

*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: Plant Genetic Resources in the EU

Main topics and outcomes of past research in plant genetic resources in the EU

Genetic diversity

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). Since the beginnings of agriculture cereals have been the main source of calories for mankind. These cereals have been bred intensively for the past hundred years, resulting in huge improvements in yield and quality (Feuillet et al., 2008). During the green revolution, for instance, high-yielding rice and wheat varieties turned India into a net food exporter (McCouch et al., 2013). Yield improvements, however, have 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). It has been estimated that 75% of the genetic diversity of crop plants has been lost in the last century (Newton et al, 2010).

Molecular markers

Since the late 1960s, it has been possible to survey molecular markers across a plant 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 required the laborious assay of several 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 crops (Glaszman *et al.*, 2010).

QTL (quantitative trait loci) mapping has been used to determine the genetic bases of more complex traits. Many QTLs associated with quantitative disease resistance have been identified, for instance for blast in rice (*Oryza sativa*), fusarium head blight in wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*), powdery mildew in wheat, late blight in potato (*Solanum tuberosum*), and bacterial spot in tomato (*Solanum lycopersicum*), but the resistance mechanisms controlled by these QTLs are still largely unknown (Kushalappa & Gunnaiah, 2013). The QTLs identified contain several genes, often also unfavourable ones, making the transfer of QTLs to elite cultivars problematic (Kushalappa & Gunnaiah, 2013).

Genome sequencing

The costs of genome sequencing have come down rapidly, resulting in vast progress in this field (Furbank & Tester, 2011). Reference genomes have been published for model plant systems such as *Arabidopsis thaliana* and *Brachypodium distachyon*, and for a range of crops, including rice (*Oryza sativa*), maize (*Zea mays*), sorghum (*Sorghum bicolor*), soya bean (*Glycine max*), cucumber (*Cucumis sativus*) and tomato (*Solanum lycopersicum*) (Furbank & Tester, 2011; Morrell et al., 2012; Varshney et al., 2010; Varshney et al., 2011). A major achievement in this field has been the sequencing of the genome of bread wheat (*Triticum aestivum*, variety 'Chinese Spring', CS42), because of the large size of the genome, its repeat content and polyploid complexity (Brenchley et al. 2012).

The advent of next-generation sequencing (NGS) technologies has resulted in the construction of genome-wide variation maps for several major crops, such as rice, maize and soybean (Qi et al., 2013). In rice, for instance, 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. Advances made include the characterization of genomic variation and genetic structure in rice populations, the analysis of the relations between genomic variations with phenotypic traits, and the investigation of the origin of cultivated rice to further understand the history of rice domestication and breeding. Because rice is a model system for crop genomics studies, these advances are highly relevant for other crops as well (Huang et al., 2013).

Another example is cucumber (*Cucumis sativus*), where a map of genome variation encompassing about 3.6 million variants was generated by deep resequencing of 115 cucumber lines sampled from 3,342 accessions worldwide. A total of 112 putative domestication sweeps were identified, with one of these regions containing a gene involved in the loss of bitterness in fruits, an essential domestication trait of cucumber. Also, a natural genetic variant in a beta-carotene hydroxylase gene was discovered that could be used to breed cucumbers with enhanced nutritional value. The results suggest that fruit crops may have undergone narrower bottlenecks during domestication than grain crops did, and that they have more strongly reduced genetic diversity in their cultivated gene pools, underlining the importance of wild germplasm in fruit breeding (Qi et al., 2013).

A third example is the comparison of a draft genome sequence of *Brassica oleracea*, a morphologically very diverse species comprising important vegetable crops such as cauliflower, broccoli, cabbages, Brussels sprouts, kohlrabi and kales, with that of its sister species *Brassica rapa*. The comparison revealed numerous chromosome rearrangements and asymmetrical gene loss in duplicated genomic blocks, asymmetrical amplification of transposable elements, differential gene co-retention for specific pathways and variation in gene expression, including alternative splicing, among a large number of paralogous (i.e. arisen from a common ancestor sequence via duplication) and orthologous (i.e. arisen from a common ancestor sequence via speciation) genes. Genes related to the production of anticancer phytochemicals and morphological variations illustrated

consequences of genome duplication and gene divergence, imparting biochemical and morphological variation to *B. oleracea* (Liu et al., 2014).

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, including the peloric (*Linaria vulgaris*), colorless non-ripening (*cnr* – *Solanum lycopersicum*) and B0 (*Zea mays*) epialleles. Studies in the model species *Arabidopsis thaliana* have provided the richest source of known epialleles (Schmitz & Ecker, 2012). A study comparing genome wide DNA methylation among 10 *Arabidopsis thaliana* lines derived 30 generations ago from a common ancestor, indicated that DNA methylation in specific regions of the genome can fluctuate over relatively short timescales, and that these sites can be considered as going through recurrent cycles of forward and reverse epimutation, which is very different from what is found at the level of the genome sequence, where reverse mutations are exceedingly rare (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).

Genetic engineering

Klümper & Qaim (2014) carried out a meta-analysis of the agronomic and economic impacts of genetically modified (GM) crops. In total, 147 original studies were included. Studies were included when they had been built on primary data from farm surveys or field trials anywhere in the world, and when they reported impacts of GM soybean, maize, or cotton on crop yields, pesticide use, and/or farmer profits. It was found that, on average, GM technology adoption has reduced chemical pesticide use by 37%, increased crop yields by 22%, and increased farmer profits by 68%. Yield gains and pesticide reductions were larger for insect-resistant crops than for herbicide-tolerant crops, and yield and profit gains are higher in developing countries than in developed countries.

Crop improvement

Since the mid-1990s, progress in conventional plant breeding has slowed down, despite the impressive yield gains of the past. Part of the reason is that only the tip of the biodiversity iceberg has been explored and used (McCouch et al., 2013). Also, conventional plant breeding has been successful in favourable environments and those which can be made favourable through interventions, but it has been less successful in low-input environments, characterized by increased G × E interactions, or in organic farming systems (Newton et al, 2010). Resistance to biotic stresses can be qualitative or quantitative. Qualitative resistance (based on monogenic inheritance) has been successfully used in plant breeding, but quantitative stress responses (based on complex polygenic inheritance) have remained largely unexplored (Kushalappa & Gunnaiah, 2013).

According to Sinclair (2011), the enormous advances in understanding the physiology, biochemistry and molecular genetics of plant response to drought have had little or no impact in improving the intrinsic behaviour of plants and crop yield. The ‘bottom–up’ approach, starting at the molecular level, is very complex, in view of the large number of interacting biochemical networks and the damping of any simple molecular change when scaling-up the complex physiological hierarchy leading to yield. An alternative is the ‘top–down’ approach, starting with a whole-crop perspective to consider how the plant community may perform in the field across a range of environments. There are, however, only few examples where the top–down approach has been successfully applied to alter biochemical or physiological activity to increase crop yield. Sophisticated tools must be developed to monitor phenotype expression at the crop level to characterize variation among genotypes across a range of environments. Once stress-tolerant cultivars are developed, regional probability distributions describing yield response across years will be necessary. This information can then aid in identifying environmental conditions for positive and negative responses to genetic modification to guide farmer selection of stress-tolerant cultivars.

According to Fischer (2011), the contribution of genomics to wheat improvement has mostly remained limited to relatively simple traits, linked to a small number of genes. This is because the complexity of the path from gene through physiology to phenotype and yield, is by far more complex than the route from agronomic management to yield. The linking of relatively simple traits like time to flowering, and aluminium and salt tolerance, in each case to a small number of genes, is much facilitated by the development of molecular gene markers, and there is some progress on the functional basis of these links, and likely application in breeding. However, with more complex crop traits, such as potential yield, progress at the gene level is negligible.

Domestication

Genetic data have become a powerful resource for answering questions on the geographical origins of crops, the number of independent domestication events for a crop, the molecular changes underlying domestication traits, and the nature of artificial selection during domestication and subsequent crop 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 (Lenser & Theissen, 2013).

Conservation

Plant diversity is currently being lost at a 100- to 1000-fold higher rate than during the recent geological past, resulting in an associated decrease in ecosystem services, such as the provision of food, fuel, biochemicals and fibre (Li & Pritchard, 2009). The conservation of genetic, species and ecosystem biodiversity in the face of global climate change and increasing human pressure on the land is mandated by the Convention on Biological Diversity (McMahon et al., 2011).

In conservation, a distinction is made between in situ and ex situ conservation. In ex situ conservation, genetic resources are conserved outside of their natural habitat or cultivation territory, in genebanks. In ex situ conservation, the variability that has been collected remains static because the natural evolution 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 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). Landraces cannot be exclusively conserved under *ex situ* conditions, and complementary on-farm and in situ conservation methods are needed to ensure both conservation and evolution (Carvalho et al., 2013).

Extensive ex-situ germplasm 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).

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 used in food and feed production, can be stored for long after dehydration and cooling, but recalcitrant seeds, such as those of cacao (*Theobroma cacao*), are sensitive to drying (Li & Pritchard, 2009).

Climate change

Climate change is predicted to bring about increased temperatures across the world in the range of 1.6°C to as much as 6°C by 2050. Although rainfall is predicted to increase globally, some areas will receive less annual rainfall, while others may receive much more. The timing of rains and crop

growing periods will also change, and the frequency and duration of extreme weather events are predicted to increase. These predicted changes in climate are expected to have impact on phenology, species distribution and agriculture (Jarvis et al, 2008; McMahon et al., 2011).

Trends in genetic resources research

Molecular markers

Marker-assisted selection (MAS) is now used routinely for 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. A danger of the use of diagnostic markers is the tendency to restrict diversity to those alleles that can be diagnosed, while good prediction of plant performance with respect to complex traits requires sets of markers that collectively explain most, if not all, of the genetic variation. Whole-genome association mapping using genome-wide, evenly distributed SNP markers is an option to achieve this (Gebhardt, 2013).

Genome 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 plant genomes or sample entire transcriptomes more efficiently and economically and in greater depth than ever before. This makes it possible to sequence hundreds or even thousands of related genomes to sample genetic diversity, and to determine the genetic basis of trait variation and adaptation (Glaszman *et al.*, 2010; Furbank & Tester, 2011; Varshney *et al.*, 2009).

Schneeberger & Weigel (2011) indicate that 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.

Genome-wide association

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 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). Methods to map intraspecific genetic variation underlying phenotypic variation in plants have been developed (Bergelson & Roux, 2010).

Omics

It is now possible to generate omics datasets for many crop species, and, although the high costs of metabolomics still limit direct application in crop improvement, developments in omics technology are helping to elucidate the biological processes that determine gene effects. The assessment of RNA (transcriptomics), protein (proteomics) and metabolite (metabolomics) levels for parental lines can deliver information on genes in the target region associated with an mRNA, protein or metabolite shift linked to the trait of interest. For example, if the target trait is a QTL associated with drought tolerance, transcriptomics will reveal genes in the region that are up- or down-regulated in response to drought stress. Proteomics may give insight into changes in protein abundance or modification in response to the stress. Metabolomics data might reveal that the region is associated with a major change in levels of an osmoprotectant, which suggests that those genes are involved in its biosynthesis (Langridge & Fleury, 2011).

Phenotyping

The phenotype is the set of structural, physiological, and performance-related traits of a genotype in a given environment. Phenotyping is the act of determining the quantitative or qualitative values of these traits. Through phenotyping, the prediction of phenotypes on the basis of genomic data becomes possible. 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 plant biomass for analysis. However, great advances are being made in the field of phenotyping, as high-throughput, non-destructive methods are being developed (Furbank & Tester, 2011; Roy *et al.*, 2011). Imaging and image processing are now major tools for phenotypic trait measurement (Dhondt et al., 2013). Infrared thermography can be used to assess the capacity of seedlings to maintain stomatal conductance under drought stress, chlorophyll fluorescence to assess the maintenance of photosynthetic function under drought stress, and digital growth analysis to monitor the progress of disease symptoms (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). As a result of these new developments, field evaluation of plant performance can be much faster, with a more dynamic, whole-lifecycle measurement, less dependent on periodic destructive sampling. 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).

'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). Flood et al. (2011) recommend that phenotyping systems should be developed that involve movement of the equipment and not the plants. Repeated measurements are to be made throughout the life cycle of the plant, and measurements that have a minimal effect on the plant are the most suited for accurate assessment of plant phenotypes.

Epigenetics

The understanding of the molecular mechanisms of epigenetic inheritance is rapidly growing, and it is now possible to profile the epigenome at high resolution. However, to determine if epigenetic variation can effectively be captured in breeding 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).

Molecular breeding and genetic engineering

Useful genes or alleles can be transferred into genotypes of the same species through molecular breeding (MB). In the MB approach, quantitative trait loci (QTLs) are identified for traits of interest, such as tolerance to abiotic stresses. Until recently, QTLs were identified by linkage mapping, but association genetics have started to supplement these efforts. Compared with linkage mapping, association mapping is a high-resolution and less expensive approach. (Varshney *et al.*, 2011). By using genetic engineering, useful genes or alleles can be transferred across different species (Varshney *et al.*, 2011).

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).

Crop improvement

Although much progress has been made in the field of biotechnology, a major challenge is the widening gap between the development of new technologies and their deployment in crop improvement. For instance, many genes for different stresses have been cloned and characterized, but there are no reports of a released transgenic variety for drought tolerance (Varshney *et al.* (2011).

Changes in the climatic and environmental conditions under which crops are grown have resulted in the appearance of new diseases, while genetic changes within the pathogen have resulted in the loss of previously effective sources of resistance. Research into plant–pathogen interactions for crop improvement currently falls into two broad approaches. The first approach focuses on the identification of the pathogen factors which are conserved across isolates, essential for pathogen survival, and unable to withstand modification. The plant resistance genes that recognize these conserved pathogen factors would theoretically have a greater possibility of remaining effective over time because of these evolutionary constraints on the pathogen factor. The second approach involves the identification and study of QTL for partial and often durable resistance. Expressing resistant alleles of these genes prevents the pathogen from exploiting its host or activates general defence mechanisms used by the plant to restrict pathogen growth. Through the cloning of these QTL it is possible to determine the mechanisms of resistance and thereby accumulate resistance genes with contrasting and complementary mechanisms that can be used in conjunction to provide long-term pathogen resistance. The informed deployment of major, race-specific and partial, race-nonspecific resistance, either by conventional breeding or transgenic approaches, will enable the production of crop varieties with effective resistance without impacting on other agronomically important crop traits (Boyd *et al.*, 2013).

According to Kushalappa & Gunnaiah (2013), a forward genetics approach (seeking to find the genetic basis of a phenotype or trait) is better suited for the discovery of genes conferring resistance to biotic stresses than reverse genetics (seeking to find what phenotypes arise as a result of particular genetic sequences), because resistance to biotic stresses can be governed by hundreds of genes.

In breeding for resistance, one should not solely look at resistance effects of alleles, but also at other effects on plant performance. In *Arabidopsis thaliana* it was found that allelic diversity at a single locus underpins differences in both vegetative growth and resistance to microbial infection and herbivory among natural strains. A hyperactive ACD6 allele strongly enhances resistance to a broad range of pathogens, but at the same time slows the production of new leaves and greatly reduces the biomass of mature leaves. Thus, accessions with this allele seem to pursue an alternative strategy, being small, but well protected, compared to other strains that are larger, but less well prepared to combat pathogens. This allele segregates at intermediate frequency both throughout the worldwide range of *A. thaliana* and within local populations, consistent with this allele providing substantial fitness benefits despite its marked impact on growth (Todesco *et al.*, 2010).

Natural genetic variation in photosynthesis is present in both crop and wild species. Modern genetics and high-throughput phenotyping make possible the detailed investigation of the natural genetic

variation in photosynthesis in a wide range of species. Incorporation of the results into breeding programs should lead to an increase in crop photosynthesis, which will ultimately lead to increased plant yields (Flood et al., 2011).

Crop diversification

Another field receiving much attention is agrobiodiversity, the diversity of agricultural systems from genes to varieties and crop species, and from farming methods to landscape composition. In many regions, agrobiodiversity is severely threatened. Pautasso et al. (2013) argue that the circulation of seed among farmers is central to agrobiodiversity conservation and dynamics, and, while agrobiodiversity conservation has received much attention over the last decades, methods to study the role of seed exchange networks have only recently begun to be considered.

Ruiz et al. (2014) argue that climate change is rapidly degrading the conditions for crop production, and that stress-tolerant species and genotypes must be identified for future agriculture. Many stress-tolerant species are traditional crops cultivated by farmers at a local scale, of which quinoa (*Chenopodium quinoa* Willd.), a seed crop traditionally produced in Bolivia and Peru, is a good example. Quinoa seeds are nutritious, and the crop remains productive on poor soils and under conditions of water shortage and high salinity. It can be adapted to diverse agroecological conditions worldwide, and its cultivation is now spreading to many other countries.

Another example of increased interest in non-traditional crops and crop varieties is emmer wheat (*Triticum dicoccon*). This crop has been cultivated during seven millennia in the Middle-East, Central and West Asia, and Europe, but has been largely replaced by hull-less species and is now a minor crop, with the exception of some countries like India, Ethiopia and Yemen, where its grain is used for preparing traditional foods. Nutritional qualities and specific taste and flavour of emmer wheat products have led to a recent development of the cultivation in some European countries, e.g. in Italy. Emmer wheat also possesses valuable traits of resistance to pests and diseases and tolerance to abiotic stresses and is increasingly used as a reservoir of useful genes in wheat breeding (Zaharieva et al., 2010).

Increased insight in the mechanisms of day length sensitivity may be useful for breeding of crop genotypes for new geographic regions at different latitudes. In potato, for instance, a regulator underlying a major-effect quantitative trait locus for plant maturity and initiation of tuber development has been identified. The gene regulates tuberization and plant life cycle length, by acting as a mediator between the circadian clock and a tuberization signal. Natural allelic variants evade post-translational light regulation, allowing cultivation outside the geographical centre of origin of potato. Potato breeders can exploit the naturally occurring variation in tuberization onset and life cycle length, allowing varietal breeding for different latitudes (Kloosterman et al., 2013).

Domestication

QTL mapping was one of the first techniques applied to understand the genetic basis of domestication traits. The limited number of QTL related to domestication traits suggest that many

plants have been domesticated via changes at relatively few loci. Despite the large amount of information obtained from QTL mapping and map-based cloning, the approach has the problem that it tends to investigate what we already think is important. In contrast, genomic scans, where diversity at molecular markers in wild and domesticated populations is compared to identify reductions in variation consistent with selection, are unbiased about the type of locus that might be identified as being important in domestication. Genome scans are becoming easier to implement because of the relative accessibility of genomic data from which markers can be developed. Genome scans have mainly been implemented to identify the proportion of genes under selection during domestication, and have provided lists of candidate domestication genes for further consideration. Plant biologists are also beginning to pursue association mapping. High-throughput resequencing techniques offer a complement to or faster implementation of the approaches mentioned above (Gross & Olsen, 2010).

The insight that during plant domestication the causal mutations for convergent changes in key traits are likely to be located in particular genes, may contribute to defining candidate genes for genetic improvement during the domestication of new plant species. Rather than starting domestication of wild plant species from scratch, promising hotspot genes might be targeted via marker-assisted breeding or transgenic technology to generate desired phenotypic changes. Such efforts may help to increase the range of arable crops available (Lenser & Theissen, 2013).

Genetic diversity

Genetic diversity has become a renewed focus, as its loss could make populations and species less able to adapt to ongoing environmental changes (Jump *et al.*, 2009). Without genetic variation there can be no natural or artificial selection, no genetic drift and no adaptation of species to changing conditions (Flood *et al.*, 2011). Alleles that may be of future benefit but confer no current advantage are likely to be at low frequency within the population relative to those that are currently advantageous, and it is these low-frequency alleles that are most likely to be lost during when population size is reduced. Therefore, the maintenance of genetic diversity within natural populations is considered a key element of maximizing their chances of survival (Jump *et al.*, 2009).

According to Roy *et al.* (2011), the variation for abiotic stress tolerance within crops has decreased due to the narrowing of genetic diversity within elite germplasm during the course of plant breeding. It has been estimated that only about 15% and 40% of the available genetic variation has been captured in modern wheat and barley varieties, respectively. There is, however, still wide genetic diversity in landraces and wild relatives of crops. This is, for instance, visible in the comparison of molecular marker diversity studies of cultivated barley (*Hordeum vulgare*) and wild barley (*H. vulgare* ssp. *spontaneum*). The latter is a possible source for abiotic stress tolerance traits, and because of the close genetic similarity to current cultivars, these traits can be introduced into commercial breeding lines. Genetic approaches to elucidating the molecular basis to abiotic stress tolerance in crops are becoming more easily achievable. A powerful approach is to identify naturally occurring variation of abiotic stress tolerance in varieties, landraces and wild relatives of a crop, and to study the traits that contribute to tolerance. Once the molecular bases of traits contributing to tolerance have been identified using this forward genetic approach, marker-assisted breeding and genetic

modification technologies can be used to introduce these traits into current, high yielding cultivars (Roy *et al.*, 2011).

Landraces and crop wild relatives (CWR) are considered necessary for broadening the limited genetic diversity of cereals, and the development of new cultivars adapted to changing environments (Carvalho *et al.*, 2013; Feuillet *et al.*, 2008; Newton *et al.*, 2010). Traits for biotic stress tolerance, such as disease and pest resistance have already been introgressed from near wild relatives of wheat, such as *Triticum urartu*, *T. monococcum*, *T. tauschii* and *Aegilops speltoides*, and from *Oryza nivara* for rice (Roy *et al.*, 2011). In a review of the current status and prospects for cereal landraces, the following points were made: 1) landraces are important resources for the development of future crops; 2) many germplasm collections of landraces of the major cereals are available, and these have been characterised to variable degrees and in different ways; 3) much of this germplasm is maintained both in long-term storage and on farm; 4) developments in genotyping technologies make the variation available in landraces ever more accessible; 5) landraces are a potential source of favourable traits such as improved nutrient use efficiency and resistance to pathogens; 6) single-gene traits are generally easily transferred from landraces to modern cultivars, but most of the desirable traits of landraces are complex and difficult to express in different genetic backgrounds; 7) participatory plant breeding has been more successful for landrace improvement in stress-prone environments than the approach used in high input breeding programmes; 8) changes in seed trade legislation are needed to promote the exploitation of diversity in landraces (Newton *et al.*, 2010).

The identification of favourable alleles in a crop relative or wild species can be complicated because of their often poor agronomic performance. In these cases, use can be made of introgression lines (ILs), series of lines, each carrying a chromosome segment from a wild relative genotype in the background of an elite genotype. Such lines enable assessment of the phenotypic impact of the wild species genome in parts while keeping the genetic background of the cultivated recipient genotype. This strategy has been widely utilized for introgression of favourable alleles from wild relatives in tomato, rice, wheat and barley (Varshney *et al.*, 2010).

Adaptation to climate change

Understanding how species and ecosystems respond to climate change has become a major focus of ecology and conservation biology. Modelling approaches provide important tools for making future projections, but, according to McMahon *et al.* (2011), current models of the climate-biosphere interface are too simplistic. To improve understanding of biodiversity-climate relations and advance the predictive ability of climate-biosphere models, they propose: (1) the creation of on-line repositories for data from existing monitoring sites; (2) the standardization of protocols for measuring biodiversity; (3) the establishment of new monitoring sites filling geographical and ecological gaps; (4) the development and application of methodological and statistical toolkits; (5) the quantification of the sensitivity of individual species to climate; (6) meta-analyses of laboratory and field observations of genetic diversity; (7) the expansion of trait databases; (8) improved parameterizations and more realistic PFT (plant functional type) classifications in global vegetation models; (9) the incorporation of demographic processes, plasticity, genetic variability and dispersal

within global vegetation models; (10) the application of hierarchical modelling approaches using multiple sources of data, to address biodiversity responses to environmental changes at different temporal and spatial scales (McMahon et al., 2011).

According to Kramer & Havens (2009), climate change, rapidly increasing habitat fragmentation, degradation, and globalization of pests and diseases mean that that not only whole-habitat restoration, but also single-species reintroduction and population augmentation will be increasingly important conservation tools. Plant conservation genetics provides tools to guide conservation and restoration efforts, measure and monitor success, and minimize extinction risk by conserving species as dynamic entities capable of evolving in the face of changing conditions. Techniques have been developed for quantifying population genetic parameters such as inbreeding, within- population genetic diversity, among-population genetic divergence, local adaptation and gene flow.

The field of 'landscape genetics', which investigates how landscape elements and environmental factors influence the spatial distribution of genetic variation, is receiving increasing attention. Main topics are the evaluation of the effects of landscape and environment on migration, dispersal and gene flow, and the exploration of the interaction between environment and adaptive genetic variation in natural populations and individuals ('landscape genomics'). It tries to identify molecular markers whose changes in allele frequencies are correlated with environmental factors potentially acting as selective pressures and enforcing directional natural selection (Holderegger et al, 2010).

Next generation sequencing (NGS) offers the opportunity to perform genomics studies on many ecologically interesting species without the requirement of a closely related genetic model organism. Thus, NGS makes it feasible to identify genetic loci responsible for adaptive evolution in non-model organisms. By integrating biogeography, field experimentation and long-term life history studies with new genomics tools, new theories can be tested and the understanding of adaptation greatly enhanced (Stapley et al, 2010).

Ex-situ conservation

The core collection concept (i.e. a limited set of accessions representing, with minimum repetitiveness, the genetic diversity of a crop species and its wild relatives) is being developed to increase the efficiency of characterisation and utilisation of collections stored in the genebanks, while preserving as much as possible the genetic diversity of the entire collection (Glaszman *et al.*, 2010; Odong et al., 2013). Definition of clear criteria for evaluation of the quality of core collections is a prerequisite for selecting high-quality core collections, but examination of the methods used in literature has shown that there are no clear guidelines on the evaluation criteria. Odong et al. (2013) recommend the use of criteria based on genetic distances between accessions, because these not only allow the simultaneous evaluation of all variables describing the accessions, but also provide intuitive and interpretable criteria, as compared with the univariate criteria generally used for the evaluation of core collections.

To facilitate material sharing within the scientific community, the 'core collection' concept has been broadened into the 'core reference set': a set of genetic stocks representative of the genetic

resources of the crop that can be used by the scientific community as a reference for characterization of its biological diversity. Core collections derived from several collections can be reduced to core reference sets of 50–500 accessions. Information concerning accessions, which ideally is included in the passport information, is essential for making these reference sets (Glaszman *et al.*, 2010).

For a range of crops, a systematic molecular characterization of germplasm from various collections has been conducted to define core reference sets. About 32,000 accessions, representing global composite collections of 21 species have been genotyped. The genotypic data were used to make core reference sets for each species, consisting of 96–359 accessions. The aim of the core reference sets was to capture about 80% of the molecular diversity present in the sampled germplasm. The next step will be to gather robust phenotypic data sets, assembled from trials conducted under a range of agro-climatic conditions and genome-wide high-density marker data on these core reference sets. These data are considered essential for the identification of phenotype/genotype associations that might deliver marker tags for agronomic traits (Varshney *et al.*, 2010).

Cryopreservation

According to Li & Pritchard (2009), it would be best to use cryopreservation (ultra-cold storage) for the long-term conservation of both recalcitrant seeds and orthodox seeds, because of evidence of less than expected longevity at conventional seed bank temperatures, innovations in the cryopreservation of species with recalcitrant seeds, and economic considerations. For instance, cryogenic storage did prolong the shelf life of lettuce (*Lactuca*) seeds with projected half-lives up to 3400 years, which is up to 20 times greater than that predicted for that species in a conventional seed bank at –20 °C.

At present, cryopreservation is, together with storage of tubers and *in vitro* culture, one of the conservation methods of cultivated potato accessions, which cannot be maintained by true potato seeds, because potato is highly heterozygous and the seeds are not true to type. Cryopreserved collections of potato shoot tips are held Germany, the Czech Republic, South Korea, Peru and the United States. Cryopreservation can also be used for virus elimination (cryotherapy) in potato and other crops. Improvement in the development and successful application of potato shoot tip cryopreservation protocols have been made over the last 30 years, but none of the cryopreservation techniques can guarantee the conservation of all genotypes (Kaczmarczyk *et al.*, 2011).

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

Conservation of genetic resources

- The availability of molecular markers and genome sequencing techniques make it possible to adjust the size, the representativeness and the general quality of ‘core’ collections of genebanks (Glaszman *et al.*, 2010).
- For technical and economical reasons, the application of cryopreservation for the long-term conservation of recalcitrant seeds as well as orthodox seeds may become more important (Li & Pritchard, 2009).
- To strengthen the complementarity of on-farm and in situ conservation, genebanks should not only serve breeders searching for accessions with specific traits, but also new users, including farmers, nature museums and growers of landraces or conservation varieties, who do not focus exclusively on crop germplasm traits, but are also interested in the landraces, their history and the traditional knowledge on their management and use. Genebanks should make efforts to share germplasm information to promote the use of conserved materials by interested entities (Carvalho *et al.*, 2013).
- National coordination of public and private conservation efforts is needed to prevent loss of genetic diversity and avoid duplication of efforts. An example is the German National Fruit Genebank, a decentralised network established in 2009 and aimed at the coordination of various germplasm collections. In the past five years, five fruit specific networks for apple, cherry, strawberry, plum and *Rubus* were successfully established. Activities include phenotypical evaluation and DNA-fingerprinting (Flachowsky & Hanke, 2014).
- It has been argued that the conservation of wild relatives of cultivated plants can be justified on the basis of their ‘option value’: their potential for the future exploitation of genetic diversity, even if their present economic value is unrecognized. Actual quantification of this option value is rare, however, and the assumption that all genetic variation in natural populations is inherently valuable is an essential extension of this concept (Jump *et al.*, 2009).
- Walck & Dixon (2009) argue that seed banks must collect wild species and condition their holdings in such a way that they can thrive in landscapes transformed by climate change. They give the example of the Millennium Seed Bank Partnership (MSBP) of the Royal Botanic Gardens in Kew, which has stored seeds from 10% of the world’s known wild flora. The primary aim should be to harvest as much genetic diversity as possible, so that collections are more likely to contain plants with traits suited to a changed climate.

Use of genetic resources

- As genome sequencing has become cheaper and more common, and reference genomic sequences have become available for a range of plants, the future of crop improvement will revolve around the comparison of individual plant genomes (‘comparative genomics’) (Morrell *et al.*, 2012).
- Understanding the molecular basis of the variation of agronomically important phenotypes will open new avenues for plant breeding, in a similar way as gene diagnosis in molecular

medicine holds the promise of more effective, personalized therapies for human diseases (Gebhardt, 2013).

- The completion of reference genome sequences for various important crops, and the ability to perform high-throughput resequencing provide opportunities for better understanding of plant domestication and accelerating crop improvement (Morrell et al., 2012).
- The application of new phenotyping 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)
- The increasing ability to connect phenotypes and genotypes will make more rapid selection and breeding possible (Morrell et al., 2012).
- Integration of genome and functional omics data with genetic and phenotypic information will make it possible to identify genes and pathways responsible for important agronomic phenotypes (Langridge & Fleury, 2011).
- New methods for improving the use of wild germplasm, making use of gene discovery, improved technologies for genetics and breeding, and a better understanding of the factors limiting practical exploitation of exotic germplasm will change existing practices and accelerate the development of new strategies for efficient germplasm utilization (Feuillet et al., 2008).

From conservation to use

McCouch et al. (2013) propose three steps to apply the recent advances in genetic resources research to ensure better utilization of gene bank materials:

1. Sequence information must be obtained from the genomes of all non-duplicate plant samples in the world's gene banks that are available under the terms and conditions of the ITPGRFA. This 'fingerprint' for each plant will serve as the basis for assessing genetic relationships and selecting subsets of material for further investigation. Sequence data should also be linked with conventional 'passport information' about collection locality and original environment.
2. The phenotypes of gene-bank accessions must be evaluated. Though it will not be possible to evaluate all gene-bank accessions in all relevant environments, even with the advent of high-throughput phenotyping technologies, it is expected that the use of sequence data in combination with phenotypic, geographical and ecological information will enable researchers to target field experiments strategically and to develop models that can predict plant performance.
3. An internationally accessible informatics infrastructure, linking seeds and genetic stocks directly to passport, genomic and phenotypic information, must be created to catalogue the diversity in the world's seed collections. Furthermore, the results from genomics and agronomic research must be connected to the communities creating new crop varieties, and scientists in both the public and the private sectors must work together to provide seeds and plants to farmers and commercial plant breeders for further crossing and testing in different environments.

Climate change adaptation

- Anticipated changes in climate and its variability, especially more extreme temperatures and changes in rainfall patterns, are expected to make crop improvement even more crucial for food production (Varshney et al., 2011).
- Climate change will have implications for genebank managers, genetic resource users, in situ biodiversity management and the maintenance of ecosystem services. Important issues are:
 - an increased need for maintaining collections of wild species, including crop wild relatives, because of an increased likelihood of extinction of narrowly adapted and endemic species;
 - new and increased demands on germplasm held in genebanks for adapting agriculture to climate change, including the need to screening for different characters;
 - a revision of breeding strategies and priorities, to make products of breeding programmes relevant to the challenges the world will be facing upon release of these improved materials;
 - the need to review and strengthen policies for promoting seed systems that encourage the maintenance of genetic diversity;
 - increased demand on international policy to facilitate access to genetic resources due to increases in interdependency brought about by shifts in climate zones globally (Jarvis et al., 2008).

A comment: very few words on minor crops and neglected species. Few word also on fruit species.