitive to drying (Li & Pritchard, 2009).

In *ex situ* conservation, the variability that has been collected remains static because the natural selection process is not allowed to continue. *In situ* conservation involves the maintenance and recovery of viable population of species in their natural surroundings, or in the case of domesticated species, in the surroundings where they have developed their distinctive properties (Newton et al., 2010). Although *ex situ* conservation of PGR is estimated to cost as little as 1% of *in situ* conservation, the two conservation approaches are viewed as complementary rather than alternatives (Li & Pritchard, 2009). To allow for the continued development of landraces, beside the static conservation under *ex situ* conditions, complementary on-farm and *in situ* management methods are needed (Carvalho et al., 2013). However, for conservation purposes in the narrow sense (conservation and availability), *ex situ* approaches proved much more effective.

During the past decade it became generally accepted in the AnGR community that *in vivo* and *in vitro* conservation strategies are complementary. FAO recently published guidelines

for cryoconservation (FAO, 2012) and for *in vivo* conservation (FAO, 2013), reviewing the state of the art for these complementary strategies. Furthermore, a general framework to choose the most appropriate (*in situ* or *ex situ*) conservation strategy has been designed (Gandini & Oldenbroek, 2007). For this purpose the use of a SWOT methodology to determine strategic directions also has an added value (Martin-Collado et al., 2013). Often there is a tension between breed development and conservation goals (Lauvie et al., 2011).

For most AnGR species, cryopreservation and use of cryopreserved semen is technically feasible, but results differ substantially between species. For example, the results of inseminating ewes with frozen semen should be improved, and cryopreservation of semen of birds is generally less successful compared to that of mammal semen. The state of development in cryobiology and reproductive technology has been reviewed in several papers in the past years (Long et al., 2014; Mara et al., 2013; Woelders & Hiemstra, 2011; Woelders et al., 2012). Cryopreservation of semen is the most common *ex situ* conservation approach. However, it has the disadvantage that genetic diversity of female animals is not captured and re-establishment of a population through backcrossing using semen will take multiple generations. Therefore, alternative options for cryopreserving different type of genetic material were proposed: embryos, gonadic tissues, primordial germ cells, stem cells and somatic cells. (Pereira & Marques, 2008; Song & Silversides, 2007a, 2007b; Tijssen et al., 2008; Groeneveld et al., 2008).

A variety of factors influence the risk status of breeds of AnGR *in vivo*, in particular the effective population size of the active breeding population and the expected change in population size or threats to the breed. In the recently published *in vivo* guidelines (FAO, 2013), different risk categories and thresholds have been defined for species with high reproductive capacity and species with low reproductive capacity. Furthermore, geographical concentration can be an important risk factor in case of disease outbreaks (Alderson, 2009). The EURECA project, funded by the European Commission, analysed in detail genetic and non-genetic factors, to assess the (self) sustainability of breeds (Hiemstra et al., 2010).

Prioritization of AnGR breeds for conservation can be supported by software, but application of these methods in practice is limited (Boettcher et al., 2010). Recently, integration of genetic, demographic and environmental factors has been made possible (Joost et al., 2010). Furthermore, an increasing amount of genomic data can be expected to support conservation decisions. In this context, maximization of conserved genetic variation on the basis of neutral genetic variation only may not be sufficient to conserve breed-level variation in important traits (Hall et al., 2012).

Genomic selection has the potential to better manage within-breed diversity in AnGR and to limit inbreeding rates. Genomic data is being used to determine molecular co-ancestry, which is a more accurate indicator for inbreeding compared to pedigree based co-ancestry. The effectiveness of strategies to maintain within-breed genetic diversity and to control the genetic background of a breed can be improved when genomic data is used (Toro et al., 2014). Relationship estimates on the basis of whole genome sequence data are significant different compared to estimates on the basis of high density SNP or pedigree data (Eynard et al., 2014).

In the FGR domain, *in situ* conservation is regarded as the most appropriate, sustainable and cost-effective conservation strategy (FAO, 2014). The main goal is the maintenance of evolutionary processes such as natural selection, genetic drift, gene flow and mutation within tree populations: 'dynamic conservation' (Koskela and Lefevre, 2013). 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, including difficulties in assessing and monitoring genetic erosion and human impacts (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 programmes). Schueler et al. (2014) argued that the vulnerability of the European network of conservation units to climate change should be assessed among others in the light of the projected impact of climate change.

*In situ* conservation of FGR can be complemented by static *ex situ* conservation, often seen as a short-term conservation strategy. It can be used, for instance in the case of rare species, when *in situ* conservation is not feasible, 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). *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). Woody species can be cryopreserved as shoot-tips, seeds or isolated embryo axes and embryogenic calluses. It has become increasingly successful since the introduction the vitrification/one-step freezing protocol using PVS2. Despite the progress made with cryopreservation, only a limited number of truly recalcitrant tree species have been successfully cryopreserved (FAO, 2014).

As the MIGR community is only based on *ex situ* methods, two of the most popular methods for conservation are lyophilisation (freeze-drying) and cryopreservation (Alexandraki et al., 2011). The preservation of micro-organisms by different drying methodologies has been used for decades. Freeze drying in particular is the preferred method for transporting and storing vast culture collections of micro-organism strain types. However, as many fungi cannot be freeze-dried due to lack of viable spores, it is concluded that cryopreservation is the best method for maintaining the genomic integrity of microorganisms in the future. (Smith and Ryan, 2012).

### 3.4 Coordination of conservation

Extensive *ex situ* PGR collections have been built up throughout the world, but their access and use for crop improvement are limited, because information on accessions is often limited to their taxonomic status and geographical origin (Glaszman et al., 2010; Varshney et al., 2010). In 2005, the Global Crop Diversity Trust initiated a consultation process leading to the

development of over 30 global crop and regional strategies for the *ex situ* conservation and utilisation of crop diversity. In a review of 26 strategies, including 18 global crop and 8 regional strategies, the primary constraints were found to be the quality and availability of accession-level information, and the availability of resources for regeneration, collecting, and research (Khoury et al., 2010). In Europe, the PGR domain was first to establish a Cooperative Program for PGR (ECPGR), a centralised repository for passport data on the collections (EURISCO) and to initiate a programme aimed at developing a 'virtual European genebank' (AEGIS).

In the AnGR domain, similarly, a European Genebank Network for Animal Genetic Resources (EUGENA) was established (Hiemstra et al., 2014) and a uniform Gene Bank Documentation system (CryoWEB) has been developed (Duchev et al., 2010). Gene bank data in CryoWEB is connected to the European breed database (EFABIS). At the global level, FAO published a set of guidelines for development and operation of national gene banks (FAO, 2012). Genetic tools and software are available to compose and optimize gene bank collections (see for example review of Boettcher et al., 2010). A number of recent studies looked into the costs of alternative AnGR strategies (Gandini et al., 2007; Silversides et al., 2012) which results can be used for decision making.

The MIGR community has been collaborating since 1982, when the European Culture Collection Curators Organization (ECCO) was established; it now brings together 64 collections (Smith, 2012). Beside this organisation, several other coordinating bodies exist such as, at the European level, the European Biological Resource Centres Network (EBRCN) and the European Consortium of Microbial Resources Centres (EMbaRC). The most recent EU initiative is the Microbial Resource Research Infrastructure (MIRRI, Schüngel et al., 2013). It aims at improving access to enhanced quality microbial resources in an appropriate legal framework, thus underpinning and driving life sciences research. MIRRI will provide coherence in the application of quality standards, homogeneity in data storage and management and sharing of workload to help to release the hidden potential of microorganisms. MIRRI will enhance existing European microbial collections linking them to non-European country partners globally and will bring added value. Furthermore, the register of culture collections of the World Data Centre for Microorganisms (WDCM, Sixth Version, 2014) presents data on more than 670 culture collections world-wide with holdings of almost 2.5 million strains of microorganisms, mainly bacteria and fungi (Smith et al., 2014). Most of these collections are small, specialized collections that supply cultures or other related services only by special agreement. Others, particularly national collections, publish catalogues listing the organisms held and provide extensive services for industrial and academic organizations (Çaktu & Türkoglu, 2011)

In the FGR domain, the Pan-European initiative EUFORGEN (European Forest Genetic Resources programme) has a crucial role in coordination of conservation efforts in Europe. It has developed 34 species distribution maps and technical guidelines for conservation and use of 36 species. It also initiated the establishment of a European information system on forest genetic resources (EUFGIS), launched in 2010, which provides georeferenced and harmonized data on dynamic conservation units of forest trees in Europe.

Recently, an assessment was carried out of the data on dynamic conservation units in the EUFGIS database for 33 European countries (Lefevre et al., 2013). It showed that the pan-European network of dynamic conservation units covers all environmental zones and that conservation efforts are highly variable among species.

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 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).

Finally, scientists from both developed and developing countries have repeatedly expressed concern about the harm that restrictive access regulations may have on scientific research (Jinnah & Jungcurt, 2009). Formerly, much of the exchange in the GR community was very informal (Dedeurwaerdere, 2010). Informal distribution of material occurred, and in some domains still occur, without any written contract. This is especially true for the smaller and more specialized collections. In recent years, this has changed dramatically, in both developing and developed countries. The adoption of rather restrictive access measures by several developing countries, as a reaction to excessive bioprospecting and patenting by developed countries, further threatens the efficacy of an informal regime. In particular, access procedures may lack transparency and be quite complicated, involving lengthy delays in obtaining genetic materials or obstructing it completely. The results will be very negative to access and benefit sharing, and thus to science (Jinnah & Jungcurt, 2009).

### 3.5 Use of diversity

#### 3.5.1 Gene and allele identification

QTL (quantitative trait loci) mapping has been used to determine the genetic bases of complex traits. Many QTLs associated with quantitative disease resistance have been identified, for instance for blast in rice, fusarium head blight in wheat and barley, late blight in potato and bacterial spot in tomato, but the resistance mechanisms controlled by these QTLs are still largely unknown (Kushalappa & Gunnaiah, 2013).

Genome-wide association (GWA) studies use a quantitative genetic approach to find genetic associations between genotype and phenotype in a population of individuals of unknown relatedness, to identify genetic loci contributing to such a phenotype (Flood et al., 2011). GWA studies have become an obvious general approach for studying the genetics of natural variation and traits of agricultural importance. They are especially useful when (in PGR)

inbred lines are available, because once these lines have been genotyped, they can be phenotyped multiple times, making it possible to study many different traits in many different environments (Atwell et al., 2010). GWA in the AnGR domain will, in generating information about the size and distribution of gene effects that control quantitative genetic variation, vastly support genomic selection (Cantet et al., 2014). Also in the FGR domain, GWA is applied; studies were reported for a few species, for example for wood characteristics in *Populus* (Porth et al., 2013) and in *Cryptomeria japonica* (Uchiyama et al., 2013).

Comparative genomics can be a powerful approach to understanding gene functions. Especially for the FGR domain, this is highly relevant, since the lack of standard reverse-genetic tools in forest trees makes understanding gene function challenging in trees. There are several FGR databases available to facilitate collaboration on comparative-genomic analyses, of which TreeGenes is the most comprehensive one (Neale et al., 2013). Comparative mapping is currently restricted to comparison between species within a genus of family, e.g. studies in *Pinus*, *Eucalyptus*, or Fagaceae, but will soon also become possible among more distantly related plant groups, as more genomes will be sequenced (for references see Abril et al., 2011).

The assessment of RNA (transcriptomics), protein (proteomics) and metabolite (metabolomics) levels can deliver information on genes in the target region associated with an mRNA, protein or metabolite shift linked to the trait of interest. It is now possible to generate omics datasets for many species, and, although the high costs of metabolomics still limit direct application in breeding, developments in omics technology are helping to elucidate the biological processes that determine gene effects. (Langridge & Fleury, 2011)

In the FGR domain, large-scale investigations in gene expression are on-going. Transcript profiling investigations are focussed on growth, biotic factors such as pathogens, insects, as well as abiotic factors, such as drought and cold acclimation (Mackay et al., 2012). Most research is conducted in *Populus* (reviewed in Grattapaglia et al., 2009), and wood formation has received most attention. The understanding of transcriptome dynamics and structure in forest trees is still poor (Neale et al., 2013; Neale and Kremer, 2011). There is only a modest contribution of proteomics to the knowledge of forest trees (Abril et al., 2011). Most of the forest tree proteomics work deals with descriptive proteomics and differential expression. The drought proteome has been examined in several tree species including poplar and spruce. Complexities in the proteome among and within individual trees highlight the diversity of drought response (reviewed in Hamanishi et al., 2011). Metabolomics in FGR is lagging behind (Harfouche et al., 2011).

The phenotyping of large numbers of genotypes makes possible the identification of trait-associated genomic regions and ultimately gene cloning, and the establishment of genetic markers that may assist in marker-based selection of germplasm (Dhondt et al., 2013). However, when relating genotypes to phenotypes, it is important to take into account genotype x environment interactions and to consider gene expression phenotypes not only as the product of a genotype, but also as a function of a particular environmental context (Hodgins-Davis et al., 2009).

Conventional phenotyping methods are often labour-intensive and destructive, usually involving the removal of biomass for analysis. However, great advances are being made in the field of phenotyping, as high-throughput, non-destructive methods are being developed for the PGR domain (Furbank & Tester, 2011; Roy et al., 2011). Imaging and image processing, including infrared thermography, are now major tools for phenotypic trait measurement (Dhondt et al., 2013; Furbank & Tester, 2011). High-throughput phenotyping systems enable several hundreds of plants to be processed daily by means of non-invasive imaging and image processing. Phenotyping systems may also comprise the means to grow plants in certain environments (controlled, semi-controlled, or uncontrolled). Automated workflows enable increased throughput when time-consuming or repetitive manual interventions and analyses can be automated (Dhondt et al., 2013). Application of these techniques in high-throughput, controlled-environment facilities has the potential to improve precision and reduce the need for replication in the field (Furbank & Tester, 2011). Despite the need for high-throughput phenotyping technologies in forest trees, phenotyping of trees in laboratory environments has limited value (Neale and Kremer, 2011).

'Phenomics' has been proposed as a novel discipline in biology, involving the gathering of high-dimensional phenotypic data at multiple levels of organization, to progress towards the full characterization of the complete set of phenotypes of a genome, in analogy with whole genome sequencing (Dhondt et al., 2013).

### 3.5.2 Landscape genomics

Landscape genomics is an approach combining genetic marker or genomic data with GIS data and may be used to improve *in situ* conservation strategies. For example, in Austria a geographic grid based gap analysis for *Picea abies* was used 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., 2013). By using NGS tools to genotype a large number of SNPs 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). 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). De Kort et al. (2014) provided empirical evidence of local adaptation to temperature in black alder (*Alnus glutinosa*) by using a landscape genomics analysis of populations sampled across a regional climate gradient with phenotypic trait assessment in a common garden experiment.

### 3.5.3 Genetic improvement

Marker-assisted selection (MAS) is now used routinely for crop plant traits conferred by single genes (Furbank & Tester, 2011). Because marker densities have increased, it has become possible to use markers to monitor the genetic background of lines, and this has provided the opportunity for accelerated backcrossing (Langridge & Fleury, 2011). The

prediction of variation in agronomic traits by diagnostic markers ultimately requires the identification of the causal genes. GWA mapping using genome-wide, evenly distributed SNP markers is an option to achieve this (Gebhardt, 2013). In the FGR domain, the use of marker-assisted selection (MAS) is not common in most tree breeding programs (except for within-family selection of elite families), due to low levels of linkage disequilibrium, the polygenic character of the traits and the variable expression of QTLs across environments (Neale and Kremer, 2011, Grattapaglia 2014).

By using genetic engineering, useful genes or alleles can be transferred across different species. While products derived with molecular breeding have been accepted and adopted, crops derived with genetic engineering struggle to gain acceptance and reach farmers' fields. It is clear that adoption of these crops will depend on public perceptions and public acceptance (Varshney et al., 2011). The same is true in the FGR domain. A diversity of species, about 33 forest tree species, have been successfully transformed and regenerated (Strauss et al., 2009). 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 were devastated by exotic diseases (Strauss et al., 2009). Although genetic engineering of forest trees is studied intensively and there is considerable potential for progress in tree improvement, commercial applications are currently limited to small-scale plantations of insect resistant poplars in China (Ruotsalainen, 2014).

#### 3.5.4 Genomic selection

The AnGR domain was the first to develop genomic selection. It 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., 2011). It has had an enormous impact on the livestock sector, starting with the dairy cattle sector. The key advantage of genomic selection is increased genetic gain through shortening of the generation interval, but investments are high (Hayes et al., 2013). Only recently, the PGR community has become interested in this approach. For the FGR domain it could be a solution as tree species have long generation times, characteristics with low heritability, traits that are expensive to measure, and traits that are expressed late in the life cycle. However, application of genomic selection in forest tree breeding is still in its infancy, there are only a few proof-of –concept studies (Grattapaglia, 2014a, b, Isik 2014 and references therein).

#### 3.5.5 Facing the global challenges

Mankind depends on fewer than a dozen of the approximately 300,000 species of flowering plants for 80% of its caloric intake, and uses only a fraction of the genetic diversity that resides within each of these species (McCouch et al., 2013). Genetic improvement via selection has been associated with a narrowing of the available genetic diversity, and concern has arisen that the prospects for continued genetic gain are becoming increasingly limited (Feuillet et al., 2008). In a time with increasing indications of climate change, this creates a major challenge.

In the AnGR domain, the importance of animal genetic diversity in the context of global livestock sector development, climate change, sustainable diets, cultural heritage and biodiversity objectives was flagged. A number of publications cover one or several of these global challenges and highlight the importance of better characterization of farm animal genetic resources and improved conservation and management strategies (Baumung & Hoffmann, 2012; Godfray et al., 2010; Hoffmann, 2010; Hoffmann, 2014; Seré et al., 2008).

Recent work has been done on valuation of breeds on the basis of market and non-market values. The total economic value (TEV) of a breed consists of the direct use value, indirect use value, option value, bequest value and existence value (Drucker, 2010; Zander et al., 2013). The concept of Payments for Ecosystem Services (PACS) associated with market failures and the public good characteristics of agrobiodiversity conservation was introduced (Narloch et al., 2011). The contribution of AnGR to key agro-ecosystem functions such as nutrient cycling, seed dispersal and habitat maintenance is widely acknowledged (Jackson et al., 2007). The Millennium Ecosystem Assessment (MEA) distinguished four groups of ecosystem services: provisioning services, regulating and habitat services, supporting services and cultural services. Livestock breeds are on one hand providers of ecosystem services and on the other hand dependent on ecosystem functions. Recently, some studies have been undertaken to collect information and evidence on the ecosystem services provided by livestock species and breeds (CGRFA, 2014).

Climate change presents a challenge to the conservation of GR in all domains (Fujisaka et al. 2009), but, due to long generation span and low mobility, especially in FGR. Therefore genetic resources get relatively much attention in forest genetics research. According to the Intergovernmental Panel on Climate Change (IPCC, 2007), climate change will alter the environmental conditions to which forest trees in Europe are adapted, although there is still considerable uncertainty in the future climate projections based on different models. (Koskela et al., 2007). These changes will also strongly impact phenology, species distribution and agricultural use of PGR (Jarvis et al., 2008; McMahon et al., 2011).

FGR 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. 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, 2007, Kremer et al., 2012). 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).

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

### 4.1 Increasing importance of genetic resources

The importance of genetic resources has increased and will continue to increase even more. The European society is becoming increasingly aware that it needs to anticipate or react to the effects of climate change. Genetic diversity is an issue appearing more and more in the media, not only in relation to climate change, but also when it concerns IP issues or diversification of food and agriculture. This literature review showed over and over again the expectation of the scientific community that technology, especially genomics, will revolutionize the way genetic resources can be understood in terms of evolution and domestication, and, more importantly, exploited in breeding and use in general. This is often put in the perspective of the increasingly important role of genetic resources and the need to anticipate to the on-going developments.

- The value of genetic resources is increasing due to global developments (incl. climate change). However support to the proper conservation of these resources is lagging behind and should be improved.
- Research on genomics in connection to genetic resources is very promising and further support for this research in the public sector is needed.
- Assuring the link between the upstream genomics research and the downstream gene bank activities deserves attention. Capacity building activities, demonstration projects and involvement of technology users in technology development projects will be necessary to assure this link.

### 4.2 Improved conservation methods and access to genetic resources

In parallel, it was observed that there is still much opportunity to improve the conservation methods, both on the technical and organisational level. Recent developments with regard to access and benefit sharing (ABS) of genetic resources has certainly not stimulated collaboration in the field of genetic resources conservation and use in Europe and even more so in the rest of the world.

- Research in the methodology of genetic resources conservation remains at a low level, support can be expected to pay off quickly via increased efficiency.
- Genetic resources conservation activities in Europe are hardly coordinated, partly due to ABS issues and funding mechanisms. Further coordination of activities has large potential for increasing the impact of the investments currently made in conservation and use.
- Access and benefit sharing of genetic resources need to be properly organised in a way that stimulates collaboration and exchange of material. Initiatives such ITPGRFA need to be supported.

### 4.3 Support for non-genomic sciences

Other areas of research relevant to genetic resources conservation, beside genomics, clearly stand in the shadow of genomics. Conservation biology, quantitative genetics, socio-economics, just to name a few, are not getting much attention, although their relevance to genetic resources conservation and use has not decreased.

- Non-genomic sciences could have a large impact on the efficacy of efforts to improve conservation and support use, and should be supported to do so.

### 4.4 Support for SMEs and NGOs

Also on the use side, a shadow of the genomic revolution can be observed. For small and medium size enterprises (SMEs) and non-governmental organisations (NGOs) it will be increasingly difficult to continue to play a role in the genetic resources arena as it becomes more and more dominated by molecular approaches.

- To assure future involvement of SMEs and NGOs in genetic resources conservation and use, these companies and organisations will have to be supported in terms of capacity building and technology application.

### 4.5 Scientific developments and societal acceptance

Finally, the implications of the scientific developments on the European society are substantial and need to be anticipated. These implications include the composition and health of the food in the shops, but also ethical aspects of technological developments (e.g. GMO's, intensive livestock keeping, conservation of clonal cells, use of pharmaceuticals and agro-chemicals).

- The debate about the impact of technology on the use of genetic resources and the resulting practices and products needs to be stimulated and supported. The public opinion should be allowed to develop on the basis of evidence and open discussion.