ISSN 1479-2621

Plant Genetic Resources 2(1); 3–21
DOI: 10.1079/PGR200430

Review article

Plant genetic resources in crop improvement[1]

B. I. G. Haussmann[1]*, H. K. Parzies[1], T. Presterl[1], Z. Sušić[1] and T. Miedaner[2]

[1] *Institute of Plant Breeding, Seed Science, and Population Genetics and* [2] *State Plant Breeding Institute, University of Hohenheim, 70593 Stuttgart, Germany*

Received 29 October 2003; Accepted 10 February 2004

Abstract

The use of plant genetic resources (PGR) in crop improvement, followed by adoption, cultivation and consumption or marketing of the improved cultivars by farmers, is one of the most sustainable methods to conserve valuable genetic resources for the future, and simultaneously to increase agricultural production and food security. The objective of this review is to summarize issues related to the use of PGR in crop improvement. Specific topics are: definition of genetic resources for crop improvement; information sources on the internet; documentation and evaluation of PGR; access to PGR, equitable sharing of profits, and material transfer agreements; impediments to the use of PGR in crop improvement; classical methods of using PGR in crop improvement (introgression, incorporation, prebreeding and wide crosses); use of landraces in breeding for specific adaptation to stress environments; utility of molecular markers and genomic research for using PGR in crop improvement (diversity assessment, mapping of quantitative trait loci (QTL) and marker-assisted selection (MAS), advanced backcross QTL analysis and introgression libraries, association studies and direct allele selection); and gene transfer. Practical examples or experimental results are given for most aspects.

Keywords: incorporation; introgression; molecular markers; prebreeding

Introduction

Agriculture today is characterized by a sharp reduction in the diversity of cultivated plants. Out of an estimated total of 30,000 edible plant species, only 30 'feed the world', with the three major crops being maize (*Zea mays*), wheat (*Triticum aestivum*) and rice (*Oryza sativa*) (FAO, 1996a). In addition to the interspecific reduction of crop diversity in agriculture, plant breeding contributes to diminution of the intraspecific diversity, through development of adapted breeding populations, selection of the 'best' genotypes, development of genetically homogeneous cultivars and promotion of few widely adapted varieties. The lack of inter- and intraspecific genetic variability among cultivated crops can lead to:

- epidemics of pests and diseases (genetic vulnerability); examples are the *Phytophthora infestans* infestation of potato (*Solanum tuberosum*) in Western Europe in 1845/1846, the *Bipolaris maydis* disaster in

* Corresponding author. E-mail: haussb@uni-hohenheim.de

[1] Dedicated to Dr Hartwig H. Geiger, Professor of Population Genetics at University of Hohenheim, Stuttgart, Germany, on the occasion of his 65th birthday.

T-cytoplasm maize in the USA in 1970 (Campbell and Madden, 1990) and the *Fusarium graminearum* epidemic in wheat and barley in the western USA (1994–1996) (FAO, 1996a);
- lack of adaptation to increasing abiotic stresses like drought or high ozone concentrations;
- lack of genetic variation for specific quality traits, e.g. starch quality in maize (Whitt *et al.*, 2002), fatty acid composition or male sterility in oilseed rape (*Brassica napus*; Hu *et al.*, 2002).

Reaching performance plateaux may be another risk. However, long-term selection experiments demonstrate that for traits such as oil or protein content in maize, simple ear-to-row selection has been effective over 90 generations without reaching a plateau (Dudley and Lambert, 1992). No new genetic variation was ever added to the initial breeding population.

As outlined by the Food and Agricultural Organization of the United Nations (FAO) in the Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources, a more efficient use of plant genetic diversity is a prerequisite to meeting the challenges of development, food security and poverty alleviation (FAO, 1996b). Concrete aims are:

- to develop cultivars that are specifically adapted to marginal or stress environments;
- to assure sustainable production in high-yielding environments through better input–output relations, i.e. through reduced application of agro-chemicals and increased nutrient and water efficiency;
- to open production alternatives for farmers through development of industrial or pharmaceutical crops.

To achieve these aims, extensive *ex situ* and *in situ* conservation of PGR must be assured. Evaluation of conserved accessions and their use by plant breeders or farmers needs to be supported and facilitated. The aim should not only be to exploit intraspecific variation within a crop but also to increase interspecific diversity in agriculture through genetic improvement and promotion of less popular, neglected or underutilized crop species (e.g. Padulosi *et al.*, 2002). Many underutilized species are particularly useful in marginal lands where they have been selected to tolerate stress conditions and contribute to sustainable production. These genetic resources need to be evaluated for their outcrossing rates, yield potential, response to inputs, agronomic value and the amount of genetic variation for specific traits, to allow more efficient genetic improvement and promotion.

The genetic improvement of PGR for specific traits, followed by successful cultivation and marketing or consumption of the improved materials, is probably one of the most sustainable ways to 'conserve' valuable genetic resources for the future.

Definition of genetic resources for crop improvement

Genetic resources can be defined as all materials that are available for improvement of a cultivated plant species (Becker, 1993). In classical plant breeding, genetic resources may also be considered as those materials that, without selection for adaptation to the target environment, do not have any immediate use for the breeders (Hallauer and Miranda, 1981). According to the extended gene pool concept, genetic resources may be divided into primary gene pool, secondary gene pool, tertiary gene pool and isolated genes (Harlan and de Wet, 1971; Becker, 1993). The primary gene pool consists of the crop species itself and other species that can be easily crossed with it. The secondary gene pool is composed of related species that are more difficult to cross with the target crop, i.e. where crossing is less successful (low percentage of viable kernels) and where crossing progenies are partially sterile. The tertiary gene pool consists of species which can only be used by employing special techniques like embryo rescue or protoplast fusion. The fourth class of genetic resources, isolated genes, may derive from related or unrelated plant species, from animals or microorganisms. The importance of the different classes of genetic resources for crop improvement depends on the target crop species. In maize, for example, genetic variation in the primary gene pool is so large that the secondary or tertiary gene pools are rarely used. In rape seed, on the other hand, genetic variation in the primary gene pool is small and breeders have to transfer important traits from *Brassica* species of the secondary and tertiary gene pool into the cultivated species (e.g. Hu *et al.*, 2002).

Facts and information sources

World-wide, 1308 genebanks are registered in the WIEWS (World Information and Early Warning System on PGR) database (http://apps3.fao.org/wiews/) and conserve a total of 6.1 million accessions, including major crops, minor or neglected crop species, as well as trees and wild plants. Of the 30 main crops, more than 3.6 million accessions are conserved *ex situ* (FAO, 1996a). Detailed information on *ex situ* conservation of PGR is available in Hawkes *et al.* (2000). Little information exists about documentation and availability of materials that are maintained *in situ*, though concepts for a sustainable *in situ*

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

conservation have been developed (Maxted *et al.*, 1997, 2002). Some of the most important organizations or networks dealing with PGR are listed below. Considerable information and additional links can be found on the individual homepages.

- The World Information and Early Warning System (WIEWS) on Plant Genetic Resources for Food and Agriculture (PGRFA) was established by FAO as a dynamic, world-wide mechanism to promote information exchange among member countries, by gathering and disseminating information on PGRFA, and as an instrument for the periodic assessment of the state of the world's PGRFA (http://apps3.fao.org/wiews/).
- The CGIAR (Consultative Group of International Agricultural Research) System-wide Information Network for Genetic Resources (SINGER) links the genetic resources information systems of the individual CGIAR Centres around the world, allowing them to be accessed and searched collectively. SINGER contains key data of more than half a million individual accessions of crop, forage and agroforestry genetic resources held in the Centre genebanks (http://www.singer.cgiar.org/).
- The International Plant Genetic Resources Institute (IPGRI) is the world's largest non-profit agricultural research and training organization devoted solely to the study and promotion of agricultural biodiversity (http://www.ipgri.cgiar.org/).
- The National Plant Germplasm System (NPGS) of the United States Department of Agriculture is a co-operative effort by public and private organizations. With its Germplasm Resources Information Network (GRIN), it seeks to facilitate the acquisition, preservation, evaluation, documentation and distribution among scientists of germplasm (http://www.ars-grin.gov/npgs/).
- The European Cooperative Program for Crop Genetic Resources Networks (ECP/GR) is a collaborative programme seeking to ensure long-term conservation and increased utilization of PGR in Europe (http://www.ecpgr.cgiar.org/).
- The European Internet Search Catalogue (EURISCO, http://eurisco.ecpgr.org/) displays passport data on *ex situ* collections maintained in Europe.
- The Mansfeld database supplies information from *Mansfeld's World Manual of Agricultural and Horticultural Crops* for an internet-searchable database (http://mansfeld.ipk-gatersleben.de/Mansfeld/).
- The European Crop Wild Relative Diversity Assessment and Conservation Forum (PGR Forum) provides a European forum for the assessment of taxonomic and genetic diversity of European crop wild relatives and the development of appropriate conservation methodologies (http://www.pgrforum.org/).
- The Information System on Genetic Resources (GENRES) provides links to internet resources related to all kind of genetic resources (plants, animals, microorganisms) within Germany (GENRES-Deutschland) or world-wide (GENRES-International) (http://www.genres.de/).

Documentation and evaluation of PGR

Genebank accessions are described by passport and characterization data, and to a variable extent also by evaluation data. Passport data include serial number, taxonomic name, collection site, date of collection and donor institute. Additional notes can refer to seed viability, number and mode of regeneration or reproduction, and information about the distribution of the sample. Germplasm passport information exchange is facilitated by the internationally standardized list of multi-crop passport descriptors (MCPD), which have been developed jointly by IPGRI and FAO (http://www.ipgri.cgiar.org/publications/pdf/124.pdf). Characterization data comprise scores for simple morphological traits like plant height, maturity date and thousand seed weight to descriptor states reflecting specific alleles for known genetic systems. Evaluation data refer to agronomic traits like grain yield, grain quality, lodging, and resistance to important pests and diseases as far as evaluated. Evaluation is a continuous process. Different people or institutions can be involved, such as genebanks, breeders, pathologists or physiologists searching for specific traits. Ideally, all data sets accompanying an accession are stored in a central database and are made available to the public; one example is the information system for evaluation data of PGR in Germany (EVA, http://www.genres.de/eva/en/index.htm). Little is known about the extent of evaluation in individual genebanks; according to FAO (1996a), the percentage of evaluated accessions ranges from 5 to 100%.

The use of PGR in crop improvement could be facilitated by systematic evaluation and documentation of the acquired data. Of particular importance are:

- information on valuable traits, e.g. resistances and specific quality traits;
- reliable information on genotype × environment interactions and specific adaptation;
- information on general and specific combining abilities and affiliation to heterotic pools (if hybrid breeding is relevant);
- on-farm evaluation to gain information about farmer's perception;
- user-friendly information and documentation systems (standard formats);

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

- creation and evaluation of core collections; and
- international co-operation.

At the outset, core collections were defined as a limited set of accessions representing, with a minimum of repetition, the genetic diversity of a crop species and its wild relatives (Frankel, 1984). In the context of an individual genebank, a core collection consists of a limited number of the accessions of an existing collection. These are chosen to represent the genetic spectrum in the whole collection and should include as much of its genetic diversity as possible (Brown, 1995). Core collections can render the evaluation process more efficient because repetition of similar entries is avoided. For more information on the establishment of core collections, the reader is referred to Hodgkin *et al.* (1995) and van Hintum *et al.* (2000).

Networking is an important component of international co-operation during the evaluation of PGR. Examples of such networking are the International Network for the Genetic Evaluation of Rice (INGER, http://www.irri.org/GRC/inger/main.htm); the Latin American Maize project (LAMP) that provides information necessary to select efficiently germplasm accessions for the Genetic Enhancement of Maize (GEM) project in the USA (http://www.public.iastate.edu/~usda-gem/); and the West Asia and North Africa Plant Genetic Resources Network (WANANET, http://www.ipgri.cgiar.org/).

Access to PGR, equitable sharing of profits and benefits, and material transfer agreements

The Convention on Biological Diversity, signed by 187 states, aims at the conservation and sustainable use of biological diversity, and an equitable sharing of profits and benefits generated by the use of genetic resources (http://www.biodiv.org/). Conservation of PGR is assured by the genebanks, hopefully soon with financial support from the Global Conservation Trust (http://www.startwithaseed.org/pages/trust.htm). Access to PGR from genebanks, public institutes and breeding companies is increasingly facilitated through international co-operation, networking during the evaluation process and publicly available information on the internet. A central question remains of how to guarantee equitable sharing of profits. Keywords in this context are 'breeder's rights', 'intellectual property rights (IPR)', 'biopiracy' and 'farmer's rights'.

Breeder's rights have been defined by UPOV (International Union for the Protection of New Varieties of Plants, http://www.upov.int/index.html). According to Article 14:

> the following acts in respect of the propagating material of the protected variety shall require the authorization of the breeder: production or reproduction (multiplication), conditioning for the purpose of propagation, offering for sale, selling or other marketing, exporting, importing, stocking for any of the purposes mentioned above. The breeder may make his authorization subject to conditions and limitations.

In addition, individual countries may impose further limits on the use of new cultivars. For example, it is not allowed in the USA to use public varieties for crossing and development of new cultivars while this action is legal in Germany and other European countries. The resulting 'essentially derived varieties' were recently investigated (Heckenberger *et al.*, 2002, 2003; Troyer and Rocheford, 2002).

Intellectual property rights have been set up in the TRIPS (Trade-Related Aspects of Intellectual Property Rights) agreement of the World Trade Organization (WTO, http://www.wto.org/english/tratop_e/trips_e/t_agm0_e.htm). Patents or monopoly rights on improved seed or genes can limit access, distribution and further use of specific PGR. This act has been contested and is considered as biopiracy when it results in the prohibition of local communities and indigenous peoples using their own seed (e.g. Friends of the Earth International, http://www.foei.org/media/2002/0419.html; Locke, 2001). The question is, who is the real owner of PGR which have been developed over thousands of years? Without doubt, local farmers, often in developing countries, have been involved in selection and *in situ* maintenance of important PGR. Therefore, farmer's rights were defined by FAO as 'rights arising from the past, present and future contribution of farmers in conserving, improving and making available PGR, particularly those in the centers of origin/diversity' (FAO, 1989). The purpose of these rights is stated to be 'ensuring full benefits to farmers and supporting the continuation of their contributions' (FAO, 1989). According to the Convention on Biological Diversity, individual countries should have sovereign rights over their own biological resources. Farmer's rights are therefore the responsibility of the individual governments. A number of non-governmental organizations try to assure realization of concrete farmer's rights. Examples are: Genetic Resources Action International (GRAIN, http://www.grain.org/about/index.cfm); South East Asia Regional Institute for Community Education (SEARICE, http://www.searice.org.ph/); the Rural Advancement Foundation International in the USA (RAFI-USA, http://www.rafiusa.org/); and the Field Alliance in Asia (http://www.thefieldalliance.org/Issues/Issues_overview.htm).

For a more detailed discussion of the issues involving breeder's rights, IPR and farmer's rights, the reader is referred to Evenson (1999), ten Kate and Laird (2000) and Swaminathan (2002). Specialized lawyers will be needed to come up with fair solutions for all parties.

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

Material transfer agreements (MTAs) define the specific situation for individual PGR. For example, MTAs from the CGIAR centres (http://www.cgiar.org/pdf/mtaeng.pdf) protect the PGR against IPR and assure continuous and free availability. However, the equal share of benefits is inadequately defined. MTAs from other institutions can refer to restricted plant materials and in this case, the user has to agree to use the material for research only; not to distribute or commercialize the plant material or materials derived from IPR-protected plant materials, and to take all reasonable precautions to prevent unauthorized propagation of any of this material or derived plant materials for the duration of the IPR (USDA, http://www.ba.ars.usda.gov/techt/mta.htm).

Impediments to the use of PGR in crop improvement

Lack of environmental adaptation of the PGR is one major reason for the limited use of genetic resources in classical plant breeding. Other reasons impeding the use of PGR in crop improvement are: huge performance difference between PGR and actual breeding materials for complex inherited traits; lack of inbreeding tolerance and unknown affiliation to heterotic pools (in the case of hybrid breeding); and genetic problems like pleiotropy, linkage between desired and undesired alleles of the PGR; and epistasis or co-adaptation of genes within both breeding population and PGR. Pleiotropy is the situation where one gene locus affects several traits. A specific allele from the PGR may be favourable for one trait, but negative for the expression of another trait that is directly or indirectly under control of the same gene locus. Strong linkage between desired and undesired alleles of the PGR makes it difficult to develop overall superior materials. With conventional backcrossing, the linkage drag (i.e. the 'baggage' which comes with the actual gene of interest) is reduced only very slowly; about 53 cM remain around the target gene in the third backcrossing generation (BC_3), and in BC_{10} the average linkage drag is still about 20 cM (assuming target locus in the centre of a 100 cM chromosome; Stam and Zeven, 1981; Welz and Geiger, 2000). If this linkage drag contains undesirable alleles from the PGR, the performance of the backcrossing products can be unsatisfactory. Epistasis or co-adaptation of genes within both breeding population and PGR means that natural or artificial selection has favoured specific combinations of alleles at different gene loci within each type of material. The specific allele combinations are lost after crossing and recombining the two types, leading to so-called 'recombination losses'. It takes several generations to establish new favourable allele combinations through selection.

In summary, all these problems can delay the development and release of a new cultivar when using PGR. This is often not acceptable to breeders because of high competition and the pressure to come up with new cultivars as quickly as possible. Lack of (long-term) financing provided explicitly for the use of PGR in crop improvement is an additional hindrance.

Classical methods of using genetic resources in plant breeding

There are three ways of using PGR in plant breeding (Simmonds, 1993; Cooper *et al.*, 2001):

- introgression involves the transfer of one or few genes or gene complexes (chromosome segments) from the PGR into breeding materials;
- incorporation (also named genetic enhancement or base broadening) describes the development of new, genetically broad, adapted populations with large variation and acceptable performance level;
- prebreeding refers to more basic research activities with the goal of facilitating use of 'difficult' materials.

Nonetheless, the three categories cannot be clearly separated from each other.

Introgression

Introgression aims at improving highly heritable qualitative traits that are governed by one or few major genes or gene complexes. Traditionally, the backcrossing method is used to introgress traits like resistances or restorer genes from wild relatives into breeding materials. The genetic problems mentioned above mostly play a minor role when introgressing major genes from PGR into high-yielding genotypes. Searching for specific traits, breeders would principally consider PGR of the primary gene pool, followed by the secondary gene pool and eventually the tertiary gene pool (Becker, 1993). Due to Darwin's observations on 'parallel variation' which correspond to Vavilov's law of 'homologous series', it is probable that a certain allele will be found in a cultivated form if genetic variation for it exists among wild relatives (Becker, 1993).

Incorporation

Incorporation, genetic enhancement or base broadening aim at increasing the genetic variation for quantitative traits in breeding materials. Various methods of population improvement can be used. The methods will vary

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

depending on the crop species (self- or cross-pollinating) and the available time-frame. Initially, selection may concentrate on adaptation traits that are highly heritable; performance traits are selected at a later stage. Diversity and recombination are maximized in the initial phase, with minimal selection intensities. According to the available time-frame, Cooper *et al.* (2001) identified three methods:

- development of synthetic or composite-cross populations (long term);
- incorporation of PGR in a region's breeding materials to reduce the effects of historical bottlenecks during the evolutional spread of the crop (medium term);
- genetic enhancement to increase the actual variation in breeding populations (short term).

To develop synthetic or composite-cross populations, a large number of accessions of different geographic origin and with maximal genetic diversity are crossed. The resulting population is divided into subpopulations (effective population size $N > 1000$) and subpopulations are grown for up to 30 generations in a number of different environments. At each site, recombination is promoted, and both natural selection and mild mass selection may contribute to adaptation of the individual subpopulations. The sum of all subpopulations has been termed 'mass reservoirs of genetic adaptability' (Simmonds, 1993; Cooper *et al.*, 2001) and is understood as a means of *in situ* maintenance of PGR. Examples are the barley (*Hordeum vulgare*) composite-cross developed at Davis, California (Cooper *et al.*, 2001), dynamic gene pool management in wheat (Goldringer *et al.*, 2001); pearl millet (*Pennisetum glaucum*) composite populations developed in Africa (Niangado, 2001); and the development of locally adapted 'farm cultivars' for ecological agriculture in Europe (Müller, 1989).

To remove the effects of historical bottlenecks during the dissemination of the crop, diverse materials from the crop's centre of origin are crossed, recombined and grown under mild selection for adaptation in the actual target environments. Bottleneck situations have, for example, appeared in potato during migration from the Andean Mountains to temperate zones and during the *Phytophthora infestans* epidemic in Europe. Within the joint potato programme of the John Innes Centre and Cornell University, a large number of 'Andigena' accessions were crossed and grown in the target environments for 10–20 years under weak mass selection. The resulting '*Neotuberosum*' materials show promising performance (Simmonds, 1993; Cooper *et al.*, 2001).

In the short-term genetic enhancement of breeding materials, PGR are selected for agronomic traits and yield performance, but not for the highest degree of genetic diversity. They are intercrossed, recombined and then selected for adaptation to the target environment. To speed up the process, selected PGR may also be crossed with the breeding materials, and selection for yield traits may be carried out in the F_2 (50% exotic genome) or BC_1 (25% exotic) generation. The optimal percentage of the exotic PGR genome (100%, 50% or 25%) in a breeder's population depends on the overall objective, available time-frame and finances, the level of adaptation of the PGR, and the yield difference between PGR and actual breeding population.

A direct use of highly unadapted PGR (100% exotic genome) requires more time and finances; large populations (effective population size $N > 1000$); mild recurrent selection methods with low selection intensity in order not to lose valuable alleles; sufficient recombination (which can be enhanced through systematic polycrossing); and isolation from actual breeding materials (Simmonds, 1993). A disadvantage of the direct adaptation of exotic PGR is the longer period (at least 10–15 years) until useful materials are available as the starting point for variety development. A definite advantage is that the new materials are not related (by descent) to the actual breeding materials; this can be particularly interesting in hybrid or population breeding. Goodman (1999) showed that direct adaptation of exotic elite material can result in materials superior to actual breeding standards. In this study, however, differences in agronomic performance between exotic and adapted maize were small, and mainly adaptation traits with high heritability had to be improved.

If the exotic materials lack alleles for adaptation to the target environment, it is necessary to cross them with adapted materials. The question whether the F_2 or the BC_1 generation is the better choice has been considered by a number of scientists. As an objective criterion, Schnell and Utz (1975) and Schnell (1983) developed the parameter 'Usefulness' (U_α) which refers to the mean genotypic performance of the selected fraction α. They showed that the Usefulness of the F_2 versus BC_1 generation depends on the population mean μ and the expected selection gain $\Delta G(\alpha)$: $U_\alpha = \mu + \Delta G(\alpha)$. The expected selection gain can be defined $\Delta G(\alpha) = i_\alpha \, h\sigma_t$, where i is the selection intensity, h is the square root of the heritability and σ_t is the square root of the genotypic variance (Schnell and Utz, 1975; Schnell, 1983). Obviously, the usefulness of the F_2 versus BC_1 population depends largely on the population mean and its genotypic variance. While adaptation and mean performance are expected to be higher in the BC_1 generation, the genetic variance is anticipated to be higher in the F_2 generation. Dudley (1984) as well as Bridges and Gardner (1987) concluded from their model calculations that

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

the BC_1 is the preferable generation if selection gain is to be made within the short term and if there is a large performance difference between actual breeding materials and PGR. The F_2 generation is superior both in the short and long term if the performance difference between PGR and adapted materials is negligible. Increased recombination should theoretically help to break linkages between desired and undesired alleles at two neighbouring gene loci, and in fact, a recombined BC_1–F_2 population derived from a cross of the Illinois Stiff Stalk Synthetic Composite with the South African Photoperiod Insensitive Composite II proved to be the most favourable foundation population for grain yield in maize (Albrecht and Dudley, 1987). However, repeated recombination of the F_2 had no significant effect on mean performance and variance in a study comparing BC_1 and F_2 foundation populations derived from 18 different crosses of adapted × exotic maize dent lines (Šimić *et al.*, 2003).

One example for genetic enhancement is the GEM project in maize, where elite cultivars from the Tropics or Europe are used to improve US materials for agronomic traits, resistance and quality (Cooper *et al.*, 2001; http://www.public.iastate.edu/~usda-gem/methods.html). Another, long-term base-broadening project in maize is the Hierarchical Open-ended Population Enrichment (HOPE) project in Canada (Kannenberg and Falk, 1995; Kannenberg, 2001).

Prebreeding and wide crosses

Prebreeding includes basic research to achieve wide crosses, and activities that facilitate the use of exotic materials or wild relatives. It can refer to both qualitative and quantitative traits and the distinction between prebreeding, introgression and incorporation is not always clear. The main objective is to provide breeders with more 'attractive' PGR that are easier to use, i.e. resistance sources in acceptable genetic background; or inbreeding-tolerant forms of outcrossing species for hybrid breeding. An example is the resistance breeding programme of the International Potato Center (CIP), within its Global Initiative on Late Blight (GILB) (Trognitz *et al.*, 2001; http://www.cipotato.org/gilb/index.htm), whose aim is to provide breeders and farmers with new sources of resistance from wild relatives. Also prebreeding activities in wheat (Valkoun, 2001) and rice (Brar and Khush, 2002) aim at transferring resistances to major diseases and insects, and tolerance to abiotic stresses, from wild relatives into cultivated forms using backcrossing and embryo rescue techniques. Homology of chromosomes is used here to incorporate single chromosomal segments carrying the resistance genes from wild relatives into elite varieties. In sugar beet (*Beta vulgaris*), a number of commercial breeders are involved in prebreeding activities, i.e. in the development of 'Base' or 'Buffer' populations from genetically extremely diverse materials (Frese *et al.*, 2001; Frese, 2002). Another example for a very innovative use of wide crosses is New Rice for Africa (NERICA) developed by the Africa Rice Center (WARDA, http://www.warda.org/). Through crossing the African rice, *Oryza glaberrima*, with Asian rice, *O. sativa*, embryo rescue and, most importantly, farmer-participatory variety selection, new rice cultivars were developed that combine positive characters (high grain yield and resistances to pests and diseases) of both rice species (http://www.warda.org/warda1/main/Achievements/nerica.htm).

One persistent question related to prebreeding is the funding question. It is increasingly difficult to acquire long-term funding for such activities whose impact may be visible only after several decades. In addition, it is unclear who should be taking the responsibility and initiative—genebanks, public or private institutions? With regard to financing, it may be worth considering support from the Global Conservation Trust (http://startwithaseed.org/) once it is established. This planned foundation for food security aims at the long-term and sustainable financing of *ex situ* conservation of PGR, 'to make them easier to use and thus more useful to farmers and professional breeders'.

Use of landraces in breeding for specific adaptation to stress environments

Breeding for wide adaptation has been found to be inappropriate for extreme stress environments, because of cross-over genotype × environment interactions appearing at low yield levels (e.g. Simmonds, 1991; Ceccarelli *et al.*, 2001; vom Brocke *et al.*, 2002a, b). Cross-over genotype × environment interactions represent the situation where newly bred 'widely adapted' cultivars are inferior to local, indigenous varieties under extreme environmental conditions. Such interactions may be considered as a hindrance to crop improvement in a target region, but they also offer new opportunities, e.g. selecting and using genotypes that show positive interaction with the location and its prevailing environmental conditions (exploitation of specific adaptation) or genotypes characterized by low frequency of crop failure (Annicchiarico, 2002).

Landraces grown in extreme areas, e.g. semi-arid to arid regions in Asia and Africa, can represent important PGR in breeding for specific adaptation (Hawtin *et al.*, 1997). They can be: donors for individual monogenic

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

traits; sources of new quantitative variation for specific adaptation to stress conditions; and breeding population or crossing partner in the development of improved, locally adapted cultivars for the same or other marginal areas. Strategies for the development of locally adapted germplasm include:

- decentralization of the breeding process from the international to the national level and from stations to farmers' fields;
- crossing of elite materials with locally adapted, farmer-preferred cultivars; development of different breeding populations for different regions;
- distribution of segregating materials to national programmes;
- farmer-participatory selection, to increase final acceptance of the improved cultivars (Ceccarelli *et al.*, 2001; Witcombe, 2001).

Utility of molecular markers and genome research for using PGR in crop improvement

'The tools of genome research may finally unleash the genetic potential of our wild and cultivated germplasm resources for the benefit of the society' (Tanksley and McCouch, 1997). The utility of molecular markers and genome research in the context of using PGR for crop improvement include:

- diversity studies to identify genetically similar or distinct accessions, and to determine individual degrees of heterozygosity and heterogeneity within populations of PGR;
- genetic mapping to identify simply inherited markers in close proximity to genetic factors affecting quantitative traits (QTL), followed by marker-assisted selection (MAS) of desired genotypes in segregating populations;
- exploitation of valuable QTL from PGR by advanced backcross QTL analysis to combine QTL analysis with the development of superior genotypes or by marker-assisted, controlled introgression of PGR into breeding materials through the development of introgression libraries;
- association studies to mine directly the allelic diversity of PGR and to identify those alleles that are beneficial for important agronomic traits.

Recent publications examining the available technologies and their application in the analysis of wild plant populations, germplasm collections and plant breeding include those by Callow *et al.* (1997), Henry (2001) and Newbury (2003).

Diversity assessment

Diversity assessment includes the (i) inventory and successive monitoring of diversity of PGR *in situ* and *ex situ* for maintaining appropriate genetic variance and establishing core collections; (ii) assessment of the mating system and the population structure along with its dynamics of locally adapted PGR; and (iii) heterotic grouping of potential lines suitable for hybrid breeding approaches. The analysis of diversity data is based on population genetic theory and mainly requires allele frequency data. Initially, diversity studies were based on morphological and agronomical traits. The increasing availability of molecular marker systems opened up new possibilities for the diversity assessment of PGR intended to be used for crop improvement (Bretting and Widerlechner, 1995; Karp *et al.*, 1997, 1998). For an efficient diversity assessment, molecular markers ideally need to be selectively neutral, highly polymorphic, codominant, well-dispersed throughout the genome, and cost- and labour-efficient (Bretting and Widerlechner, 1995; Van Treuen, 2000). Genetic markers complying with these requirements are protein markers (i.e. isoenzymes) and DNA markers such as restriction fragment length polymorphisms (RFLPs) and microsatellites or simple sequence repeats (SSRs). Because the development of the latter two marker types requires prior knowledge of DNA sequences, a number of universal, dominant molecular marker types such as randomly amplified polymorphic DNA (RAPD) and amplified fragment length polymorphisms (AFLPs) have also been employed in PGR diversity studies. However, the latter are not suitable to assess, for instance, the mating behaviour or heterozygosity of PGR. Different indices exist for the measurement of diversity, partitioning of diversity within and between crop populations and the genetic distance between them (i.e. differentiation). They have recently been reviewed by Mohammadi and Prasanna (2003). It seems noteworthy that comparing data achieved with different molecular marker types and even measured at different marker loci of the same type is ambiguous as diversity measures are relative rather than absolute (Ennos, 1996).

To analyse the significance of different marker types, crop plants and geographic regions used in diversity studies with relevance to plant breeding, a survey of the literature was conducted in the Commonwealth Agricultural Bureau Abstracts (CAB Abstracts, 1984–2003) for all of the three following keywords 'genetic diversity', 'plant genetic resources' and 'plant breeding' which yielded a total of 2432 records for the period 1984–2003 (Tables 1 and 2). Although the figures in Tables 1 and 2 are not exhaustive, they may exemplify some trends. The majority of studies are based on isoenzyme

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

Table 1. Significance of selected marker types and plant species mentioned in published diversity studies with relevance to plant breeding between the years 1984 and 2003[a]

Selected marker types	Hits[b]	Selected crop groups	Hits	Selected crop species	Hits
Isozymes/allozymes	517	Trees	464	Wheat	197
RAPDs	313	Cereals	327	Rice	137
Morphological traits	259	Legumes	253	Barley	103
Microsatellites/SSRs	109	Fruits	185	Maize	92
RFLPs	104	Vegetables	108	Potato	64
AFLPs	60			Soybean	47
SNPs	2			Sorghum	39
				Cassava	21
				Sunflower	12
				Banana	11

[a] Source: CAB Abstracts.
[b] Total hits do not add up to 2432 because of selective search criteria.

Table 2. Changes of the significance of selected marker types and geographic regions mentioned in published diversity studies with relevance to plant breeding between the years 1984 and 2003[a]

	Periodicals (number of years)				
	1984–1989 (6)	1990–1994 (5)	1995–1998 (4)	1999–2003 (5)	1984–2003 (20)
No. of diversity studies published	17	417	841	1157	2432
Selected marker types used[b]					
Isoenzymes	2	97	242	212	517
RAPD	0	14	114	185	313
SSR	0	1	20	88	109
Geographical origin[bc]					
Asia	3	110	260	386	759
Europe	2	56	171	227	456
Africa	1	51	103	115	270
North America	2	55	110	89	256
South America	3	32	47	68	150
Australia	0	21	46	59	126

[a] Source: CAB Abstracts.
[b] Total hits do not add up to 2432 because of selective search criteria.
[c] Geographical origin of plant species inferred from occurrence of the search term in abstracts.

markers and RAPDs which have been criticized (Wolfe and Liston, 1999) for not being selectively neutral (isoenzymes) or having a low reproducibility (RAPDs). While RAPD-based studies are still on the increase, it seems that isoenzyme-based studies have reached a peak and are now on the decrease (Table 2). Diversity studies based on SSRs seem to increase, which may be due to the availability of DNA-sequence data and improved and more efficient techniques (i.e. use of automated sequencers, multiplexing).

Diversity studies comprised a wide range of species of which surprisingly tree species accounted for the majority followed by cereals (wheat, rice and barley) and legumes. Over the last 20 years, the number of diversity studies conducted increased almost linearly and there seems to be no end to this trend. Figures in Table 2 suggest that the development of the isoenzyme marker technique may have triggered research activities in this field considerably. The number of hits for the geographical origin of plant species used in these studies may reflect their quantitative occurrence in the centres of diversity, however, regional differences are substantial and it seems justified to conclude that diversity studies of PGR of African origin are underrepresented and have not increased at the same rate as for Asian or European PGR (Table 2).

About 772 examples of studies assessing the population structure and its dynamics of locally adapted PGR were found in a further search of CAB Abstracts (1984–2003), of which almost half (355) were dealing with tree species (e.g. Grattapaglia *et al.*, 1998). The remainder included a wide range of plant species from neglected cereals such as pearl millet (vom Brocke *et al.*, 2002a) to legumes such as lima beans (*Phaseolus lunatus*; Maquet *et al.*, 1997) and wild relatives of cereals (Nevo, 1998). The focus of this type of diversity study

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

seems to be on crop evolution and conservation rather than on immediate crop improvement.

A different objective of molecular diversity studies is heterotic grouping of genotypes suitable for hybrid breeding approaches. The principle behind this approach is the search for a correlation between genetic distance and heterosis, i.e. the more distant two genotypes of a crop species are genetically, the more heterozygosity and therefore heterosis can be expected in the hybrid resulting from a cross between them (Melchinger *et al.*, 1999). Yet, the effect on heterosis and hybrid performance needs to be distinguished, since high heterosis does not necessarily mean high hybrid yield. The prediction of hybrid performance from genetic distances based on evenly dispersed molecular markers seems to be complicated because yield QTL are located in particular chromosomes and are not distributed evenly over the genome (Jordan *et al.*, 2003). However, if genotypes are somewhat related (intra-pool versus inter-pool hybrids), the correlation between genetic distance and hybrid performance seems to increase (Boppenmaier *et al.*, 1993). A study of maize (Vuylsteke, 1999; Vuylsteke *et al.*, 2000) based on 53 inter-pool hybrids showed that diversity measures related to those markers that were linked to yield QTL explained 59–62% of the variation in hybrid performance for grain yield. Diversity measures related to markers spread across the whole genome explained in the same set of hybrids 28–44% of the hybrid performance only. This method of correlating the trait of interest with a hybrid value that represents the total contribution of selected markers of the respective trait seems to have a good potential for use in the selection of hybrids with high heterosis. In the meantime, this was verified in sorghum (*Sorghum bicolor*; Jordan *et al.*, 2003). A model combining phenotypic trait data and parental diversity on particular linkage groups explained 71% of the variation in grain yield.

In terms of using PGR in hybrid breeding, a two-step procedure seems appropriate: (i) establishing heterotic groups based on molecular markers and (ii) testing the performance of hybrids produced from crosses among representative genotypes from each heterotic group. The approach has been proven useful for outbreeding species such as maize (Reif *et al.*, 2003a, b) and seems promising for clonal species (i.e. asexually propagated hybrids; Schnell, 1978) and even for inbreeding crop species such as wheat (Jordaan, 1998; Cui *et al.*, 2002) and rice (Zhang *et al.*, 1995). Corroborating this appraisal, a literature search in CAB Abstracts (1984–2003) yielded 140 hits for the term 'heterotic groups' of which 116 related to maize. The remaining 24 studies related to crops such as sunflowers (*Helianthus annuus*), legumes, wheat, rice, brassicas and bananas (*Musa* spp).

Co-ordinated efforts by molecular geneticists, breeders and curators are required to exploit fully the molecular marker information gained in diversity studies in the final crop improvement. An example of extensive and co-ordinated networking among scientists from 11 countries in Africa, North and South America, and Europe is the Cassava Molecular Diversity Network (MOLCAS). The goals of MOLCAS are to enhance cassava (*Manihot esculenta*) productivity by: collection and molecular marker study of genetic variation and heterotic grouping of cassava landraces in Africa and Latin America; elucidation of the genetic diversity and differentiation of landraces in Africa compared to what exists in the crop's primary centre of diversity; exploitation of this information in systematic improvement of the crop including hybrid breeding; capacity building in the sub-Saharan region for molecular diversity assessment and interpretation (http://www.ciat.cgiar.org/molcas/).

Genetic mapping and marker-assisted selection

Marker-assisted selection (MAS) can help (i) to select individuals carrying molecular markers that are linked to the trait of interest instead of performing extensive phenotypic tests (foreground selection) and (ii) to reduce undesired parts of the donor genome including the linkage drag (background selection). Foreground selection requires a tight linkage between the trait of interest and its flanking markers that are being selected for. Background selection necessitates genotyping with a larger number of markers that cover the whole genome.

MAS has proven efficient for the transfer of simply inherited qualitative traits from PGR into elite materials using backcrossing procedures. It is particularly useful for traits that are recessive, that can be assessed only after flowering and that are very difficult and expensive to assess. An example is MAS for fertility restoration in rye. To restore male fertility in cytoplasmic genic male sterility (CMS)-based rye hybrids restorer genes are urgently needed, but absent in adapted materials (Geiger *et al.*, 1995). To evaluate this trait, a 2-year test-crossing and evaluation procedure is crucial. Effective restorer genes were recently detected in materials originating from old Argentinian rye cultivars and Iranian primitive rye accessions. Only one to two loci of each of these sources explain 70% of the phenotypic variation (Miedaner *et al.*, 2000). After establishing PCR-based markers for two restorer loci, this trait could be successfully introgressed into hybrid rye breeding materials (Stracke *et al.*, 2003).

By using a combination of foreground and background selection, the transfer of a monogenic trait from a PGR into a breeding line may be completed within three to four instead of the usual six generations of backcrossing

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

with the same proportion of the recurrent parent genome (Ragot *et al.*, 1995; Frisch *et al.*, 1999a). Frisch *et al.* (1999a, b, 2000, 2001a–c) developed models and strategies for optimal application of MAS (both foreground and background selection) to transfer one or two genes from a donor PGR into a recipient genotype. Important parameters that were optimized for given marker intervals around the target gene(s) are the minimum number of individuals to be genotyped, the minimum number of data points in the genotyping and the allocation of marker analyses to different backcross generations.

MAS for multigenic, quantitative traits at first requires the identification of the genomic regions (QTL) that affect the trait of interest. In classical QTL mapping, a segregating population (e.g. F_2, F_3 or recombinant inbred population) is developed from two inbred lines. This mapping population is evaluated for the trait(s) of interest. Simultaneously, the population is genotyped with a number of markers and a genetic map is constructed from the marker data. In the final QTL analysis, data are analysed for co-segregation of particular markers with the trait of interest. Several recent reviews of QTL mapping techniques are available: Broman (2001), Knapp (2001), Doerge (2002), Hackett (2002) and Kearsey (2002). QTL analysis is then followed by transfer of favourable QTL alleles into elite materials via pure MAS or MAS combined with phenotypic selection.

However, for complex, quantitative traits, the efficiency of QTL mapping and MAS is not uncontested. There are a number of risks which can result in MAS becoming a 'money-absorbing system' (Melchinger and Utz, personal communication). For example, there may be no selection gain because of: unreliable QTL estimates (too few QTL with highly overestimated effects); QTL not being expressed in new genetic backgrounds; recombination between marker and QTL; unfavourable alleles of other genes linked to good QTL alleles; too high costs for marker analyses. It is therefore essential to: use large mapping populations; genotype the mapping population with good genome coverage; assess phenotypic values in multi-environment field trials; cross-validate the gained data; verify QTL effects using independent population samples, near-isogenic lines or different genetic backgrounds; assure close linkage between marker and QTL and verify the linkage by a phenotypic test of all three to four generations; increase the marker density around the QTL to allow reduction of the linkage drag; and to optimize individual procedures while taking into account economic parameters (Geiger and Welz, 2000). For quantitative traits, where many loci of minor effects are responsible, it is very difficult to obtain reliable, unbiased QTL estimates (e.g. Beavis, 1998; Melchinger *et al.*, 1998; Utz *et al.*, 2000). Prospects for MAS are therefore more promising for oligogenic traits that are largely determined by few QTL with large effects (Melchinger, 1990).

One example for successfully using molecular markers for introgressing highly valuable genes from a PGR is the resistance to *Fusarium* head blight derived from the Chinese wheat variety Sumai 3 from subtropical environments. Due to its high effectiveness, this source is used in the temperate wheat-growing areas of North America and Europe. Grain yield, lodging resistance and quality are, however, very poor and Sumai 3 is highly susceptible to powdery mildew and cereal rusts. It has a prominent QTL for head blight resistance on chromosome 3BS, explaining 15–40% of the phenotypic variance, depending on the crossing partner and the test environments. An additional QTL was detected on chromosome 5A, explaining 23% of the phenotypic variance (Buerstmayr *et al.*, 2003). Both QTL could be readily detected in other crosses involving a Sumai 3 descendant and were successfully used for MAS in elite spring wheat material (Miedaner, unpublished data). MAS allows rapid introgression with three generations per year, while for phenotype-based selection, a time-consuming field test with artificial infection at and several disease scoring dates after flowering is required.

Advanced backcross QTL analysis and introgression libraries

QTL analysis can also be performed in backcross generations derived from crosses of exotic PGR with elite materials. The advanced backcross QTL analysis (AB-QTL; Tanksley and Nelson, 1996) combines QTL analysis with the development of superior genotypes and has been shown to be particularly useful for a trait transfer from poorly adapted germplasm. AB-QTL is therefore of special importance in the use of PGR for crop improvement. The starting point is a segregating generation of a cross between an exotic parent and an elite line that is analysed with as many molecular markers as possible. The QTL mapping procedure is delayed until one of the advanced backcross generations ($\geq BC_2$) when lines or test-crosses are evaluated across environments.

To date, the AB-QTL strategy has been applied in several crops such as tomato, rice and barley. Favourable QTL alleles originating from four different wild relatives were identified for important agronomic traits in tomato, improving the fruit yield in the range from 17 to 34% (Tanksley *et al.*, 1996; Fulton *et al.*, 1997, 2000; Bernacchi *et al.*, 1998a). In two AB-QTL studies in rice, two wild-species QTL alleles on chromosomes 1 and 11 were associated with a yield increase of 17% and 18%, respectively (Xiao *et al.*, 1996, 1998). The QTL on chromosome 1 was validated in an additional cross using the same *Oryza rufipogon* donor accession (Moncada *et al.*,

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

2001). A recent AB-QTL study of the introgression from *Hordeum spontaneum* into cultivated barley (Pillen *et al.*, 2003) reported several valuable donor QTL for quantitative traits, such as the number of days to heading and lodging at flowering stage. In one case, the *Hordeum spontaneum* allele was associated with a yield increase of 7.7% averaged across six test environments.

Once favourable QTL alleles from an exotic donor are identified, one or two additional backcrossing and selfing generations are needed to derive QTL-bearing near-isogenic lines (QTL-NILs). These carry recurrent parent alleles throughout their genome except for the specific target QTL (Tanksley and Nelson, 1996). QTL-NILs can be used to verify observed QTL effects as well as commercial lines improved for one or more quantitative traits compared to the original recurrent elite line. Bernacchi *et al.* (1998b) validated the effects of exotic tomato QTL in QTL-NILs. In field evaluations at five locations worldwide, 22 QTL-NILs out of 25 tested showed the phenotypic improvement predicted in the previous AB-QTL analysis.

In contrast to the AB-QTL method, Eshed and Zamir (1994) suggested the approach of establishing a population of NILs such that the donor chromosome (DC) segments are evenly distributed over the whole recipient genome. Ideally, the total genome of the exotic donor is comprised in the established set of NILs. This NIL population, named introgression library, consists of a set of lines, each carrying a single marker-defined DC segment introgressed from an agriculturally unadapted source into the background of an elite variety (Zamir, 2001).

The procedure of establishing an introgression library implies systematic transfer of DC segments from a PGR (donor) into an elite line (recurrent parent) by marker-aided backcrossing. Additional self-pollination and marker-based selection lead to NILs homozygous at DC segments. Such NILs differ from the elite line by only a small, defined chromosomal segment, and phenotypic differences between a line in the library and the nearly isogenic elite line are associated with the single DC segment.

The approach of using an introgression library in broadening the genetic base of breeding material was firstly applied to tomato. Eshed and Zamir (1995) established an introgression line population originating from a cross between cultivated tomato and the wild species *Lycopersicon pennelli*. They genotyped 600 BC_1 lines with RFLP markers and selected 50 introgression lines in which each line carried a different DC segment. In total, the line population covered 97.5% of the exotic donor genome. The lines were further backcrossed and selfed, and the resulting BC_2S_1 population was subjected to QTL analysis. On chromosome 1, a donor QTL allele increasing the soluble solid yield was detected and, in the meantime, cloned (Fridman *et al.*, 2000). Further introgression libraries in plants have been established in *Brassica oleracea* (Howell *et al.*, 1996; Ramsay *et al.*, 1996) and *L. hirsutum* (Monforte and Tanksley, 2000). Two rye marker-based introgression libraries derived from a cross between an elite inbred line and an Iranian primitive rye (Altevogt 14160) are considered a powerful and highly efficient tool to characterize and exploit genetic resources in rye breeding (Miedaner and Sušić, unpublished).

Both introgression library and AB-QTL approaches provide a valuable opportunity to extract quantitative trait alleles for modern crop varieties from exotic PGR. Their main advantage is that the exotic genome is introgressed into the elite line only as small, well-defined DC segments. This reduces unfavourable effects that often impede the use of PGR in practical breeding programmes. If less-adapted exotic resources are used, the introgression library approach might be more effective due to the fact that the final NILs of the introgression library carry a lower number of shorter and more precisely marker-defined DC segments than QTL-NILs obtained by AB-QTL analysis. The disadvantage of a higher number of longer DC segments in QTL-NILs lies in the higher linkage drag which can mask positive effects of favourable DC segments. Additionally, the whole donor genome is recognized and chances to detect most of the valuable segments are higher in the introgression libraries. Nevertheless, the AB-QTL approach is usually less expensive because it requires only one generation of marker analysis. Both approaches are recommended as starting points for high-resolution mapping and the isolation and functional characterization of candidate genes located in DC segments of economic importance. However, higher inputs in marker analyses for introgression library development make these NILs more useful in respect of further research towards functional genomics. Which of these two approaches will be applied in breeding programmes for improving quantitative traits depends mostly on the used PGR, as well as on the available budget.

Association studies and direct allele selection

The increased insight into the molecular organization and sequence of plant genomes leads to new methods to mine directly the allelic diversity of PGR. The aim of such studies is to associate sequence polymorphisms within genes or across genomes with phenotypic variants to detect superior alleles affecting agronomically important traits. Such valuable alleles detected within germplasm collections can subsequently be transferred to elite breeding materials via marker-assisted backcrossing using allele-specific markers (direct allele selection; Sorrells and Wilson, 1997). The major advantage

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

of association studies over classical QTL mapping experiments is that no segregating population has to be established from two inbred lines and that the results are not limited to the specific mapping population but can cover the full allelic variation available in natural or breeding populations or genebank accessions (Jannink *et al.*, 2001; Jannink and Walsh, 2002).

Associations between DNA sequence polymorphisms and phenotypic trait variation can either occur when the polymorphisms are directly responsible for the functional differences between the alleles of the respective genes or when the analysed polymorphisms are in linkage disequilibrium (LD) with the functional alleles. LD is a statistical association of alleles at two loci within a population (Falconer and Mackay, 1996). Different measures exist to describe the extent of LD within a population. One of the most widely used measures is the square of the correlation coefficient (r^2) between two loci (Pritchard and Przeworski, 2001; Ardlie *et al.*, 2002). In large populations under random mating, LD decays over time due to the effect of recombination. Factors such as mutation and migration or population admixture will create LD, and inbreeding, genetic drift and selection (co-adaptation) will maintain or increase levels of LD. In 102 maize inbred lines, Remington *et al.* (2001) observed a rapid decay of $r^2 < 0.1$ within 1.500 bp, but rates of decline were variable among the six genes analysed. For the selfing species *Arabidopsis thaliana*, Nordborg *et al.* (2002) found that LD decayed in global samples within 250 kb, which is equivalent to 1 cM. However, Tian *et al.* (2002) observed a complete LD decay within less than 10 kb. As a preliminary result one may conclude that extent and structure of LD depend on the species, the population and its demographic history (e.g. population bottlenecks or founder effects) and the genomic region studied (Nordborg and Tavaré, 2002; Flint-Garcia *et al.*, 2003). In plant populations, a larger degree of LD is expected to be maintained in selfing compared to outcrossing species (Nordborg, 2000). Furthermore, studies in human populations revealed that recombination frequencies are not evenly distributed across the genome (Wall and Pritchard, 2003). To size up the feasibility of association mapping in plant species, further studies on the LD structure in natural and breeding populations have to be conducted. Software for analysing LD in populations is available on the web (e.g. Arlequin, http://lgb.unige.ch/arlequin/; DNASP, http://www.ub.es/dnasp/).

The basic idea of association mapping can be investigated using two strategies. One approach is first to identify candidate genes (i.e. from available data bases or gene expression studies) and to re-sequence those candidate genes in plants derived from diverse germplasm accessions. The maize gene *dwarf8*, a candidate gene for flowering time and plant height, was used by Thornsberry *et al.* (2001) in a first association study with a crop species. They sequenced *dwarf8* in a representative set of 92 inbred lines and found polymorphisms within the gene to be strongly associated with flowering time. This group of researchers also developed the software TASSEL (http://www.maizegenetics.net/bioinformatics/index.htm) for analysing LD and for performing association mapping in populations of inbred lines. Österberg *et al.* (2002) conducted a detailed comprehensive study on *Brassica nigra*, where variation in the *COL1* gene showed up to be associated with flowering time. The candidate gene approach has been extensively used in human populations to screen various disease genes. Lohmueller *et al.* (2003) gave a detailed review on the reproducibility of such studies.

The second approach is to analyse a set of randomly chosen molecular markers, evenly distributed across the genome. If such markers are in LD with the genes controlling the trait variation, one will also detect a significant association. The practicability of this approach strongly depends on the level and structure of LD. Low levels of LD would be favourable for a high resolution fine mapping within candidate genes, but limit the feasibility of genome-wide association studies. For the human genome it was estimated that between 300,000 and one million single nucleotide polymorphism (SNP) markers will be necessary to scan the genome for associations (Gabriel *et al.*, 2002). So many factors seem to influence the success of such an approach that, at the current stage, it is challenging to design an appropriate study. A first attempt to use this approach in plants was reported for *Beta vulgaris* ssp. *maritima* using 440 AFLP markers in 106 individual plants from four natural populations (Hansen *et al.*, 2001). Two markers were detected showing significant association with the bolting gene, which is responsible for the vernalization requirement. The bolting gene mapped to a region with suppressed recombination which was known from previous studies. Therefore, LD in this particular region was expected to be extensive. The feasibility of this approach to map unknown genes responsible for phenotypic trait variation in plant species has to be proven in further studies.

Two different experimental designs are mainly considered for association studies: the case-control study and the transmission disequilibrium test (TDT). Both designs are suited for the candidate gene approach as well as for genome scans. A brief review of the two designs and their analysis was published by Lewis (2002). Case-control studies compare the allele frequencies of two groups of individuals sampled from the population(s). Groups are distinguished on the basis of their divergent phenotypes (e.g. expression of a disease) and allele frequencies are compared across cases and

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

controls. In case-control studies, population structure in germplasm accessions, which may be unknown to the researcher, can cause spurious associations. Statistical methods were developed by Pritchard *et al.* (2000) and Falush *et al.* (2003) to detect such population structures using a few molecular markers evenly spread across the genome. Relevant software is available at: http://pritch.bsd.uchicago.edu/. Another way to avoid the problem of population structure is the TDT. This design uses family triplets, analysing an offspring that expresses the trait of interest, and both parents. The TDT tests for distortion in transmission of alleles from a heterozygous parent to the offspring. One disadvantage of the TDT compared to the case-control design is the higher number of individuals necessary for genotyping. Many programs can analyse TDT data, one example is the software from Spielman *et al.*: http://genomics.med.upenn.edu/spielman/TDT.htm (Spielman *et al.*, 1993; Spielman and Ewens, 1998).

Jannink *et al.* (2001) and Bink *et al.* (2002) suggested that the association mapping approach be combined with QTL mapping in segregating F_1-derived populations. They suggested working with extended pedigrees founded by multiple individuals to consider the allelic diversity available in natural and breeding populations, and to make useful alleles available for crop improvement.

In conclusion, current advances in genomic research could finally lead to genetic resources collections becoming 'genebanks' in the truest sense of the word (Kresovitch *et al.*, 2002).

Gene transfer

Gene transfer is independent on crossing barriers and may therefore increase the usable genetic variation of and beyond the tertiary gene pool. The principal steps for a transfer of genes from any species into cultivated crops are: gene isolation, gene cloning, gene transfer and final expression studies in greenhouse and field trials across several generations of progeny. The details of gene transfer go beyond the scope of this review. Recent reviews on the topic include those by Repellin *et al.* (2001), Bhat and Srinivasan (2002), Francois *et al.* (2002), Galili *et al.* (2002) and Lessard *et al.* (2002). Within the next 10–20 years, transformation research hopes to reach the following goals: controlled integration and stable expression of transferred genes; targeted manipulation of multigenic characters; efficient production of transgenes; transgenes without or with harmless selection markers (Lütticke, personal communication); and efficient transformation of cell organelles to assure maternal inheritance and therefore avoid unwanted horizontal gene transfer (Daniell *et al.*, 2002).

Classical examples for the use of gene transfer are the improvement of insect resistance through transfer of *bt* genes from *Bacillus thuringiensis* in crops like tobacco, tomato, maize, rice, cotton and soybean; the improvement of virus resistance through transfer of viral coat proteins in tomato and potato; and the creation of herbicide-resistant crops through transfer of bacterial or fungal genes into sugar beet, tomato and rape seed. There are also increasing efforts to improve stress tolerance of crops through transfer of genes for increased osmoregulation, heat shock proteins, phytohormone synthesis and other traits from different organisms into cultivated plants. More information and numerous references on genetic engineering of stress tolerance can be found on the website http://www.plantstress.com.

Outlook

Integrated approaches are necessary to increase diversity in agriculture. Therefore, plant breeders should closely co-operate with farmers to see which crops best fit into current farming systems. In addition, collaboration with food scientists, food technologists and the industry in general may open new uses of selected crops or cultivars and therefore open production alternatives for the farmers. As shown in this review, numerous different methods are available for the use of PGR in crop improvement. The choice mainly depends on the crop, the trait(s) of interest, availability of molecular markers, the chosen time-frame and on the available finances. A combination of advanced, molecular techniques with classical and farmer-participatory breeding methods will most likely achieve the desired impact. In order to enhance the utilization of PGR in crop improvement, the Global Plan of Action (FAO, 1996b) proposed a number of measures, among them expanded creation, characterization and evaluation of core collections; increased genetic enhancement and base-broadening efforts; development and commercialization of underutilized species; development of new markets for local varieties and 'diversity-rich' products and concomitant efficient seed production and distribution; comprehensive information systems for PGR; and promoting public awareness of the value of PGR for food and agriculture. All of us, farmers, breeders, agronomists, politicians and donor agencies should remember these recommendations and contribute to their implementation in an integrated manner.

Acknowledgements

The financial support of Bettina I. G. Haussmann by the German Research Foundation (DFG) is gratefully

acknowledged. Sincere thanks are extended to Dr Dale E. Hess (Purdue University) for proof-reading the manuscript before submission. The authors also wish to thank two anonymous reviewers for their most useful comments on the manuscript.

References

Albrecht B and Dudley JW (1987) Evaluation of four maize populations containing different proportions of exotic germplasm. *Crop Science* 27: 480–486.

Annicchiarico P (2002) Defining adaptation strategies and yield stability targets in breeding programmes. In: Kang MS (ed.) *Quantitative Genetics, Genomics and Plant Breeding*. Wallingford: CAB International, pp. 365–383.

Ardlie KG, Kruglyak L and Seielstad M (2002) Patterns of linkage disequilibrium in the human genome. *Nature Reviews Genetics* 3: 299–309.

Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: Paterson AH (ed.) *Molecular Dissection of Complex Traits*. New York: CRC Press, pp. 145–162.

Becker H (1993) *Pflanzenzüchtung*. Stuttgart: Verlag Eugen Ulmer.

Bernacchi D, Beck Bunn T, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D and Tanksley S (1998a) Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theoretical and Applied Genetics* 97: 170–180.

Bernacchi D, Beck Bunn T, Emmatty D, Eshed Y, Inai S, Lopez J, Petiard V, Sayama H, Uhlig J, Zamir D and Tanksley S (1998b) Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. *Theoretical and Applied Genetics* 97: 381–397.

Bhat SR and Srinivasan S (2002) Molecular and genetic analysis of transgenic plants: considerations and approaches [Review]. *Plant Science* 163: 673–681.

Bink MCAM, Uimari P, Sillanpää MJ, Janss LLG and Jansen RC (2002) Multiple QTL mapping in related plant populations *via* a pedigree-analysis approach. *Theoretical and Applied Genetics* 104: 751–762.

Boppenmaier J, Melchinger AE, Seitz G, Geiger HH and Herrmann RG (1993) Genetic diversity for RFLPs in European maize inbreds: III. Relation to performance of crosses within *versus* between heterotic groups for grain traits. *Plant Breeding* 111: 217–226.

Brar DS and Khush GS (2002) Transferring genes from wild species into rice. In: Kang MS (ed.) *Quantitative Genetics, Genomics and Plant Breeding*. Wallingford: CAB International, pp. 197–217.

Bretting PK and Widerlechner MP (1995) Genetic markers and horticultural germplasm management. *Hortscience* 30: 1349–1356.

Bridges WC and Gardner CO (1987) Foundation populations for adapted by exotic crosses. *Crop Science* 27: 501–506.

Broman KW (2001) Review of statistical methods for QTL mapping in experimental crosses. *Lab Animal* 30: 22–52.

Brown AHD (1995) The core collection at the crossroads. In: Hodgkin T, Brown AHD, van Hintum TJL and Morales EAV (eds) *Core Collections of Plant Genetic Resources*. Rome: IPGRI/John Wiley and Sons, Sayce Publishing, pp. 3–19.

Buerstmayr H, Steiner B, Hartl L, Griesser M, Angerer N, Lengauer D, Miedaner T, Schneider B and Lemmens M (2003) Molecular mapping of QTLs for *Fusarium* head blight resistance in spring wheat. II. Resistance to fungal penetration and spread. *Theoretical and Applied Genetics* 107: 503–508.

CAB Abstracts (1984–2003) *Commonwealth Agricultural Bureau Abstracts*. Wallingford: CAB.

Callow JA, Ford-Lloyd BV and Newbury HJ (eds) (1997) *Biotechnology and Plant Genetic Resources: Conservation and Use*. Wallingford: CABI Publishing.

Campbell CL and Madden LV (1990) *Introduction to Plant Disease Epidemiology*. New York: John Wiley & Sons.

Ceccarelli S, Grando S, Amri A, Asaad FA, Benbelkacem A, Harrabi M, Maatougui M, Mekni MS, Mimoun H, El-Einen RA, El-Felah M, El-Sayed AF, Shreidi AS and Yahyaoui A (2001) Decentralized and participatory plant breeding for marginal environments. In: Cooper HD, Spillane C and Hodgkin T (eds) *Broadening the Genetic Base of Crop Production*. Wallingford: CABI Publishing in co-operation with FAO and IPGRI, CAB International, pp. 115–135.

Cooper HD, Spillane C and Hodgkin T (2001) Broadening the genetic base of crops: an overview. In: Cooper HD, Spillane C and Hodgkin T (eds) *Broadening the Genetic Base of Crop Production*. Wallingford: CABI Publishing in co-operation with FAO and IPGRI, CAB International, pp. 1–23.

Cui GH, Ni ZF, Wu LM, Li YQ and Sun QX (2002) Study on wheat heterotic groups. V. Relationship between micro-satellite marker-based genetic distance and heterosis between *Triticum aestivum* and *T. spelta*. *Journal of Triticeae Crops* 22: 5–9.

Daniell H, Khan MS and Allison L (2002) Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology [Review]. *Trends in Plant Science* 7: 84–91.

Doerge RW (2002) Mapping and analysis of quantitative trait loci in experimental populations. *Nature Reviews Genetics* 3: 43–52.

Dudley JW (1984) Theory for identification and use of exotic germplasm in maize breeding programs. *Maydica* 29: 291–407.

Dudley JW and Lambert RJ (1992) 90-Generations of selection for oil and protein in maize. *Maydica* 37: 81–87.

Ennos RA (1996) Utilising genetic information in plant conservation programmes. In: Hochberg ME, Clobert J and Barbault R (eds) *Aspects of the Genesis and Maintenance of Biological Diversity*. Oxford: Oxford University Press, pp. 278–291.

Eshed Y and Zamir D (1994) A genomic library of *Lycopersicon pennellii* in *L. esculentum*: a tool for fine mapping of genes. *Euphytica* 79: 175–179.

Eshed Y and Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* 141: 1147–1162.

Evenson RE (1999) Intellectual property rights, access to plant germplasm, and crop production scenarios in 2020. *Crop Science* 39: 1630–1635.

Falconer DS and Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th edn. New York: Longman Scientific and Technical.

Falush D, Stephens M and Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.

FAO (1996a) *Report on the State of the World's Plant Genetic Resources for Food and Agriculture*, prepared for the International Technical Conference on Plant Genetic Resources, Leipzig, Germany, 17–23 June 1996. Rome: Food and Agricultural Organization of the United Nations.

FAO (1996b) *Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources and the Leipzig Declaration*, adopted by the International Technical Conference on Plant Genetic Resources, Leipzig, Germany, 17–23 June 1996. Rome: Food and Agricultural Organization of the United Nations.

FAO (1989) *Annex to the International Undertaking on Plant Genetic Resources*, Conference Resolution 5/89. Twenty-Fifth Session of the FAO Conference, Rome, 11–29 November 1989. Rome: Food and Agricultural Organization of the United Nations.

Flint-Garcia SA, Thornsberry JM and Buckler ES IV (2003) Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology* 54: 357–374.

Francois IEJA, Broekaert WF and Cammue BPA (2002) Different approaches for multi-transgenic-stacking in plants [Review]. *Plant Science* 163: 281–295.

Frankel OH (1984) Genetic perspectives of germplasm conservation. In: Arber W, Llimensee K, Peacock WJ and Starlinger P (eds) *Genetic Manipulation: Impact on Man and Society*. Cambridge: Cambridge University Press, Part III, Paper No. 15.

Frese L (2002) Combining static and dynamic management of PGR: a case study of *Beta* genetic resources. In: Engels JMM, Rao VR, Brown AHD and Jackson MT (eds) *Managing Plant Genetic Diversity*. Wallingford: IPGRI/CABI Publishing, pp. 133–147.

Frese L, Desprez B and Ziegler D (2001) Potential of genetic resources and breeding strategies for base-broadening in *Beta*. In: Cooper HD, Spillane C and Hodgkin T (eds) *Broadening the Genetic Base of Crop Production*. Wallingford: CABI Publishing in co-operation with FAO and IPGRI, CAB International, pp. 295–309.

Fridman E, Pleban T and Zamir D (2000) A recombinant hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proceedings of the National Academy of Sciences USA* 97: 4718–4723.

Frisch M and Melchinger AE (2001a) The length of the intact donor chromosome segment around a target gene in marker-assisted backcrossing. *Genetics* 157: 1343–1356.

Frisch M and Melchinger AE (2001b) Marker-assisted backcrossing for simultaneous introgression of two genes. *Crop Science* 41: 1716–1725.

Frisch M and Melchinger AE (2001c) Marker-assisted backcrossing for introgression of a recessive gene. *Crop Science* 41: 1485–1494.

Frisch M, Bohn M and Melchinger AE (1999a) Minimum sample size and optimal positioning of flanking markers in marker-assisted backcrossing for transfer of a target gene. *Crop Science* 39: 967–975.

Frisch M, Bohn M and Melchinger AE (1999b) Comparison of selection strategies for marker-assisted backcrossing of a gene. *Crop Science* 39: 1295–1301.

Frisch M, Bohn M and Melchinger AE (2000) PLABSIM: software for simulation of marker-assisted backcrossing. *Journal of Heredity* 91: 86–87.

Fulton TM, Beck Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D and Tanksley SD (1997) QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. *Theoretical and Applied Genetics* 95: 881–894.

Fulton TM, Grandillo S, Beck Bunn T, Fridman E, Frampton A, Lopez J, Petiard V, Uhlig J, Zamir D and Tanksley SD (2000) Advanced backcross QTL analysis of a *Lycopersicon esculentum* × *Lycopersicon parviflorum* cross. *Theoretical and Applied Genetics* 100: 1025–1042.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ and Altshuler D (2002) The structure of haplotype blocks in the human genome. *Science* 296: 2225–2229.

Galili G, Galili S, Lewinsohn E and Tadmor Y (2002) Genetic, molecular and genomic approaches to improve the value of plant food and feeds. *Critical Reviews in Plant Sciences* 21: 167–204.

Geiger HH and Welz HG (2000) Marker-assisted selection. II Quantitative traits. In: Haussmann BIG, Geiger HH, Hess DE, Hash CT and Bramel-Cox P (eds) *Application of Molecular Markers in Plant Breeding* (training manual). Patancheru: ICRISAT, http://www.icrisat.org/text/research/grep/homepage/mol/molecular.htm.

Geiger HH, Yuan Y, Miedaner T and Wilde P (1995) Environmental sensitivity of cytoplasmic genic male sterility (CMS) in *Secale cereale* L. In: Kück U and Wricke G (eds) *Genetic Mechanisms for Hybrid Breeding. Advances in Plant Breeding* 18: 7–17.

Goldringer I, Enjalbert J, David J, Paillard S, Pham JL and Brabant P (2001) Dynamic management of genetic resources: a 13-year experiment on wheat. In: Cooper HD, Spillane C and Hodgkin T (eds) *Broadening the Genetic Base of Crop Production*. Wallingford: CABI Publishing in co-operation with FAO and IPGRI, CAB International, pp. 245–260.

Goodman MM (1999) Broadening the genetic diversity in maize breeding by use of exotic germplasm. In: *The Genetics and Exploitation of Heterosis in Crops*. Madison: ASA-CSSA-SSSA, pp. 139–148.

Grattapaglia D, Ciampi AY, Gaiotto FA, Squilassi MG, Collevatti RG, Ribeiro VJ, Reis AM, Gandara FB, Walter BM, Brondani RPV, Bruns S, Mantell S, Tragardh C and Viana AM (1998) DNA technologies for forest tree breeding and conservation. In: Bruns S, Mantell S and Tragardh C (eds) *Recent Advances in Biotechnology for Tree Conservation and Management*. Proceedings of an IFS Workshop, Florianopolis, Brazil, 15–19 September 1997. Stockholm: International Foundation for Science (IFS), pp. 50–61.

Hackett CA (2002) Statistical methods for QTL mapping in cereals. *Plant Molecular Biology* 48: 585–599.

Hallauer AR and Miranda JB (1981) *Quantitative Genetics in Maize Breeding*. Ames: Iowa State University Press.

Hansen M, Kraft T, Ganestam S, Säll T and Nilsson NO (2001) Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. *Genetical Research* 77: 61–66.

Harlan JR and de Wet JMJ (1971) Toward a rational classification of cultivated plants. *Taxon* 20: 509–517.

Hawkes JG, Maxted N and Ford-Lloyd BV (2000) *The Ex Situ Conservation of Plant Genetic Resources*. Dordrecht: Kluwer Academic Publishers.

Hawtin G, Iwanaga M and Hodgin T (1997) Genetic resources in breeding for adaptation. In: Tigerstedt PMA (ed.)

Adaptation in Plant Breeding. Amsterdam: Kluwer Academic Publishers, pp. 277–288.

Heckenberger M, Bohn M, Ziegle JS, Joe LK, Hauser JD, Hutton M and Melchinger AE (2002) Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties. I. Genetic and technical sources of variation in SSR data. *Molecular Breeding* 10: 181–191.

Heckenberger M, van der Voort JR, Melchinger AE, Peleman J and Bohn M (2003) Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties: II. Genetic and technical sources of variation in AFLP data and comparison with SSR data. *Molecular Breeding* 12: 97–106.

Henry RJ (ed.) (2001) *Plant Genotyping*. New York: CABI Publishing.

Hodgkin T, Brown AHD, van Hintum TJL and Morales EAV (1995) *Core Collections of Plant Genetic Resources*. Chichester: John Wiley & Sons.

Howell PM, Marshall DF and Lydiate DJ (1996) Towards developing intervarietal substitution lines in *Brassica napus* using marker-assisted selection. *Genome* 39: 348–358.

Hu Q, Hansen LN, Laursen J, Dixelius C and Andersen SB (2002) Intergeneric hybrids between *Brassica napus* and *Orychophragmus violaceus* containing traits of agronomic importance for oilseed rape breeding. *Theoretical and Applied Genetics* 105: 834–840.

Jannink JL and Walsh B (2002) Association mapping in plant populations. In: Kang S (ed.) *Quantitative Genetics, Genomics and Plant Breeding*. Wallingford: CAB International, pp. 59–68.

Jannink J, Bink MC and Jansen RC (2001) Using complex plant pedigrees to map valuable genes. *Trends in Plant Science* 6: 337–342.

Jordaan JP (1998) Hybrid wheat in Africa? Proceedings of the Tenth Regional Wheat Workshop for Eastern, Central and Southern Africa, University of Stellenbosch, September, pp. 465–472.

Jordan D, Tao Ygodwin I, Henzell R, Cooper M and McIntyre C (2003) Prediction of hybrid performance in grain sorghum using RFLP markers. *Theoretical and Applied Genetics* 106: 559–567.

Kannenberg LW (2001) HOPE, a hierarchical, open-ended system for broadening the breeding base of maize. In: Cooper HD, Spillane C and Hodgkin T (eds) *Broadening the Genetic Base of Crop Production*. WAllingford: CABI Publishing in co-operation with FAO and IPGRI, CAB International, pp. 311–318.

Kannenberg LW and Falk DE (1995) Models for activation of plant genetic resources for crop breeding programs. *Canadian Journal of Plant Science* 75: 45–53.

Karp A, Edwards KJ, Bruford M, Funk S, Vosman B, Morgante M, Seberg O, Kremer A, Boursot P, Arctander P, Tautz D and Hewitt GM (1997) Molecular technologies for biodiversity evaluation: opportunities and challenges. *Nature Biotechnology* 15: 625–628.

Karp A, Isaak PG and Ingram DS (eds) (1998) *Molecular Tools for Screening Biodiversity*. London: Chapman and Hall.

Kearsey MJ (2002) QTL analysis: problems and (possible) solutions. In: Kang MS (ed.) *Quantitative Genetics, Genomics and Plant Breeding*. Wallingford: CAB International, pp. 45–58.

Knapp SJ (2001) Mapping quantitative trait loci. In: Phillips RL and Vasil IK (eds) *DNA Markers in Plants*. Amsterdam: Kluwer Academic Publishers, pp. 59–99.

Kresovitch S, Luongo AJ and Schloss SJ (2002) 'Mining the Gold': finding allelic variants for improved crop conservation and use. In: Engels JMM, Rao VR, Brown AHD and Jackson MT (eds) *Managing Plant Genetic Diversity*. Wallingford: IPGRI/CABI Publishing, pp. 379–386.

Lessard PA, Kulaveerasingam H, York GM, Strong A and Sinskey AJ (2002) Manipulating gene expression for the metabolic engineering of plants [Review]. *Metabolic Engineering* 4(1): 67–79.

Lewis CM (2002) Genetic association studies: design, analysis and interpretation. *Briefing in Bioinformatics* 3: 146–153.

Locke C (2001) Has bioprospecting gone wild? *Red Herring Magazine* 95, 15 April, http://www.redherring.com/mag/issue95/250018225.html.

Lohmueller KE, Pearce CL, Pike M, Lander ES and Hirschhorn JN (2003) Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genetics* 33: 177–182.

Maquet A, Bi IZ, Delvaux M, Wathelet B and Baudoin Jp (1997) Genetic structure of a Lima bean base collection using allozyme markers. *Theoretical and Applied Genetics* 95: 980–991.

Maxted N, Ford-Lloyd BV and Hawkes JG (eds) (1997) *Plant Genetic Conservation—The In Situ Approach*. London: Chapman & Hall.

Maxted N, Guarino L, Myer L and Chiwona EA (2002) Towards a methodology for on-farm conservation of plant genetic resources. *Genetic Resources and Crop Evolution* 49: 31–46.

Melchinger AE (1990) Use of molecular markers in breeding for oligogenic disease resistance. *Plant Breeding* 104: 1–19.

Melchinger AE, Utz HF and Schoen CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* 149(1): 383–403.

Melchinger AE, Coors JG and Pandey S (1999) Genetic diversity and heterosis. In: Coors JG (ed.) *The Genetics and Exploitation of Heterosis in Crops*. Proceedings of an International Symposium, CIMMYT, Mexico City, Mexico, 17–22 August 1997. Mexico City: CIMMYT, pp. 99–118.

Miedaner T, Glass C, Dreyer F, Wilde P, Wortmann H and Geiger HH (2000) Mapping of genes for male-fertility restoration in 'Pampa' CMS winter rye (*Secale cereale* L.). *Theoretical and Applied Genetics* 101: 1226–1233.

Mohammadi SA and Prasanna BM (2003) Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop Science* 43: 1235–1248.

Moncada PP, Martínez CP, Borrero J, Chatel M, Gauch H Jr, Guimaraes E, Tohme J and McCouch SR (2001) Quantitative trait loci for yield and yield components in an *Oryza sativa* × *Oryza rufipogon* BC_2F_2 population evaluated in an upland environment. *Theoretical and Applied Genetics* 102: 41–52.

Monforte AJ and Tanksley SD (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: a tool for gene mapping and gene discovery. *Genome* 43: 803–813.

Müller KJ (1989) Die Entwicklung von Basispopulationen bei Getreide für biologisch-dynamische Betriebe. *Zeitschrift Lebendige Erde* 5(89): 364–367.

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press

Nevo E (1998) Genetic diversity in wild cereals: regional and local studies and their bearing on conservation *ex situ* and *in situ*. *Genetic Resources and Crop Evolution* 45: 355–370.

Newbury HJ (2003) *Plant Molecular Breeding*. Oxford: Blackwell Publishing/CRC Press.

Niangado O (2001) The state of millet diversity and its use in West Africa. In: Cooper HD, Spillane C and Hodgkin T (eds) *Broadening the Genetic Base of Crop Production*. Wallingford: CABI Publishing in co-operation with FAO and IPGRI, CAB International, pp. 147–157.

Nordborg M (2000) Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. *Genetics* 154: 923–929.

Nordborg M and Tavaré S (2002) Linkage disequilibrium: what history has to tell us. *Trends in Genetics* 18: 83–90.

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad JJ, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA and Weigel D (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics* 30: 190–193.

Österberg MK, Shavorskaya O, Lascoux M and Lagercrantz U (2002) Naturally occurring indel variation in the *Brassica nigra* COL1 gene is associated with variation in flowering time. *Genetics* 161: 299–306.

Padulosi S, Hodgkin T, Williams JT and Haq N (2002) Underutilized crops: trends, challenges and opportunities in the 21st century. In: Engels JMM, Rao VR, Brown AHD and Jackson MT (eds) *Managing Plant Genetic Diversity*. Wallingford: IPGRI/CABI Publishing, pp. 323–338.

Pillen K, Zacharias A and Léon J (2003) Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 107: 340–352.

Pritchard JK and Przeworski M (2001) Linkage disequilibrium in humans: models and data. *American Journal of Human Genetics* 69: 1–14.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Ragot M, Sisco PH, Hoisington DA and Stuber CW (1995) Molecular-marker-mediated characterization of favorable exotic alleles at quantitative trait loci in maize. *Crop Science* 35: 1306–1315.

Ramsay LD, Jennings DE, Bohuon EJR, Arthur AE, Lydiate DJ, Kearsey MJ and Marshal DF (1996) The construction of a substitution library of recombinant backcross lines in *Brassica oleracea* for the precision mapping of quantitative trait loci. *Genome* 39: 558–567.

Reif JC, Melchinger AE, Xia XC, Warburton ML, Hoisington DA, Vasal SK, Beck D, Bohn M and Frisch M (2003a) Use of SSRs for establishing heterotic groups in subtropical maize. *Theoretical and Applied Genetics* 107: 947–957.

Reif JC, Melchinger AE, Xia XC, Warburton ML, Hoisington DA, Vasal SK, Srinivasan G, Bohn M and Frisch M (2003b) Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. *Crop Science* 43: 1275–1282.

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM and Buckler ES IV (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences USA* 98: 11479–11484.

Repellin A, Båga M, Jauhar PP and Chibbar RN (2001) Genetic enrichment of cereal crops *via* alien gene transfer: new challenges. *Plant Cell, Tissue and Organ Culture* 64: 159–183.

Schnell FW (1978) Progress and problems in utilizing quantitative variability in plant breeding. *Plant Breeding Research and Development* 7: 32–43.

Schnell FW (1983) Probleme der Elternwahl—Ein Überblick. In: *Bericht über die Arbeitstagung der Vereinigung österreichischer Pflanzenzüchter*, Gumpenstein, 22–24 November 1983. Gumpenstein, Austria: BAL, pp. 1–11.

Schnell FW and Utz HF (1975) F_1 Leistung und Elternwahl in der Züchtung von Selbstbefruchtern. In: *Bericht über die Arbeitstagung der Vereinigung österreichischer Pflanzenzüchter*, Gumpenstein, 25–27 November 1973. Gumpenstein, Austria: BAL, pp. 243–248.

Šimić D, Presterl T, Seitz G and Geiger HH (2003) Comparing methods for integrating exotic germplasm into European forage maize breeding programs. *Crop Science* 43: 1952–1959.

Simmonds NW (1991) Selection for local adaptation in a plant breeding programme. *Theoretical and Applied Genetics* 82: 363–367.

Simmonds NW (1993) Introgression and incorporation. Strategies for the use of crop genetic resources. *Biological Reviews* 68: 539–562.

Sorrells ME and Wilson WA (1997) Direct classification and selection of superior alleles for crop improvement. *Crop Science* 37: 691–697.

Spielman RS and Ewens WJ (1998) A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *American Journal of Human Genetics* 62: 450–458.

Spielman RS, McGinnis RE and Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus. *American Journal of Human Genetics* 52: 506–516.

Stam P and Zeven AC (1981) The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. *Euphytica* 30: 227–238.

Stracke S, Schilling AG, Förster J, Weiss C, Glass C, Miedaner T and Geiger HH (2003) Development of PCR-based markers linked to dominant genes for male-fertility restoration in Pampa CMS of rye (*Secale cereale* L.). *Theoretical and Applied Genetics* 106: 1184–1190.

Swaminathan MS (2002) The past, present and future contributions of farmers to the conservation and development of genetic diversity. In: Engels JMM, Rao VR, Brown AHD and Jackson MT (eds) *Managing Plant Genetic Diversity*. Wallingford: IPGRI/CABI Publishing, pp. 23–31.

Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J and Beck-Bunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theoretical and Applied Genetics* 92: 213–224.

Tanksley SD and McCouch S (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277: 1063–1066.

Tanksley SD and Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theoretical and Applied Genetics* 92: 191–203.

ten Kate K and Laird SA (2000) *The Commercial Use of Biodiversity—Access to Genetic Resources and Benefit Sharing*, 2nd edn. London: Earthscan Publications.

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D and Buckler ES IV (2001) *Dwarf8* polymorphisms associate with variation in flowering time. *Nature Genetics* 28: 286–289.

Tian D, Araki H, Stahl E, Bergelson J and Kreitman M (2002) Signature of balancing selection in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA* 99: 11525–11530.

Trognitz BR, Bonierbale M, Landeo JA, Forbes G, Bradshaw JE, Mackay GR, Waugh R, Huarte MA and Colon L (2001) Improving potato resistance to disease under the Global Initiative on Late Blight. In: Cooper HD, Spillane C and Hodgkin T (eds) *Broadening the Genetic Base of Crop Production*. Wallingford: CABI Publishing in co-operation with FAO and IPGRI, CAB International, pp. 385–398.

Troyer AF and Rocheford TR (2002) Germplasm ownership: related corn inbreds. *Crop Science* 42: 3–11.

Utz HF, Melchinger AE and Schön CC (2000) Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. *Genetics* 154: 1839–1849.

Valkoun JJ (2001) Wheat pre-breeding using wild progenitors. *Euphytica* 119: 17–23.

van Hintum TJL, Brown AHD, Spillane C and Hodgkin T (2000) Core collections of plant genetic resources. *IPGRI Technical Bulletin* 3: 5–48.

Van Treuen (2000) *Molecular Markers*. Wageningen, The Netherlands: Centre for Genetic Resources (CGN), http://www.cgn.wageningen-ur.nl/pgr/research/molgen/.

vom Brocke K, Christinck A, Presterl T, Weltzien-Rattunde E and Geiger HH (2002a) Effects of farmers' seed management on performance and adaptation of pearl millet in Rajasthan, India. *Euphytica* 130: 267–280.

vom Brocke K, Presterl T, Christinck A, Weltzien-Rattunde E and Geiger HH (2002b) Farmers' seed management practices open up new base populations for pearl millet breeding in a semi-arid zone of India. *Plant Breeding* 121: 36–42.

Vuylsteke, MJR (1999) Genetic analysis of maize by using the AFLP® method. PhD thesis, University of Wageningen.

Vuylsteke M, Kuiper M and Stam P (2000) Chromosomal regions involved in hybrid performance and heterosis: their AFLP-based identification and practical use in prediction models. *Heredity* 85: 208–218.

Wall JD and Pritchard JK (2003) Haplotype blocks and linkage disequilibrium in the human genome. *Nature Reviews Genetics* 4: 587–597.

Welz HG and Geiger HH (2000) Marker-assisted selection. I Qualitative traits. In: Haussmann BIG, Geiger HH, Hess DE, Hash CT and Bramel-Cox P (eds) *Application of Molecular Markers in Plant Breeding (training manual)*, http://www.icrisat.org/text/research/grep/homepage/mol/molecular.htm. Patancheru: ICRISAT.

Whitt SR, Wilson LM, Tenaillon MI, Gaut BS and Buckler ES (2002) Genetic diversity and selection in the maize starch pathway. *Proceedings of the National Academy of Sciences USA* 99: 12959–12962.

Witcombe JR (2001) The impact of decentralized and participatory plant breeding on the genetic base of crops. In: Cooper HD and Hodgkin T (eds) *Broadening the Genetic Base of Crop Production*. Wallingford: CABI Publishing in co-operation with FAO and IPGRI, CAB International, pp. 407–417.

Wolfe AD and Liston A (1999) Contributions of PCR-based methods to plant systematics and evolutionary biology. In: Soltis DE, Soltis PS and Doyle JJ (eds) *Molecular Systematic of Plants II: DNA Sequencing*. Boston: Kluwer Academic Publishers, pp. 43–86.

Xiao J, Grandillo S, Ahn SN, McCouch SR, Tanksley SD, Li J and Yuan L (1996) Genes from wild rice improve yield. *Nature* 384: 223–224.

Xiao J, Li J, Grandillo S, Ahn S, Yuan L, Tanksley SD and McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* 150: 899–909.

Zamir D (2001) Improving plant breeding with exotic genetic libraries. *Nature Reviews Genetics* 2: 983–989.

Zhang QF, Gao YJ, Maroof MAS, Yang SH, Li JX and Zhang QF (1995) Molecular divergence and hybrid performance in rice. *Molecular Breeding* 1: 133–142.

https://doi.org/10.1079/PGR200430 Published online by Cambridge University Press