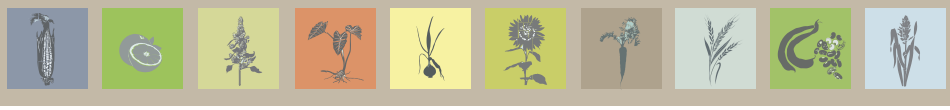

Appendix 3

The state-of-the-art: methodologies and technologies for the identification, conservation and use of plant genetic resources for food and agriculture

A3.1 Introduction

The magnitude and structure of the genetic diversity of a population determine the ability of that population to adapt to its environment through natural selection. This is because when genetic diversity is low, the possible combinations of genes that can confer fitness, and hence, adaptation to variations in environmental conditions are reduced, decreasing the probability of successful individuals arising in the population. Thus, a population in nature (or managed in a protected area) needs to have adequate genetic diversity to sustain its continued existence in the face of the continually changing biotic and abiotic components of its ecosystem.

A parallel scenario depicting natural populations takes place in crop improvement programmes with regard to available heritable variation within the germplasm. Breeders seek and recombine genetic variability in their breeding populations and screen for desired traits or characteristics that enable the crop to be successful in target environments or against targeted pests or pathogens. Breeders therefore need access to adequate genetic diversity for success in breeding programmes.

Underlying these scenarios (variations in nature and in germplasm collections for breeding), superficially conceptualized as 'diversity is good' in nature and in crop improvement programmes, there are many complicated issues. A fundamental imperative is the need to distinguish phenotypic diversity (net result of the interaction between both heritable and non-heritable components of variation) from genetic (heritable) diversity. Other issues relate to strategies for finding genetic diversity, maintaining, measuring and monitoring it, as well as devising mechanisms for exploiting it most efficiently. The processes of both scenarios can be further complicated by the biology of the species which encompass its breeding system, whether it is annual or perennial, the ploidy levels, and its ecological tolerances. The extent to which these factors are understood therefore has an impact on the ability of researchers to develop breeding or conservation strategies for the species in question.

There are also non-biological issues that can complicate management practices for both natural populations and breeding materials; these include organizational, policy, legal and economic issues. There are also issues of scale - ranging from national through regional to global - with respect to collaboration, incentives and efficiencies that facilitate conservation and use of genetic resources.

The objective of this Appendix is to summarize primarily the status of scientific knowledge, practices and technologies pertaining to genetic diversity that have arisen since the first SoW report published in 1998 which had a similar summary presented in Annex 1. The status of the social enabling environment also will be addresses as its components impact directly on national capacities for the conservation and use of genetic resources.

Annex 1 to the first SoW report clearly set out the importance of genetic diversity in the context of both the conservation and use of plant germplasm; the contrasts between qualitative and quantitative genetic variation and the different emphasis placed on these by curators and users of genetic resources; the means and techniques for conservation; the various breeding strategies and their roles and challenges with respect to breeding goals and finally, the legal and economic issues that can promote or deter conservation and use of genetic resources. This Appendix will not repeat that information but will focus on new developments since the publication of first SoW report.

A3.2 Advances in knowledge of genetics relevant to PGRFA conservation and use

The principal advances in the understanding and application of heredity in the management of PGRFA in the past 12 years emanate from the immense strides that have been made in molecular biology during the period especially with regard to genomics, the study of the totality of an individual's genetic makeup (genome). With the ability to sequence whole genomes in a timely and cost-efficient manner, the period has been characterized by an ever increasing volume of publicly accessible DNA, gene and protein sequence information. This has been complemented by the incredible advancements in the scopes for

APPENDIX 3

both the generation and analysis of data to degrees that were unfathomable a couple of decades ago. This paradigm contrasts sharply with the significantly narrower scope of the understanding of heredity that had hitherto been possible using classical genetics in isolation.

Genomics and the related fields of proteomics (the study of proteins), metabolomics (the study of metabolites) and the more recent phenomics (study of phenotypes in relation to genomics) have developed from the confluence of classical genetics, automated laboratory tools for generating molecular data, and methods of information management, especially bioinformatics. Advances in taxonomy and systematics, largely attributable to refined information arising from the use of molecular biology approaches in genome characterizations, have led to better understanding of the structure of genepools, relationships within and between taxonomic groupings and, in some cases, to the reversal of hitherto assigned taxonomic classifications. These novel fields of the biological sciences have direct implications for germplasm management (e.g. in the designation of core collections) and in determining the needs for further collections of genetic resources. Furthermore, molecular data, being environment-neutral, are particularly useful in devising crop improvement strategies including pre-breeding activities as they are particularly suited for trawling through the genepool for new sources of gene alleles.

The contributions of genomics and the other – omics to basic biology have been equally profound as their judicious applications continue to lead to better understanding of metabolic processes, their key components and pathways. This allows researchers to ultimately achieve greater precision in the identification of genes and their alleles for use in crop improvement. Quite importantly also, molecular biology techniques are permitting better and more precise understanding of adaptation and evolution making it possible therefore to delineate reliably neutral genetic diversity from adaptive genetic diversity, and the role different markers can play in identifying and using genetic diversity.

With the current pervasive ability to use appropriate molecular approaches to identify genome segments that discriminate between individuals (known as molecular markers) and apply statistical algorithms to identify precisely the genome locations of these "landmarks", molecular markers are now the tools of choice for both tracing the inheritance of target regions of genomes in plant breeding programmes (marker assisted selection) and for characterizing germplasm collections. The routine use of molecular tools in analysing germplasm collections in PGRFA management, will lead to enhanced efficiencies in the management of collections. Advantages would include greater ease in the identification and elimination of duplicates (or other levels of redundancies) in germplasm collections and at the same facilitate the creation of core collections.

Another area of PGRFA management that has been profoundly impacted by the applications of molecular biology techniques is population genetics. This is on account of the widespread use of molecular data in the study of populations (diversity and structure). The heavy reliance on molecular data in population genetics has spawned the neologism, population genomics. It is becoming commonplace, for instance, to identify specific loci under natural selection and thus of adaptive importance merely by sampling at a population level . It has also become quite routine to track gene expression (based on transcript profiling; or transcriptomics), even at tissue levels, under different environmental influences (biotic and abiotic) and under a time series regimen. Such a strategy, in addition to permitting the identification of genes that modulate particular phenotypic expression, also leads to the elucidation of the functions of genes and their interactions with other genes. The refined understanding of genes and their functions and the tools being generated in this manner will prove invaluable as efforts are invested in crop improvement programmes to develop varieties that will thrive in spite of the extreme climatic conditions expected as consequences of global climate change and variations.

One specific example of the striking contrast between what was considered possible in 1995 and what is possible now comes from Annex 1 of the first SoW report, where it was stated that the direct application of DNA sequencing was more useful in the identification of a gene or genes than for analysing a complete genotype. The conclusion at the time

was that there was only "a very limited possibility to sample many variants for PGRFA characterization". Today, with improvements in technology, especially with regard to high throughput platforms for DNA extraction, amplifying and visualizing DNA (and RNA) fragments, with sequencing DNA fragments (and whole genomes), significantly enhanced computing capacity (data storage and analysis) and the suite of custom analytical software, it has become routine to characterize large numbers of accessions for polymorphisms (differences in sequences) at thousands of DNA loci across the genome.[1]

Another area of great progress since 1995 is the identification of conserved linear order of genes on chromosomes, a phenomenon known as synteny. This has been established not only between closely related species but also with more distant taxons and even between species that differ by large differences in genome sizes. Synteny has now been documented for many taxons in such families as the *Fabaceae, Poaceae*, *Solanaceae* and *Brassicaceae*. These findings have provided the impetus for the investment of a significant amount of effort in comparative genomics with the goal of leveraging gene sequence information from model species for the identification of genes in taxons other than the model species. Microsynteny (similarity between taxons in the ordering of nucleotide sequences along the same chromosome) has only become measurable with the availability of copious amounts of genome sequence data that are now available in the public domain. The demonstrated instances of macrosynteny (similarity between taxons in the ordering of large numbers of genes along the same chromosome) suggest therefore that there are ancestral genomic segments conserved across many taxons. The implication is that molecular markers identified in those segments could be used in genone characterizations even across the different taxons. Of course, the utility of synteny will always be subject to the influences of chromosome rearrangements.[2]

In general, the increased understanding of, and the enhanced ability to study, genetic diversity within species, populations and genepools with respect to distribution and structure have been key developments since the first SoW report. It is now established that nucleotide sequence polymorphism provide valuable information for understanding and deploying genetic diversity for crop improvement. The utility of these polymorphisms, as molecular markers, is even enhanced when the polymorphism occurs within a target gene (yielding functional markers). Representative examples are presented below.

A3.3 Advances in biotechnology relevant to PGRFA conservation and use

The initial applications of molecular biology in the characterization of plant genomes included single gene sequencing, the development and use of restriction fragment length polymorphism (RFLP) markers and low-density dotblot types of DNA arrays (or northern blots). The state of knowledge initially also favoured the one-gene, one-phenotype paradigm. All of these were in place at the time of the first SoW report but were quickly supplanted by whole-genome sequencing, widespread use of molecular genetic markers based on PCR, the single nucleotide polymorphism (SNP) markers, and medium-density arrays (for gene discovery and function elucidation). Currently, comparative whole genome sequencing (using multiple related species), extremely high-density genotyping (involving re-sequencing of individuals) and whole-genome arrays for monitoring genome-wide transcription, alternative (or differential) splicing, are but a few of the examples of new molecular biology tools that are revolutionizing the depth and breadth of genome analysis of crop germplasm. Also, the one-gene, one-phenotype paradigm is giving way to a new philosophy of a dynamic genome responding globally to developmental pathways and environmental signals.[3]

Speed, scale and size are the parameters that are most impacted upward by technological advances. Speed or throughput has increased significantly in many diverse activities ranging from DNA extraction, through polymerase chain reactions to microarray transcriptome profiling. Scale of approach has also increased significantly as exemplified by the numbers of molecular markers that can be used to assay individual DNA samples simultaneously; the numbers

APPENDIX 3

of progeny from mutation events or recombination events that can be screened for low probability responses; or the numbers of samples that can be handled simultaneously with robotics. In general, the manageable sizes and scopes of many activities and assays have increased significantly; the number of nucleotide base pairs that can be amplified or sequenced, the extent of coverage of genome in any assay, the density of molecular markers (number of markers per centiMorgan) on a molecular genetic linkage map, the lengths of fragments inserted into bacterial artificial chromosome (BAC) libraries, and lengths of contigs that can be assembled while comparing sequence data are a few examples of such increases.

Interestingly, increases in scope and scale have progressed in tandem with concomitant enhancements in efficiency levels as costs and time per unit data point have been reduced significantly; equipment and supplies have become cheaper and therefore lent themselves to wider access to research facilities of varying levels of budget, infrastructure and human resource capacities. However, it is also noteworthy that the net result of the increases in speed, scale and size and decreases in cost and time itself is a new kind of bottleneck - massive amounts of data that must be stored, processed, analysed, interpreted and displayed. Developments in computing hardware and software are addressing this bottleneck very satisfactorily as researchers usually have a wide array of choices in information technology paraphernalia for managing molecular data.

Genome sequencing has also continued apace with the aforementioned advances in the science of molecular biology and innovations in the ancillary technology platforms. The first fully sequenced plant genome was *Arabidopsis thaliana* in 2000.[4] This species has a small genome and has become a model plant species for research in biology and genetics. The second plant species sequenced was a crop species, rice - the sequences of two different genotypes of rice were published in 2002 (*Oryza sativa* indica[5] and *O. sativa* japonica[6]). Also, the first tree genome sequenced was a species of poplar (*Populus trichocarpa*) in 2006.[7] Also in 2006, the draft sequence of the genome of *Medicago truncatula* was published.[8] This species provides a genome model for legumes. The other crop genomes that have been sequenced were those of sorghum (*Sorghum bicolor*), grape (*Vitis vinifera*) and papaya (*Carica papaya*); all three sequences were published in 2007.[9] In 2008, draft sequences for soybean (*Glycine max*)[10] and *Arabidopsis lyrata*[11] were published. *Arabidopsis lyrata* is a close relative of *A. thaliana*, but with a larger genome. Most recently (2009), the sequences for *Brachypodium distachyon*[12] (a new model species for temperate grasses and herbaceous energy crops) and for maize (*Zea mays*)[13] were published. Box A3.1 identifies several other higher plant species for which genome sequencing projects are underway (as of early 2010).[14] In addition to full genome sequencing, large amounts of sequence data are available for many plant species; these result from the sequencing of sizeable fragments of their genomes (e.g. the sequencing of BAC libraries or whole chromosomes). Examples of crop species (or species closely related to crops) with substantial deposits of DNA sequences in publicly accessible databases are *Brassica rapa*, *Carica papaya*,

Box A3.1
List of plant species with ongoing genome sequencing projects in 2010[15]

Amaranthus tuberculatus, Aquilegia coerulea, A. formosa, Arabadopsis arenosa, Arundo donax, Beta vulgaris, Brassica napus, B. oleracea, B. rapa, Capsella rubella, Chlorophytum borivilianum, Citrus sinensis, C. trifoliata, Cucumis sativus, Dioscorea alata, Eucalyptus grandis, Gossypium hirsutum, Glycyrrhiza uralensis, Hordeum vulgare, Jatropha curcas, J. tanjorensis, Lotus japonicus, Madhuca indica, Malus x domestica, Manihot esculenta, Millettia pinnata, Mimulus guttatus, Miscanthus sinensis, Musa acuminata, Nicotiana benthamiana, N. tabacum, Oryza barthii, Panicum virgatum, Phoenix dactylifera, Pinus taeda, Ricinus communis, Solanum demissum, S. lycopersicum, S. phureja, S. pimpinellifolium, S. tuberosum, Theobroma cacao, Triphysaria versicolor, Triticum aestivum, Vigna radiata and *Zostera marina.*

Gossypium hirsutum, Glycine max, Hordeum vulgare, Lotus japonicus, Medicago truncatula, Sorghum bicolor, Solanum lycopersicum, Triticum aestivum, Vitis vinifera and *Zea mays*.[16] Another source of sequence information is the collections of expressed sequence tags (ESTs, produced by the sequencing of complimentary DNA or cDNA libraries) that are being generated for many crops. Maize, wheat, rice, barley, soybean and *Arabidopsis* have the largest collections of EST sequences for plants; over one million ESTs have been published for each of these plant species.[17]

The development of new DNA sequencing technology[18] has been driven by both publicly and privately funded research and development activities in human genomics. Lagging behind, but benefiting greatly nonetheless from progress being made in human genomics, is the application of these technologies to plant research in general, and more specifically, to research relevant to crop improvement, plant evolution and PGR conservation. Steady advances are being made in both the hardware and software for genome sequencing[19] and it is envisaged that in the near future, whole genome sequencing will become so widely affordable as to be the genome characterization strategy of choice. To buttress this prognosis, the so-called next generation sequencing platforms (i.e. the newer methods that are not based on the Sanger method of 1997, namely, Roche's 454 sequencer and Illumina's SOLEXA sequencer, but are rather based on the more cost-effective and faster pyrosequencing technology), are continually gaining acceptance and hence larger shares of the sequencing market.

A3.4 Assessing and analysing genetic diversity

There are currently many strategies for assessing genetic diversity and structure of plant populations. Many were in use at the time when the first SoW report was published and are still valuable; these include pedigree analysis and replicated field experiments (to quantify heritable variations and their components). The molecular tools used for germplasm characterization and diversity studies in 1995 included, isozyme, RFLPs, Random Amplified Polymorphic DNA (RAPD), Simple Sequence Repeat (SSR) and Amplified Fragment Length Polymorphism (AFLP) markers. With more widespread genome sequencing and generation of ESTs, SSR markers have become easier to generate and thus more widely used. Developments in high throughput marker screening systems, especially platforms that are amenable to automation and varying degrees of multiplexing, have also led to greater ease and increased efficiencies for using PCR-based markers including AFLPs. Quite importantly, the ability to discover SNPs, a marker type that is fast becoming the preferred option in high throughput systems, with ease in all parts of genomes is a direct result of significantly enhanced sequencing capacity. SSRs and the more recent SNPs are suitable for genotype fingerprinting.[20] SNPs offer the promise of higher map resolution, higher throughput, lower cost and a lower error rate compared with SSR markers.[21]

An additional feature of markers such as SNPs and SSRs is the possibility to transfer them from the genotypes in which they are identified to related materials for which sequence information is not available, without the need to re-sequence[22]. Fingerprinting individuals for SNPs dispersed throughout a genome or a particular section of interest has become a very powerful way to characterize collections such as breeding materials (including segregating populations) and genebank accessions.[23]

The utility of SNP-based genome characterization for crop improvement and genebank (*in situ* and *ex situ* materials) may be compromised in situations where sequence information is not available. In such cases, SNPs would not be an option; a high throughput microarray assay procedure, Diversity Array Technology (DArT), may be a suitable alternative. DArT technology discriminates between individuals based on polymorphisms from their simultaneous comparisons to a defined common genomic representation. It is a low-cost high-throughput system that requires minimal DNA per individual and at the same time provides comprehensive genome coverage even in organisms without any DNA sequence information.[24] Since the proof of concept with Rice in 2001, DArT has been employed for high throughput analyses in many genera including barley, *Musa* and Eucalyptus.

APPENDIX 3

For instance, DArT markers were as useful at revealing genetic relationships among 48 *Musa* accessions (derived from two wild species with different genome compositions) as other markers were, but with a lower cost and greater resolution and speed.[25]

Qualitative traits (such as many disease resistances and stress tolerances) and quantitative traits (such as indices of yield and productivity) are typically the targets for improvement in plant breeding programmes and for characterization of genebank collections. Obtaining this information for collections of individuals is laborious and costly, involving screening in the presence of pathogens and stresses in replicated field experiments with adequate sample sizes. The utility of molecular markers that could serve as proxies for this type of laborious and expensive studies is obvious.

Both natural and artificial selections are directed at genes. Though selection is a locus-specific force, it creates a pattern of variation involving few loci in specific regions of the genome. Variation in the traits governed by genes should therefore be a measure of the adaptive genetic diversity or adaptive potential of a population or breeding genepool. A majority of molecular markers only measure neutral genetic variation, i.e. variations in sections of the genome not involved with coding for genes or in the regulation of genes and hence, assumed not to be under natural selection pressure. These patterns of genetic variation are genome wide. Due to the fact that molecular methods are fast and relatively cheap, surveys of molecular marker variations are becoming widespread and attractive as means for evaluating genetic diversity across populations or genepools. There are even greater benefits when gene-based markers are used for analysis. A relevant advance in the past decade is that the relationships between adaptive genetic diversity and neutral genetic diversity are becoming much clearer.[26]

Unfortunately, many neutral molecular markers are not usually indicative of the adaptive potential of populations or accessions they are used to characterize (for example, RFLPs, RAPDs, AFLPs and SSRs).[27] In some cases, they have been used inappropriately for this purpose with the assumption that neutral markers and quantitative adaptive variation are positively correlated. There are uses of neutral molecular markers that are appropriately of value for conservation and use of genetic resources. When the patterns of genetic variation at many neutral molecular markers randomly scattered throughout a genome can be measured, they can be very useful for providing a measure of processes within ecosystems such as gene flow, genetic drift and migration or dispersal, which act on the entire genome; these are important for population biology, for monitoring progress in maintaining species in protected areas, or for testing the success of spatial connections between reserves.[28]

With the many recent, reasoned enunciations of the distinctions between types of molecular markers and the appropriateness of their respective usages for conservation and utilization of genetic resources, it is expected that any report on the deployment of molecular markers should provide a rationale for the type of marker used with respect to the objective of the work.[29] An example of investigating the utility of specific marker types for specific uses was an analysis in barley of three types of markers (EST-derived SSRs, EST-derived SNPs and AFLPs) for use in diversity analyses in breeding, natural populations and genebank materials. No one marker type was best for all studied uses.[30]

Given the ability to work with raw genomic sequence, the comprehensive pattern of DNA polymorphisms within a species can now be appreciated. *Arabidopsis thaliana* is the most thoroughly studied plant at this level since its genome was sequenced. There is an abundant natural variation for both neutral DNA markers and also for those loci that cause phenotypic changes.[31] Building on this model will be increasingly possible for crop species themselves as the genomic sequences become readily accessible. SNPs derived from ESTs were used successfully for cultivar identification in melon; this provides an example of the deployment of DNA-level polymorphism for genome characterization where few genomic tools exist other than ESTs and genetic maps based on early molecular markers.[32]

As researchers take advantage of these innovations, it needs to be emphasized that strategies adopted for estimates of genetic diversity have to be suitable to the objectives for the conservation and use of the genetic resources. To illustrate, if the aim for assaying

a number of populations of a species for diversity - as measured by a neutral molecular marker - is to accord higher priority for conservation to the most diverse populations with the assumption that this will also conserve the greatest adaptive genetic diversity, the researcher may decide that relatively few populations might be needed to capture the greatest amount of the neutral genetic diversity. A possible pitfall in this scenario is that if, for instance, the other populations were abandoned to the exclusion of the few diverse populations, significant amounts of adaptive genetic diversity, which is not distributed uniformly among all populations, would thus be lost. This would then be contrary to the originally stated objective for the assessment of genetic diversity.[33]

Molecular markers are also increasingly been used in more downstream applications. For instance, in addition to serving as tools for conserving and using genetic resources, markers have been used successfully to investigate the genetic impacts of traditional farmer practices which are often poorly documented. A case study involving yams in Benin showed that the traditional practice by farmers of selecting spontaneous wild yams from areas surrounding farms and cultivating them resulted in the creation of new varieties with new genetic combinations. These new variants arose as a direct result of sexual reproduction between wild and cultivated yams as the alleles could be traced to the progenitors. The markers used in this study were SSRs. It was therefore deduced that the mix of a cycle of sexual reproduction followed by the traditional vegetative propagation (using tubers) leads to the large-scale cultivation of the best genotypes while facilitating the introgression of potential diversity that could be useful for future adaptation.[34]

A3.5 Conservation technologies and strategies

An aspect of the use and conservation of PGRFA that has remained largely without significant advances since the first SoW Report is the orthodox seed storage conditions. Current recommendations for temperature and humidity are still the same as those developed before the first SoW report. Since then, however, the country reports that are part of this SoWPGR-2 and the crop-specific conservation strategy developed by the GCDT, call attention to the concerns for backlogs in accession testing and regeneration. For instance, it has been reported that viability testing results have indicated the need for regeneration after shorter periods of storage than were currently the norm. It is possible that, as one researcher has shown, humidity is the more critical of the two storage factors, and that seeds are exposed to higher humidity levels in the seed packaging materials than are optimal with resulting losses of viability.[35] Given the room for potential enhancements in efficiencies in seed storage, it is probably time to apply the innovative tools of biology to decipher the seemingly complex interactions in the seed storage container types, temperature and humidity regimen matrices.[36]

In the past 12 years, there have been progressively increasing reports of the assessments of the utility of molecular markers as reliable tools for managing conserved diversity in genebanks. One example of this kind of study was the use of AFLP markers to assess the extent of within-accession genetic diversity for the self-fertilizing species, lettuce, at the Centre for Genetic Resources (CGN), in the Netherlands. Two plants each with a total of 1390 accessions, (comprising six cultivar types) were screened by the array of available markers. Overall, there was a very low average probability (about one percent) that the two plants of an accession would differ. However, this probability differed among the cultivar types. The types composed of accessions that are primarily modern cultivars had probabilities of difference between the two plants of about 0.5 percent, while the two types composed of accessions which are mainly landraces had probabilities greater than one. This information would be useful in determining whether and how the observed level of diversity for each accession should be maintained in future generations of the accession.[37]

The utility of molecular markers in contributing to decisions in strategies for managing conserved diversity has also been demonstrated amply with field collections. Fingerprinting techniques have been used to determine identity and redundancy in large field collections. For example, at the ICGT in

APPENDIX 3

Trinidad and Tobago, over 2 000 crop accessions are maintained as a field collection, with each accession being represented by as many as 16 individual trees, with an overall average of six trees per accession. Multi-locus SSR fingerprinting was successfully used to resolve ambiguities arising from mislabelling of plants, a critical problem in such extensive operations.[38]

An emerging trend, for the past 12 years has been the maintenance of DNA banks of PGRFA. There have been reported cases of DNA libraries of germplasm accessions, mapping populations, breeding materials, etc. that are retrieved at will for use in subjecting the materials to molecular assays. This practice is bound to become more pervasive as the costs for molecular assays and the requisite facilities become cheaper in turn rendering this technology option more accessible to practitioners in this field. It is indicative of this trend that more formal repositories for plant DNA have been established under the auspices of botanic gardens (with examples including the RGB Kew DNA Bank or the DNA bank at the Berlin Botanic Garden and Botanical Museum) or as stand-alone entities (e.g. the Australian Plant DNA Bank and the National Institute of Agrobiological Sciences [NIAS] DNA Bank, Japan). In addition to the usual data management platforms for classical germplasm accessions, an associated bioinformatics facility is required for a DNA bank in order to accommodate the management of molecular data such as sequence and marker information for each accession. DNA banks could also serve as sources of genetic information from endangered taxa without the need for additional germplasm prospection.[39]

A3.6 Breeding methodologies

Upfront, it is worthy to emphasize that the use of genomic tools in the different facets of PGRFA management has not reduced the importance of phenotypic characterization of breeding materials, mapping populations and natural populations, or genebank accessions. On the contrary, thorough and accurate phenotyping remains as important as it has ever been and is key to the utility of molecular data as markers have value only as long as they are accurately linked to phenotypes.

Early efforts to develop large numbers of molecular markers, high-density genetic maps and appropriately structured mapping populations have now begun to enhance the efficiency of the genetic improvement of many crop species. The results from numerous mapping studies provide greatly improved estimates for the number of loci, allelic effects and gene action controlling traits of interest.[40] There have been several major advances in the incorporation of molecular techniques in crop breeding strategies since the publication of the first SoW report. These advances have led to the paradigm of molecular breeding, the collective term that encompasses marker assisted selection and recombinant DNA technologies as crop improvement strategies.

MAS

This refers to the novel crop improvement strategy of using molecular markers (genome landmarks) to aid decision making in the screening of breeding materials. This paradigm shift has been greatly facilitated by high-throughput methods for identifying and using molecular markers on a large scale, including information technology infrastructure, and by inter-disciplinary approaches that make phenotyping and trait characterizations possible across several environments. Firm verifications of the co-segregation of the trait of interest with one of the many possible types of DNA markers precede the use of the marker to select for the trait in breeding materials. MAS is becoming a valuable tool for many different crops with its utility expected to greatly increase as molecular biology assays become more cost efficient.[41] Marker development has been greatly facilitated by improvements in the genome locations of gene alleles that control traits. The advances in the construction of molecular genetic linkage maps, in building physical maps and more recently, association mapping, contribute to continually populate the repertoire of useful molecular markers for crop improvement.

Association mapping, also known as linkage disequilibrium (LD) mapping or association analysis and the most novel of the mapping methods, is a population-based survey used to link sequence polymorphisms (usually SNPs) to phenotypic variations

based on linkage disequilibrium (non-random association between alleles at linked loci) without the necessity for creating structured segregating mapping populations. By mapping nearby SNPs it is therefore possible to ascertain the genome locations of genes associated with a trait without cloning the genes. Causal SNPs identified through high density association maps are usually subsequently confirmed through functional assays. There are three main advantages of association mapping over linkage analysis: increased mapping resolution, reduced research time and greater allele number.[42]

The deployment of these strategies has been restricted primarily to crop improvement institutions which have also developed the capacity to produce sequence information for their target crops. National PGR conservation and utilization programmes are increasingly enhancing expertise and general capacity in plant biotechnology as documented in the country reports published as part of this SoWPGR-2.[43] International and other national efforts at capacity and infrastructure building have contributed to this emerging trend. However, full deployment of advanced breeding strategies, bioinformatics and genomics capabilities has not taken place in developing countries and even in many developed countries, they are only possible through collaboration with other national or international genomics projects.

The challenge within a breeding programme would be the devising of appropriate strategies for the many different scenarios that call for the integration of molecular biology techniques in PGRFA[44]. For instance while, marker-assisted backcrossing might require a few markers for genotyping hundreds of samples (backcross progenies) for a particular simply inherited trait as would the screening for introgressed elements or GMO constructs, genetic characterization or fingerprinting on the other hand would require hundreds to thousands of markers in order to be effective. In general, a genomics research service centre would be required for programmes characterized by extensive marker diversity, high throughput and large sample sizes. This requirement for high start-up investment costs probably explains the preponderance of MAS applications in large multinational breeding companies to the exclusion of the publicly funded entities.

Genetic transformation

Methods based on recombinant DNA, i.e. molecules containing DNA sequences, derived from more than one source, are used to create novel genetic variations. In crop improvement, this has involved the incorporation of exogenous DNA or RNA sequences, using either biolistics or vectors, into the genome of the recipient organism which as a result expresses novel and agronomically useful traits. The new variants are referred to as genetically modified organisms or GMOs. Transgenic crops were first grown on a commercial scale in the mid-1990s about the time of the publication of the first SoW report. Since then, the commercially grown GMOs have been four commodity crops, namely, maize, soybean, canola and cotton). By 2008, these collectively accounted for over 99.5 percent of transgenic crop production (James, 2008[45]). Interestingly also, only two transformation events, i.e. herbicide tolerance and insect resistance or their combinations were expressed in these crops. This implies therefore that more than 25 years after the first successful production of transgenic plants, the scope of the utility of genetic transformation as a routine crop improvement strategy remains limited in spite of the obvious potentials of this technology. The drawbacks include the lack of efficient genotype-independent regeneration systems for most crops and probably most limiting of all factors is the associated IPR restrictions. When GMOs have remained the exclusive preserve of private sector breeding enterprises in developed countries it has restricted (with patents) several components of the research and development efforts in the lead up to production of the transgenic crops. The interesting emerging trends - that could ultimately precipitate the review of the place of IPR protections in PGRFA - are that GMO crops are currently grown in developing countries as exemplified in the cultivation of transgenic soybean in South America and the cultivation of transgenic cotton in both India and China (James, 2008; Glover 2007,[46] 2008[47]).

As more developing countries acquire the requisite capacity for dealing with the statutory regulations governing the cultivation of GMOs, especially in line with biosafety regulations as enunciated in the

APPENDIX 3

Cartagena Protocol on Biosafety, concerted efforts will need to be targeted at building capacity to navigate the IPR restrictions that effectively impeded the exploration of the full potentials of transgenesis in PGRFA. Moving forward, it is surmised that research efforts on the other hand will target the refining of plant regeneration systems and, quite importantly, expanding the scope of agronomic traits that can be improved using genetic transformation. So far, the stacking of several transformation events and getting them to express phenotypes in one recipient organism has remained impractical. Removing the technological barriers will be key to taking advantage of genetic transformation to address polygenic traits, especially those related to climate change and variations such as drought and salinity. Removing this bottleneck will also be important for gene pyramiding.

A3.7 Bioinformatics

One consequence of the relative ease for generating molecular genetic data has been the need for ever increasing capacity for electronic data storage, analysis and retrieval systems. Currently, the data storage requirements are estimated in petabytes, about three orders of magnitude greater than what was commonly in use in 1995. A trend in cost reductions for bioinformatics facilities is that costly mainframe computer installations for bioinformatics work have been mainly replaced at genomics centres by computer server farms comprised of off-the-shelf, ordinary PCs or servers harnessed together to provide equal or greater computing capacity at lower cost and with built-in Central Processing Unit (CPU) redundancy. These units are conditioned to ensure greater reliability even with individual unit failures. Access to such storage and analysis systems is increasingly made available by the incorporation of internet servers within the system.

It is the combination of creative software engineering, open-source operating systems and database software, the advent of the ubiquitous access to, and use of, the Internet and both private and public investment that have led to the availability of reliable tools to manage genomics laboratories and hence the capacity to store, analyse, distribute and interpret the massive datasets generated from sequencing projects and molecular biology based activities.

New algorithms and statistics are continually necessary to study relationships among data sets. Maps are the most common formats for presenting genetic information and developing software for producing and displaying maps has remained one of the most active fields of research and development in molecular biology. Advances in bioinformatics will continue to be necessary to facilitate the analysis of genomic data and the integration of genomics information with data from the related fields of transcriptomics, proteomics, metabolomics and phenomics.

Collaborative genome projects have led to the creation of databases that store data centrally but are accessible globally. Integral to such efforts are collections of genomic resources whose inventories and access are components of the genome database. Funding for such projects has been largely within the public sector (nationally and internationally).

A3.8 Policy, organizational and legal considerations

The major international instrument impacting PGR conservation and use since 1995 was the ITPGRFA which was adopted in 2001 and came into force in 2004.[48] This agreement, aimed at improving upon the Convention on Biological Diversity, obligates parties to the Treaty to develop legislation and regulations to fulfill its mandates to facilitate conservation, exchange and use of the genetic resources covered by the ITPGRFA. The development of a specialized funding mechanism for the ITPGRFA subsequently took place and the GCDT was created in 2004. Currently, the GCDT is raising an endowment and additional funding for upgrading national germplasm collection facilities, building capacity and strengthening information systems. Special focus has been on the collaborative development of regional and global crop conservation strategies.[49] A major development in the exchange of PGRFA since the first SoW report has been the SMTA that provides the Contracting Parties

with a multilateral system for executing the exchange of crop germplasm.

National and international research funding bodies, in recognizing the need for collaboration for successful genomics projects have tailored some of their funding programmes to specifically underwrite collaborative efforts. The results have been public investments in sequencing centres, databases of genomic data, tools for analyses and public access, typically via the internet. The ability to sustain or increase such investments will depend on the status of the global and national economies. While there was a fall in world gross product in 2009, the first since World War II, prospects seem to be improving for a recovery in 2010.[50]

The technical advancements in DNA fingerprinting may have relevance for intellectual property protection to the extent that it is possible to unambiguously identify cultivars. SNP fingerprinting will be precise and applicable in a high-throughput process; however, widespread application is still limited to crops with SNP databases. More widely used to date are fingerprinting platforms based on SSR markers or even AFLP and RAPD markers.[51]

Concerns for the protection of inventors' IPR in endeavors relating to PGRFA were initially restricted to the safeguarding of PBR. At the national levels, this safeguard was provided through different forms of legislations that vested IPR over new crop varieties with the developer, namely, the plant breeder. Efforts at harmonizing these national laws resulted in the 1961 International Convention and Union for the Protection of New Varieties of Plants (UPOV) and its revised Acts of 1972, 1978 and 1991. This was followed by the WTO Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) which was signed in 1994. TRIPS had specific provisions for IPR protection relating to innovations in agricultural produce (crops and animals). Efforts at engendering IPR at both the national and international levels always had the stated aim of facilitating access to inventions in a fair and equitable manner. It is axiomatic that the net results of such well intentioned interventions have been further restrictions to access.

Inventions in biotechnology, including those relating to PGRFA, have spawned such an unprecedented rash of patents as to amount to a virtual gridlock in efforts to access biotechnological innovations. Since the first SoW report, the profile of biotechnology in food and agriculture has continued to rise especially with the near ubiquity of GMO crops either in commercial production or in trial stages in many parts of the world. The patent protections for the crops and even the materials used in developing them, such as the sequences of the gene constructs, have been notoriously restrictive. For instance, it is such IPR issues that have impeded the widespread use of the genetically engineered high beta-carotene rice, the golden rice, as public good. Considering the moral imperatives of safeguarding food security, it is surprising that a lot more efforts have not been invested in breaking these gridlocks.

The options for accessing proprietary biotechnologies by national research organizations are severely limited as the costs are usually prohibitive. The alternatives, normally requiring accessing the technologies without permission, would involve the exploitation of loopholes in patent and protected variety jurisdictions. International public research entities, notably the centers of the Consultative Group for International Agricultural Research, have also been successful with negotiating royalty-free access. A pioneering regional effort, the African Agricultural Technology Foundation, has also been able to broker access to IPR protected biotechnologies that impact on the ability of national programmes to harness the full potentials of their PGRFA. In general, the current efforts at accessing such technologies under IPR regimens have been piecemeal, expensive and clearly call for concerted international collaborations. The starting point will be education and capacity building in order to deal with the very complex issues involved.

A3.9 Future perspectives

The future will present multiple challenges to crop performance that can be met effectively by a combination of the development of resilient and hardy crop varieties (modifying crop genomes through plant breeding, preferably facilitated by efficient molecular approaches) and the introduction of suites

APPENDIX 3

of mitigating factors into agronomic management practices. In order to increase the reliability of predictions of crop performance based on molecular genetic information, new tools that enhance the ability for more precise linkage of molecular profiles (genotypes) to performance (phenotypes) will have to be readily accessible to the researcher.

Gaps in knowledge abound that must also be addressed. For instance, the subtleties of phenotypic plasticity in the face of a changing environment and the layers of genetic redundancy that characterize biological systems remain largely uncharted. The concerted application of the myriad tools and procedures that are both available now and are under development holds great promise for deciphering these processes and thereby enhance the ability to more efficiently manage PGRFA in the face of the daunting challenges of an increasingly variable climate, increasing world populations and competing demands for diverting foodstuffs to non-traditional uses in fuel, animal feeds and fibre industries.

The cumulative progress achieved to date in genomics and its ancillary scientific and technological endeavors has only provided the very beginning of understanding for the way in which a genotype confers a particular set of attributes to a living organism. Today, it is possible to dissect a complex phenotype and to determine where individual genes or, more correctly, QTL are physically located along the chromosomes. Information about DNA markers linked to QTL represents a powerful diagnostic tool that enables a breeder to select for specific introgressions of interest. As more genes of interest are cloned, identified or mapped and their contributions to complex biological systems are better understood, there will be many opportunities for creative "synthesis" of new varieties. It is possible that some of the opportunities will involve genetic engineering approaches, where new information about genes, gene regulation and plant responses to the environment may be used in innovative ways to fine-tune existing plant varieties so that they utilize resources more efficiently, provide greater nutritional value, or simply taste better.

A continuing need will be that of extending molecular crop improvement strategies and capacities to under-studied and under-funded crops (the so called orphan crops) but which ironically remain the bulwarks for the food security of a huge percentage of humankind. Achieving a widespread and routine application of novel biotechnologies to orphan crops, with the attendant potential for extensive positive impacts on human welfare, therefore represents an irresistible opportunity not only to those dedicated to public goods but to humanity in general. The current unacceptably high level of food insecurity need not remain so, nor get worse; the judicious management of PGRFA – while taking advantage of the novel tools and advancements - holds the key to reversing the trend.

The immediate steps will involve the investment of resources in empirical studies with the aim of attaining an understanding of the biological processes that underpin the phenotypes of the crops themselves.[52] To date, the species sequenced or for which sequencing is underway, represent only about 13 plant families. There is a compelling need to make inroads into the balance of over 600 plant families for which genome sequences have not commenced as the benefits of whole genome sequence data are proving incalculable. Precisely, many orphan crop species and others need to become candidates for sequencing.

None of these advances in technological innovations lessens the need for collections of PGR. In fact, to make the best use of the new tools, new strategies may be necessary to capture even greater genetic diversity and for maintaining that diversity during conservation and regeneration of samples. Genebanks remain vital and are in need of increased support.[53]

Also, parallel progress in genome analysis of plant pests and pathogens should lead to greater insights into mechanisms of disease and pest resistance. Global climate change and variations will present some predictable challenges to agricultural production systems (for example, higher temperatures, drought, flood, stronger winds and increased and new pests and pathogens). To address those challenges, research should make full use of available molecular tools and strategies not only to improve productivity but also to reduce impact on the environment, increase carbon sequestration and substitute for fossil fuels.[54]

References

1 **Metzker, M.L.** 2010. Sequencing technologies—the next generation. *Nature Reviews Genetics* 11:31-46. While this survey has a focus on human genomics, the conclusions about sequencing capabilities and capacities are relevant to plant genomics.

2 **Delseny, M.** 2004. Re-evaluating the relevance of ancestral shared synteny as a tool for crop improvement. *Current Opinions in Plant Biology* 7:126-131.

3 The characterization of progress in genomic technology in this paragraph as a series of waves derives from this review: **Borevitz, J.O. & Ecker, J.R.** 2004. Plant genomics: The third wave. *Annu. Rev. Genom. Hum. Genet.* 5:443-447. While this survey of what has been and what will be possible for plant genomics is based on progress with *Arabidopsis thaliana*, there is much of relevance here to plant genomics in general.

4 **The Arabidopsis Genome Initiative.** 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature,* 408:796-815.

5 **Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L. & Yang, H.** 2002. A draft sequence of the rice genome (*Oryza sativa* ssp. *indica*). *Science,* 296:79-92.

6 **Goff, S.A., Ricke D., Lan, T. H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, ., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R. M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A. & Briggs, S.** 2002. A draft sequence of the rice genome (*Oryza sativa* ssp. *japonica*). *Science,* 296:92-100.

7 **Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D.L., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S.,Déjardin, A., dePamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjärvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leplé, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouzé, P., Ryaboy, D., Schmutz, J.,**

APPENDIX 3

Schrader, J., Segerman, B., Shin, A., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y. & Rokhsar, D. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science,* 313:1596-1604.

8 http://medicago.org/genome/

9 See: http://www.phytozome.net/sorghum; http://www.phytozome.net/grape.php; and http://www.phytozome.net/papaya.php

10 http://www.phytozome.net/soybean.php

11 http://genome.jgi-psf.org/Araly1/Araly1.info.html

12 http://brachypodium.pw.usda.gov/

13 http://maizesequence.org/index.html

14 Good entry points for access to sequence databases and genome browsers for plants are PlantGDB at http://www.plantgdb.org/ and Phytozome at http://www.phytozome.net/.

15 The listed taxa come from the NCBI Entrez Genome Project site at http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi?taxgroup=11:|12:Land%20Plants&p3=12:Land%20Plants.

16 http://www.ncbi.nlm.nih.gov/nucgss

17 http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html

18 **Strausberg, R.L., Levy, S. & Rogers, Y.-H.** 2008. Emerging DNA sequencing technologies for human genomic medicine. *Drug Discovery Today* 13:569-577. Although presented in the context of human genomics, the three major sequencing technologies described are in use in crop plant research today and the forecast of emerging ones is equally relevant.

19 **Metzker, M.L.** 2010. Sequencing technologies—The next generation. *Nature Reviews Genetics* 11:31-46. A more recent review of the same three technologies along with details of a new platform expected in 2010.

20 **Angaji, S.A.** 2009. Single nucleotide polymorphism genotyping and its application on mapping and marker-assisted plant breeding. *African Journal of Biotechnology,* 8:908-914.

21 **Jones, E., Chu, W.-C., Ayele, M., Ho, J., Bruggeman, E., Yourstone, K., Rafalski, A., Smith, O.S., McMullen, M.D., Bezawada, C., Warren, J., Babayev, J., Basu, S. & Smith, S.** 2009. Development of single nucleotide polymorphism (SNP) markers for use in commercial maize (*Zea mays* L.) germplasm. *Molecular Breeding,* 24:165-176.

22 **Vezzulli, S., Micheletti, D., Riaz, S., Pindo, M., Viola, R., This, P., Walker, M.A., Troggio, M. & Velasco, R.** 2008. An SNP transferability survey within the genus *Vitis*. *BMC Plant Biology* 8:128-137. Genomic information from one *V. vinifera* cultivar for which sequencing information was available was leveraged to inform other closely related cultivars and wild forms in that species without the need for resequencing. However, utility was limited for other species of *Vitis*.

23 **Spooner, D., van Treuren, R. & de Vicente, M.C.** 2005. Molecular markers for genebank management. IPGRI Technical Bulletin No. 10. International Plant Genetic Resources Institute [now Bioversity International, Inc.]. Rome, Italy.

24 **Jaccoud, D., Peng, K., Feinstein, D. & Kilian, A.** 2001. Diversity arrays: A solid state technology for sequence information independent genotyping. *Nucleic Acids Research* 29:e25-e31. Describes the technique with a case study of its use with rice.

25 **Risterucci, A.-M., Hippolyte, I., Perrier, X., Xia, L., Caig, V., Evers, M., Huttner, E., Kilian, A. & Glaszmann, J.C.** 2009. Development and assessment

of Diversity Arrays Technology for high-throughput DNA analyses in *Musa*. *Theor. and Applied Genet.,* 119:1093-1103.

[26] **González-Martínez, S.C., Krutovsky, K.V. & Neale, D.B.** 2006. Forest tree population genomics and adaptive evolution. *New Phytologist* 170:227-238. Provides a review of differences among marker types.

[27] **FAO.** 2001. Forest genomics for conserving adaptive genetic diversity. Paper prepared by K. Krutovskii and D.B. Neale. Forest Genetic Resources Working Papers, Working Paper FGR/3 (July 2001). Forest Resources Development Service, Forest Resources Division. FAO, Rome (*unpublished*).

[28] **Holderegger, R., Kamm, U. & Gugerli, F.** 2006. Adaptive versus neutral genetic diversity: Implications for landscape genetics. *Landscape Ecology* 21:797-807.

[29] For example, a thorough discussion of several types of markers and many different uses of them is provided by **De Vincente, M.C., Guzman, F.A., Engels, J.M.M. & Rao, V.R.** 2006. Genetic characterization and its use in decision-making for the conservation of crop germplasm. p. 129-138 *in* J. Ruane and A. Sonnino (eds.) *The role of biotechnology in exploring and protecting agricultural genetic resources*. UN Food and Agriculture Organization. Rome, Italy.

[30] **Varshney, R.K., Chabane, K., Hendre, P.S., Aggarwal, R.K. & Graner, A.** 2007. Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. *Plant Science,* 173:638-649.

[31] Op cit. Endnote 4.

[32] **Deleu, W., Esteras, C., Roig, C., González-To, M., Fernández-Silva, I., Gonzalez-Ibeas, D., Blanca, J., Aranda, M.A., Arús, P., Nuez, F., Monforte, A.J., Picó, M.B. & Garcia-Mas, J.** 2009. A set of EST-SNPs for map saturation and cultivar identification in melon. *BMC Plant Biology,* 9:90-98.

[33] **Bonin, A., Nicole, F., Pompanon, F., Miaud, C. & Taberlet, P.** 2007. Population adaptive index: A new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology* 21:697-708. Combines an analysis of the differences among neutral and adaptive diversity with a presentation of a 'population adaptive index' proposed as a way to allow use of many molecular markers distributed throughout the genome (a measure only possible because of advances in biotechnology) that will allow pinpointing localized variations in the pattern of diversity thus detecting loci supposedly under natural selection and thus of adaptive significance.

[34] **Scarelli, N., Tostain, S., Vigouroux, Y., Agbangla, C., Daïnou, O. & Pham, J.-L.** 2006. Farmers' use of wild relative and sexual reproduction in a vegetatively propagated crop. The case of yam in Benin. *Molecular Ecology,* 15:2421-2431.

[35] **Gómez-Campo, C.** 2006. Erosion of genetic resources within seed genebanks: The role of seed containers. *Seed Science Research,* 16:291-294.

[36] **Pérez-García, F., González-Benito, M.E. & Gómez-Campo, C.** 2007. High viability recorded in ultra-dry seeds of *Brassicaceae* after almost 40 years of storage. *Seed Science and Technology* 35:143-153. This paper presents data on the impact of humidity and quality of storage materials on seed longevity.

[37] **Jansen, J., Verbakel, H., Peleman, J. & Van Hintum, T.J.L.** 2006. A note on the measurement of genetic diversity within genebank accessions of lettuce (*Lactuca sativa* L.) using AFLP markers. *Theor. and Applied Genet.,* 112:554-561.

[38] **Motilal, L.A., Zhang, D., Umaharan, P., Mischke, S., Boccara, M. & Pinney, S.** 2009. Increasing accuracy and throughput in large-scale microsatellite fingerprinting of cacao field germplasm collections. *Tropical Plant Biology,* 2:23-37.

[39] **Rice, N., Cordeiro, G., Shepherd, M., Bundock, P., Bradbury, L., Pacey-Miller, T., Furtado, A. &**

APPENDIX 3

Henry, R. 2006. DNA banks and their role in facilitating the application of genomics to plant germplasm. *Plant Genetic Resources* 4:64-70. Australian Plant DNA Bank: http://www.dnabank.com.au/; NIAS DNA Bank: http://www.dna.affrc.go.jp/; RBG Kew DNA Bank: http://data.kew.org/dnabank/homepage.html; The DNA Bank in Berlin-Dahlem, at the Botanic Garden and Botanical Museum (BGBM): http://www.bgbm.org/bgbm/research/dna/.

40 **Moose, S.P. & Mumm, R.H.** 2008. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology,* 147:969-977.

41 **Guimarães, E.P., Ruane, J., Scherf, B.D., Sonnino, A. & Dargie, J.D.** (eds.) 2007. *Marker-assisted selection: Current status and future perspectives in crops, livestock, forestry and fish*. UN Food and Agriculture Organization. Rome, Italy.

42 **Zhu, C., Gore, M., Buckler, E.S. & Yu, J.,** 2008. Status and prospects of association mapping in plants. *The Plant Genome,* 1:5-20.

43 For example, according to Country reports, molecular markers are in use for crop improvement in Argentina, Azerbaijan, Brazil, China, Croatia, Czech Republic, Egypt and Indonesia.

44 **Bagge, M. & Lübberstedt, T.** 2008. Functional markers in wheat: Technical and economic aspects. *Molecular Breeding,* 22:319-328.

45 **James, C.** 2008. *Global status of commercialized biotech/GM crops: 2008*. ISAAA Brief No 39. Available online: www.isaaa.org/resources/publications/briefs/39/default.html

46 **Glover, D.** 2007. *Monsanto and smallholder farmers: A case-study on corporate accountability.* IDS Working Paper 277. University of Sussex, UK, Institute of Development Studies.

47 **Glover, D.** 2008. *Made by Monsanto: The corporate shaping of GM crops as a technology for the poor*. STEPS Working Paper 11. Brighton: STEPS Centre. Available online: www.steps-centre.org/PDFs/GM Crops web final_small.pdf.

48 See Chapter 7

49 See Chapter 6 and Appendix 4.

50 **United Nations.** 2010. *World economic situation and prospects 2010*. Department of Economic and Social Affairs, United Nations. New York NY USA.

51 **Romero, G., Adeva, C. & Battad II, Z.** 2009. Genetic fingerprinting: Advancing the frontiers of crop biology research. *Philippine Science Letters* 2:8-13. This review summarizes the status of deploying fingerprinting with different markers, with examples from crops and situations in the Philippines.

52 **Nelson, R.J., Naylor, R.L. & Jahn, M.M.** 2004. The role of genomics research in improvement of "orphan" crops. *Crop Science,* 44:1901-1904.

53 See Chapters 3 & 4. For an outspoken advocacy of wider collecting and conservation strategies, see **Walck, J. & Dixon, K.** 2009. Time to future-proof plants in storage. *Nature,* 462:721.

54 The Brazil Country report, Chapter 9, offers a very effective discussion of these issues and a rationale for the contribution of genetic resources to sustainable development and food security.