Aus dem Institut für
Pflanzenzüchtung, Saatgutforschung und Populationsgenetik
der Universität Hohenheim
Fachgebiet: Angewandte Genetik und Pflanzenzüchtung
Prof. Dr. A.E. Melchinger

GENETIC DIVERSITY IN ELITE LINES AND LANDRACES
OF CIMMYT SPRING BREAD WHEAT
AND HYBRID PERFORMANCE OF CROSSES AMONG
ELITE GERMPLASM

Dissertation
zur Erlangung des Grades eines
Doktors der Agrarwissenschaften
vorgelegt der
Fakultät für Agrarwissenschaften
der Universität Hohenheim

Von Dipl. Agr. Biol.
Susanne Dreisigacker
aus Birkweiler

Stuttgart-Hohenheim
2004

Die vorliegende Arbeit wurde am 16.11.2004 von der Fakultät III – Agrarwissenschaften I, der Universität Hohenheim als „Dissertation zur Erlangung der Grades eines Doktors" der Agrarwissenschaften angenommen.

Tag der mündlichen Prüfung:

Dekan:		Prof. Dr. K. Stahr
Berichterstatter,	1. Prüfer:	Prof. Dr. A.E. Melchinger
Mitberichterstatter,	2. Prüfer:	Prof. Dr. R. Blaich
	3. Prüfer:	Prof. Dr. H-P. Piepho

TABLE OF CONTENTS

[1] Dreisigacker S., P. Zhang, M.L. Warburton, M. Bohn, M. Van-Ginkel, D.Hoisington, and A.E. Melchinger. 2004. SSR and pedigree analysis of genetic diversity in CIMMYT wheat lines targeted to different megaenvironments. Crop Sci. 44:381-388.

[2] Zhang, P., S. Dreisigacker, A.E. Melchinger, J.C. Reif, M. van Ginkel, A. Mujeeb-Kazi, D. Hoisington, M.L. Warburton. 2005. Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheats and their backcross-derived lines using SSR markers. Molecular Breeding 15:1-10.

[3] Dreisigacker S., P. Zhang, M.L. Warburton, B. Skovmand, D.Hoisington, and A.E. Melchinger. 2005. Genetic diversity among and within CIMMYT hexaploid wheat landrace accessions and implications for plant genetic resource management. Crop Sci. 45:653-661.

[4] Dreisigacker S., A.E. Melchinger, P. Zhang, K. Ammar, C. Flachenecker, D. Hoisington, and M.L. Warburton. 2005. Relationship between genetic diversity estimates and heterosis in CIMMYT wheat germplasm. Euphytica (in press).

LIST OF ABBREVIATIONS

AMOVA	analysis of molecular variance
AFLP	amplified fragment length polymorphism
CIMMYT	International Maize and Wheat Improvement Center
COP	coefficient of parentage
EST	expressed sequence tag
GCA	general combining ability
GS	genetic similarity
LC	landrace cultivar
LP	line *per se* performance
MB	multiple bands
ME	megaenvironment
MRD	modified Rogers' distance
NARS	national agricultural research stations
PCoA	principal coordinate analysis
PC	principal coordinate
PIC	polymorphism information content
RAPD	random amplified polymorphic DNA
RD	Rodgers distance
RFLP	restricted fragment length polymorphism
SBL	synthetic backcross-derived lines
SCA	specific combining ability
SHW	synthetic hexaploid wheat
SNP	single nucleotide polymorphism
SSR	simple sequence repeat

1 GENERAL INTRODUCTION

Wheat - A major crop

Wheat is one of the most important and widely-cultivated crops in the world. The global consumption of wheat, which is close third after rice and maize, continuously increased during the past decades (Fig. 1) and is expected to increase to 600 million tons in 2003/2004 (Foreign Agricultural Service, 2002). Wheat is used mainly for human consumption and supports nearly 35% of the world population. It is nutritious, easy to store and transport and can be processed into various types of food. The demand for wheat is expected to grow faster than any other major agricultural crop. To meet the needs of the growing world population, the forecast demand for the year 2020 varies between 840 (Rosegrant et al., 1995) and 1050 million tons (Kronstad, 1998). Due to land limitations, the enhancement of wheat production must come from higher absolute yields, which can only be met by the concerted action of scientists involved in diverse agricultural disciplines and in particular by increased efforts in plant breeding (Braun et al., 1998). In addition to continuous investments in conventional breeding methods, biotechnological tools and the better understanding of the current and expanded genetic diversity should be considered for raising the yield frontier in wheat.
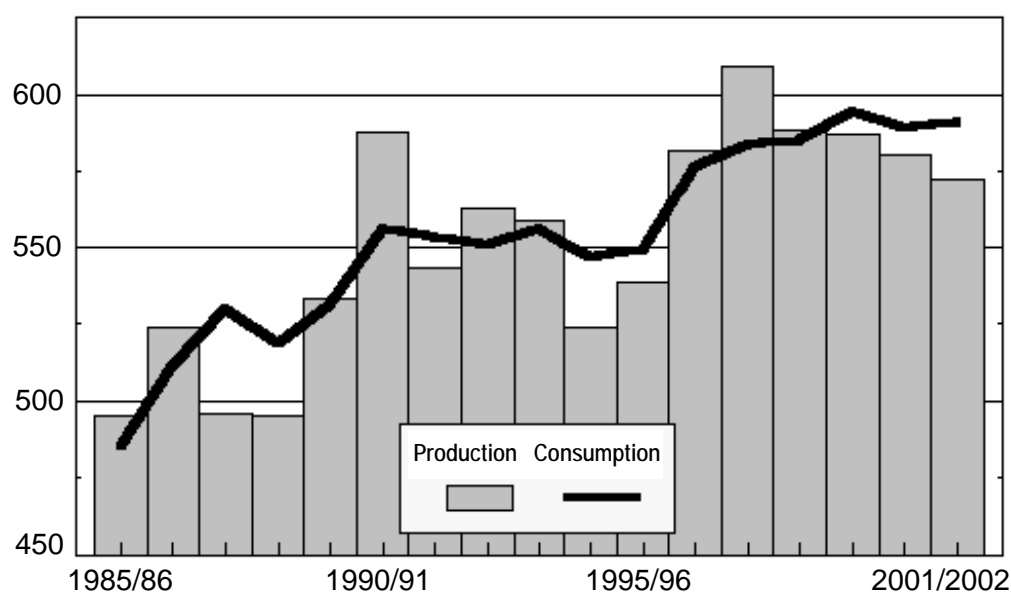

Figure 1. Development of global wheat production and consumption across the past two decades. Source: www.fas.usda.gov.

The origin and genetic characteristics of cultivated bread wheat

Wheat belongs to the genus *Triticum* and originated about 10 000 years ago in the 'Fertile Crescent', one of the most diversified regions in the world, comprising a wide array of different habitats (Morris and Sears, 1967). Of current commercial importance in agriculture are durum and bread wheat. They are products of natural hybridization of perennial wild types, none of which is cultivated on a large scale today. Wild emmer (*Triticum turgidum* L. *ssp. dicoccides,* 2n = 28) has been identified as the donor of the A and B genomes of durum and bread wheats (Fig. 2, Kimber and Feldman, 1987). Tetraploid wheat later outcrossed with goat grass (*T. tauschii,* 2n = 14) resulting in bread wheat (*T. aestivum* L. em Thel, 2n = 42) with the additional D genome (Kihara, 1944). The origin of wild emmer is still a matter of controversy, but there is a general conclusion that its A genome comes from the wild diploid wheat, Einkorn (*T. monococcum* L., 2n = 14, Boissier, 1884) and its B genome is related to the genome coming from a species of *Aegilops*. The wild types of wheat can still be seen in some regions of Iran, Iraq, and Turkey.

Each of the three different bread wheat genomes contributes seven chromosomes. Sears (1954) indicated that they show a similar physical characterization across the genomes, also defined as homoeologous groups. Genes from homoeologous groups can compensate each other, which makes wheat highly tolerant to genetic changes e.g., mutations or losses of individual chromosomal segments (Kimber and Feldman, 1987). Furthermore, this means that wheat can contain considerable genetic variation. The homoeologous groups allow breeders to accumulate favourable alleles (up to 6 per locus) for the enhancement of desired traits. Recombination between homoeologous chromosomes is suppressed, leading to a pairing pattern similar to that of diploid crops (Sears, 1976).

Because of its allopolyploid nature, the genomes of bread wheat show a high homology with those of several diploid and tetraploid wild species (Kimber, 1993). Consequently, genes from wild wheat species can be introgressed into cultivated wheat through recombination of the homoeologous chromosomes, and undesirable gene linkages can often be broken using repeated backcrossing to cultivated wheat (Valkoun, 2001). Moreover, chromosome recombination enables a simultaneous gene transfer from different

chromosomes, as well as introgression of polygenic traits, in which the genes are dispersed on different chromosome segments.
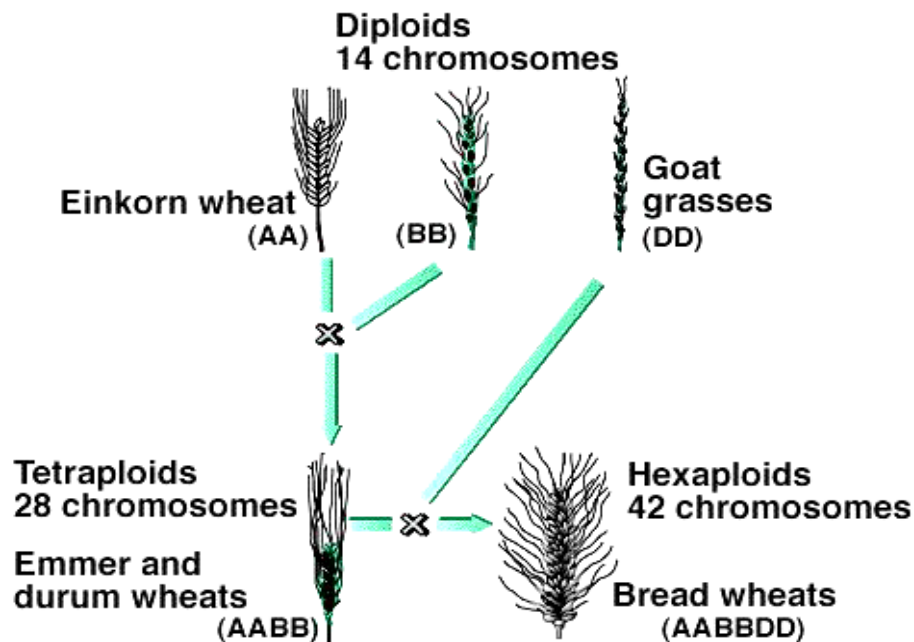

Figure 2. The origin of cultivated wheat types. Source: Hancock (1994).

The impact of CIMMYT in wheat production

The evolution of a modern system of improvement in wheat is often linked to the "Green Revolution", which had its origin in the transfer of Mexican semi-dwarf varieties developed by the Rockefeller Foundation research program to India and Pakistan (Hanson et al., 1982). The cultivation of semi-dwarf varieties with higher yields due to an increased harvest index and better lodging as well as disease resistance, doubled and even tripled the production of food grains in these countries in a very short period of time. The initial seed transfer was followed by the establishment of the International Maize and Wheat Improvement Center (CIMMYT) in Mexico in 1966. Over the last 50 years, CIMMYT originated wheat germplasm has continuously enhanced yield potential and has been distributed to a large number of developing countries (Rajaram, 1995). Breeding efforts basically addressed resistance to major diseases, industrial grain quality, and wheat adaptation to different agro-climatological regions. The exchange of CIMMYT wheat germplasm with National Agricultural Research Stations (NARS) in many developing countries has led to the development of an international wheat improvement system.

The substantial impact of CIMMYT in the developing world became apparent in a recent study of Heisey et al. (2002), which considered the distribution of CIMMYT bread wheats as well as the number of cultivars released per million hectares and year. Slightly more than 62% of the area grown to bread wheat in developing countries was sown to cultivars with CIMMYT ancestry. Restricting the area to spring bread wheat, which is the dominant type of wheat in developing countries, the percentage of area sown to CIMMYT related lines amounted to 88%. Between 1991 and 1997, 84% of the spring bread wheat cultivars released by NARS programs were derived from CIMMYT crosses or included at least one CIMMYT parent in their pedigree.

Owing to their long-term nature and relatively high budgets for long-term projects, international agricultural research centers like CIMMYT hold the responsibility for *ex situ* conservation for 22 major crops (Cohen et al., 1991). The preservation of wheat genetic resources at CIMMYT started in 1968 (Hoisington et al., 1999). Since 1994, the collection has been designated "in trust to humanity" under the auspices of FAO, and is today the largest collection worldwide including about 168 000 *Triticeae* accessions. The collection includes hundreds of wild species, landraces, locally grown and modern cultivars as well as breeding lines. Researchers often term CIMMYT's collection as a "gold mine" of wheat germplasm.

Dimensions of genetic diversity

Nearly 30 years ago, the scientific community raised an alarm about "genetic erosion" (National Research Council, 1972). Harlan (1972) first used this term to describe the potentially disastrous narrowing of the germplasm base caused by improved food crops. The replacement of traditional farmer's varieties with more uniform modern varieties entails the risk of genetic vulnerability, because mutations in disease or insect populations or changes of environmental conditions can result in a drastic crop loss. This risk was brought sharply into focus in 1970 with the outbreak of the Southern corn leaf blight (National Research Council, 1972). The disease drastically reduced corn yields in the United States and was caused by the extensive use of a single sterile cytoplasm, which was associated with disease susceptibility. First signs that germplasm with a narrow genetic base might lead to disasters in wheat came from several severe epidemics of shoot fly (*Atherigona spp.*) and karnal bunt (*Tilletia indica*) in the 1970s in India (Dalrymple, 1986).

Taking these epidemiological and economic aspects into account, there was a general agreement that the wealth of genetic variation provided by nature have to be harnessed for future genetic improvement of crops.

Several types of diversity can be measured in the context of breeding programs (Smale et al., 1996). Spatial and temporal diversity reflects the distribution of crops regarding to the area planted and the average livetime and turnover of cultivars, respectively. Apparent and latent genetic diversity are directly related to the performance of crops. Measures of apparent diversity are manifested in phenotypic differences of populations or cultivars in the field (Souza et al., 1994). Latent diversity refers to parentage analysis and molecular measurements that are not necessarily expressed in crop performance or phenotypes. Parentage analysis can be applied when extensive pedigree records are available. Coefficients of parentage (COP) summarize the genealogical information from an array of cultivars. Originally developed by Wright (1922) and Malecot (1948), the COP for two cultivars estimates the expected percentage of loci in common by descent. St. Martin (1982) adapted the COP analysis to autogamous crops by assuming that (i) each cultivar is completely homozygous, (ii) cultivars without common parentage are unrelated, and (iii) parents contribute equally to the offspring despite inbreeding and selection.

Genetic diversity based on molecular markers has been studied in plants for about three decades. The most comprehensive early studies were performed with isoenzymes (Hamrick and Godt, 1990), which provided many insights into population structure and breeding systems. However, although these markers allowed large numbers of samples to be analyzed, only a limited number of loci could be scored and the comparison of samples from different species and laboratories were problematic. During the past decade the focus shifted to surveys at the level of DNA. Using DNA markers, diversity is measured as the average allelic divergence between any two individuals for given loci. Sufficiently large numbers of samples allow robust analyses of open questions in population genetics.

The application of SSR markers in wheat

The development of DNA markers in wheat is somewhat problematic due to three features. First, the size of the wheat genome (16×10^9 bp, compared to barley or maize with 5×10^9 bp), which makes the application of several marker techniques difficult. Second, the hexaploid nature of wheat adds complexity to many marker assays. Three sets

of bands usually appear (often in the same size range), which are difficult to manage and interpret. Third, there is a generally low level of polymorphism in wheat relative to other cereal crops. This implies that a larger number of markers must be screened than in the case of rice, barley or maize (Chao et al., 1989; Lui et al., 1990). Furthermore, the level of polymorphism is not consistent across genomes and crosses. Commonly, the D genome tends to have the poorest marker coverage (Chalmers et al. 2001). Most work on wheat to date has used random amplified polymorphic DNA (RAPD), restriction fragment length polymorphisms (RFLPs), amplified fragment length polymorphisms (AFLP), and simple sequence repeats (SSRs).

Bohn et al. (1999) and Parker et al. (2002) compared these marker systems across 11 and 124 wheat varieties, respectively. Their results suggest that AFLP and SSR markers are the most effective in detecting polymorphisms. However, linkage maps for the dominant AFLP marker system are limited. The specific genome being assayed by the AFLP markers are unknown and translocations from other species (which are abundant in CIMMYT wheats) are overrepresented with AFLPs (Warburton et al., 2002). Furthermore, AFLPs are costly and complex to use. SSRs are therefore the currently most popular marker system in wheat.

The objectives of this study were to:

(i) investigate the genetic diversity in current CIMMYT elite materials targeted to different megaenvironments based on SSR markers and pedigree analysis.

(ii) evaluate the use of CIMMYT synthetic hexaploid wheats and their backcross-derived lines as a source to enhance the genetic variation in breeding programs,

(iii) determine the genetic diversity among and within wheat landrace accessions and evaluate the use of SSRs as a tool for genetic resources management, and

(iv) assess the prospects of hybrid wheat as one strategy to enhance the yield potential of wheat.

References

Boissier, P.E. 1884. Flora orientalis, v. 5, pp. 673-679, Geneva, Basel and Lyon.

Bohn, M., F.H. Utz, and A.E. Melchinger. 1999. Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. Crop Sci. 39:228-237.

Braun, H.-J., T.S. Payne, A.I. Morgounov, M. van Ginkel, and S. Rajaram. 1998. The challenge: One billion tons of wheat by 2020. pp. 33-40. *In* A.E. Slinkard (ed.) Proc. Int. Wheat Genet. Symp., 9th, Saskatoon, Canada. 2-7 Aug. 1998. Extension Division, University of Saskatchewan, SK.

Chalmers, K.J., A.W. Campbell, J. Kretschmer, A. Karakousis, P. Henschke, S. Pierens, N. Harker, M. Palotta, G.B. Cornish, M.R. Shariflou, L.R. Rampling, A. McLauchlan, G. Daggard, P.J. Sharp, T.A. Holton, M.W. Southerland, R. Appels, and P. Langridge. 2001. Construction of three linkage maps in bread wheat, *Triticum aestivum*. Aust. J. Agric. Res. 52:1089-1119.

Chao, S., P.J. Sharp, A.J. Worland, E.J. Warham, R.M.D. Koebner, and M.D. Gale. 1989. RFLP-based genetic linkage maps of wheat homoeologous group 7 chromosomes. Theor. Appl. Genet. 78:495-504.

Cohen, J.I., J.T. Willams, D.L. Plucknett, and H. Shands. 1991. *Ex situ* conservation of plant genetic resources: Global development and environmental concerns. Science 253:866-872.

Dalrymple, D.G. 1986. Development and spread of high-yielding wheat varieties in developing countries, 7th ed. US Agency for Int. Develop., WA.

Foreign Agricultural Service. 2002. Foreign Agricultural Service Circular Series, FG 11-02, November 2002, FAS online: http://www.fas.usda.gov/grain/circular/2002/11-02/graintoc.htm.

Hamrick, J.L., and M.J.W. Godt. 1990. Alloenzyme diversity in plant species.pp. 43-63. *In* Brown, A.H.D., M.T. Clegg, A.L. Kahler, B.S. Weir (eds.) Plant population genetics, breeding, and genetic resources. Sinauer Associates Inc., Sunderland, MA.

Hanson, H., N.E. Borlaug, and R.G. Anderson. 1982. Wheat in the third world. Westview Press, Inc., Boulder, CO.

Harlan, J.R. 1972. Genetics of disaster. J. Environ. Qual. 1:212-215.

Harlan , J.R. 1987. The early history of wheat: Earliest traces to the sack of Rome. *In* L.T Evans and W. Peacock (eds.) Wheat science: Today and tomorrow. Cambridge University Press, Cambridge, U.K.

Heisey, P.W. 2002 Impacts of international wheat breeding research in developing countries 1966-1997. CIMMYT, Mexcio D.F, Mexico.

Hoisington, D., M. Kahairallah, T. Reeves, J.-M. Ribaut, B. Skovmand, S. Taba, and M. Warburton. 1999. Plant genetic resources: What can they contribute toward increased crop productivity. Proc. Natl. Acad. Sci. 96:5937-5943.

Kihara, H., 1944. Die Entdeckung der DD-Analysatoren beim Weizen. Agric. Hort. 19:889-890 (in German).

Kimber, G. 1993. Genomic relationships in *Triticum* and the availability of alien germplasm. pp. 9-16. *In* Damania, A.B. (ed.) Biodiversity and wheat improvement, John Wiley & Sons, Chichester, UK.

Kimber G., and M. Feldman, 1987. Wild wheat. An introduction. pp. 129-131. Special Report 353, College of Agriculture, University of Missouri, Colombia.

Kronstad, W.E. 1998. Agricultural development and wheat breeding in the 20th century. Pp. 1-10. *In* Braun, H.-J., F. Altay, W.E. Kronstad, S.P.S. Benival, and A. McNAb (eds.) Wheat: Prospects for global improvement. Proc. of the 5th Int. Wheat Conf., Ankara, Turkey, Developments in Plant Breeding v. 6. Kluwer Academic Publishers, Dordrecht.

Langridge, P.E.S. Lagudah, T.A. Holton, R. Appels, P.J. Sharp, and K.J. Chalmers. Trends in genetic and genome analysis in wheat: a review. Aust. J. Agric. Res. 52:1043-1077.

Lui, Y.G., N. Mori, K. Tsunewaki. 1990. Restriction fragment polymorphism (RFLP) analysis in wheat. I. Genomic DNA library construction and RFLP analysis in common wheat. Jap. J. Genet. 65:367-380.

Malecot, G. 1948. Les mathematiques de l'heredite. Masson et Cie, Paris.

Morris, R. and E.R. Sears. 1967. The cytogenetics of wheat and its relatives. *In* Quinsberry, K.S. and L.P. Reiz (eds.). Wheat and wheat improvement. Agronomy Monograph 13. Am. Soc. of Agron. Inc., Madison, WI.

National Research Council. 1972. Genetic vulnerability of major crops. National Academy of Science, WA.

Parker, G.D., P. N. Fox, P. Langridge, K. Chalmers, B. Whan, and P. Ganter. 2002. Genetic diversity within Australian wheat breeding programs based on molecular and pedigree data. Euphytica 124:293-306.

Rajaram, S. 1995. Wheat germplasm improvement: Historical perspectives, philosophy, objectives, and missions. pp. 1-10. *In* Rajaram, S., and G.P. Hettel (eds.) Wheat Breeding at CIMMYT: Commemorating 50 years of research in Mexico for global wheat improvement. Wheat special report No. 29, CIMMYT, Mexico D.F., Mexico.

Rosegrant, M.W., A. Agcaoili-Somlilla, and N. Perez. 1995. Global food projections to 2020. Discussion paper v. 5. IFPRI, WA.

Sears, E.R. 1954. The aneuploids of bread wheat. Ann. Rev. Genet., 10:31-51.

Sears, E.R. (1976). Genetic control of chromosome pairing in wheat. Annu. Rev. Genet. 10:31-51.

Smale, M. 1996. Understanding of global trends in the use of wheat diversity and international flows if wheat genetic resources. Economics working paper 96-02. CIMMYT. Mexico D.F., Mexico.

Souza, E., P.N. Fox, D. Byerlee, and B. Skovmand. 1994. Spring wheat diversity in irrigated areas of two developing countries. Crop Sci. 34:774-783.

St. Martin, S.K. 1982. Effective population size for the soybean improvement program in maturity groups 00 to IV. Crop Sci. 22:151-152.

Valkoun, J.J. 2001. Wheat pre-breeding using wild progenitors. Euphytica 119:17-23.

Warburton, M.L., B. Skovmand, and A. Mujeeb-Kazi. 2002. The molecular genetic characterization of the 'Bobwhite' bread wheat family using AFLPs and the effect of the T1BL.1RS translocation. Theor. Appl. Genet. 104:868-873.

Wright, S. 1922. Coefficients of inbreeding and relationship. Am. Nat. 56:330-33.

CROP BREEDING, GENETICS & CYTOLOGY

SSR and Pedigree Analyses of Genetic Diversity among CIMMYT Wheat Lines Targeted to Different Megaenvironments [1]

S. Dreisigacker, P. Zhang, M. L. Warburton, M. Van Ginkel, D. Hoisington, M. Bohn, and A. E. Melchinger*

ABSTRACT

Improved bread wheat (*Triticum aestivum* L.) cultivars for diverse agroecological environments are important for success in the effort to increase food production. In the 1980s, CIMMYT introduced the megaenvironment (ME) concept to breed wheats specifically adapted to different areas. Our objective was to analyze the genetic diversity among 68 advanced CIMMYT wheat lines targeted to different MEs by using 99 simple sequence repeats (SSRs) and the coefficient of parentage (COP). The average number of alleles detected was higher for the 47 genomic SSRs (5.4) than for the 52 SSRs derived from expressed sequence tags (EST) (3.3), but gene diversity between MEs was similar for both types of markers. No significant differences among the five MEs were observed for the means of SSR-based genetic similarities (GS), calculated as $1 -$ Rogers' distance, and COP values. Both measures showed a low correlation ($r = 0.43$). High levels of genetic diversity were found within the germplasm targeted to each ME. However, principle coordinate analysis based on modified Rogers' distances did not separate the genotypes according to their targeted MEs. We conclude that presence of a single core germplasm can reflect large phenotypic differences. A sufficient number of diverse breeding lines for each ME is required because MEs generally combine various production areas. SSRs represent a powerful tool to quantify genetic diversity in wheat, but genotypic differentiation for adaptation to specific MEs in the CIMMYT program could not be proven.

WHEAT, together with maize (*Zea mays* L.) and rice (*Oryza sativa* L.), is one of the three major food crops in the world. It is grown in a variety of environments, ranging from fully irrigated (e.g., northern India, Egypt), to high rainfall (e.g., northwestern Europe, eastern Africa, southern zone of Latin America), and drought-prone regions (e.g., U.S. Great Plains, most of Australia, parts of Argentina). In these areas wheat production experiences a range of biotic and abiotic stresses and crop improvement requires precise focusing on the needs of the crop in each area, the producers, the processing industry, and the consumers (Lantican et al., 2002).

More than one half of the wheat production environments are located in developing countries, which fall within the mandate of CIMMYT. In the 1980s, CIMMYT introduced the concept of breeding for different MEs. A ME is defined as a large, not necessarily contiguous area, which usually encompasses more than one country and is frequently transcontinental. It is characterized by similar biotic and abiotic stress conditions, cropping systems, and consumer demands (Rajaram et al., 1994). Twelve MEs have been classified, six of which are focused on efficient selection of better-adapted spring bread wheat, the dominant type of wheat in developing countries. The concept has permitted expanding breeding efforts relevant within each ME. In breeding for enhanced adaptation, adequate genetic diversity is a prerequisite for any crop improvement program. The genetic progress through selection is directly related to the variability present in the gene pool, and the quality of the genes contributed by the parents.

The COP is an indirect measure of genetic diversity among genotypes based on the probability that alleles at a certain locus are identical by descent. Calculation of COP values rests on simplifying assumptions regarding the relatedness of ancestors, parental contribution to the offspring, and absence of selection and genetic drift, which are not met under breeding conditions (Cox et al., 1985; Cowen and Frey, 1987). In contrast, molecular markers measure diversity directly at the DNA level. In studies of autogamous crops with low levels of apparent genetic variability such as wheat, soybean [*Glycine max* (L.) Merr.], and rice, SSRs proved to be a suitable marker system. They are generally genome specific, abundant, codominant in nature, and show a fairly uniform distribution over the genome. SSRs have been applied in many aspects of genetic diversity analyses such as genetic differentiation caused by selection (Stachel et al., 2000), fingerprinting of genotypes to analyze the structure of germplasm collections (Parker et al., 2002; Huang et al., 2002), and the analysis of temporal changes in diversity (Donini et al., 2000; Christiansen et al., 2002).

Traditional methods to develop SSRs are based on isolating and sequencing genomic libraries, which contain putative SSR tracts (Adams et al., 1992). A novel source for generating SSRs is provided by screening EST databases available online (Kota et al., 2001). This recent approach allows researchers to shift from the use of anonymous markers with unknown effect on the phenotype to markers physically associated with coding

S. Dreisigacker and A.E. Melchinger, Inst. of Plant Breeding, Seed Science, and Population Genetics, Univ. of Hohenheim, 70593 Stuttgart, Germany; M. Bohn, Dep. of Crop Science, Univ. of Illinois, Urbana, IL 61801; P. Zhang, M. van Ginkel, M.L. Warburton, and D. Hoisington, CIMMYT, Mexico D.F., Mexico. Received 17 March 2003. *Corresponding author (melchinger@pz.uni-hohenheim.de).

Abbreviations: AMOVA, analysis of molecular variance; CIMMYT, International Maize and Wheat Improvement Center; COP, coefficient of parentage; EST, expressed sequence tag; GS, genetic similarity; ME, megaenvironment; SSR, simple sequence repeat.

regions, which may more accurately reflect the effects of selection, both natural and artificial.

The objectives of the present study were to (i) evaluate the use of genomic and EST-derived SSRs for determining the genetic diversity among advanced spring bread wheat lines from the CIMMYT breeding program, (ii) compare genetic distances based on SSRs with the COP estimates of these wheat lines, and (iii) determine the diversity for SSRs within and among sets of lines targeted to different MEs.

MATERIALS AND METHODS

Plant Materials

A total of 68 CIMMYT advanced spring bread wheat lines from crosses made during 1989 to 1996 were chosen for this study (Table 1). Most of the lines were bred by a "modified bulk" procedure described by Van Ginkel et al. (2002). Seed of outstanding F_7 lines was harvested in bulks for subsequent yield trials. These yield trials were grown in replicated and latinized α-lattice designs at Cd. Obregon (Sonora, Mexico) or Toluca (Mexico State, Mexico) in 2000 and 2001 under conditions simulating the different MEs for which they are being bred (e.g., full irrigation, reduced irrigation, drought, and heat stress, etc.). On the basis of their performance, 8 to 15 advanced lines were selected from yield trials representative of the first five spring bread wheat MEs (Table 2). The lines were chosen as candidates for further evaluation at international testing sites (Van Ginkel et al., 2002). Progenies from three crosses (Alucan/Duluca, PF869107/CEP8825//Milan and Babax/Amadina//Babax) were identified and selected for more than one ME.

SSR Analyses

DNA extraction was performed with the CTAB method of Saghai-Maroof et al. (1984) modified according to CIMMYT Applied Biotechnology Center's Manual of Laboratory Protocols (Hoisington et al., 1994). Twenty seeds per advanced line were grown in the greenhouse and after 2 wk young leaves were harvested from 5 to 10 plants per line. Leaves were bulked for DNA extraction to assess the genetic variability within each line as described by Gilbert et al. (1999). Quality and quantity of the isolated DNA was determined on 1% (w/v) agarose gels by comparing bands to known concentrations of λ DNA.

SSR information was obtained from two different sources: 46 SSRs were collected from a conventional genomic library (genomic SSRs) developed at IPK Gatersleben by Röder et al. (1998 and unpublished data) and 51 SSRs derived from ESTs (EST-SSR) with the prefix "*DuPw*" were kindly provided by DuPont, Wilmington, DE (Dupont, unpublished data; Eujayl et al., 2002). The SSRs from both sources were distributed equally over the genome. In addition, the 1BS/1R translocation EST-SSR marker *Taglgap* (Devos et al., 1995) and the SSR marker *WMC56* developed by the Wheat Microsatellite Consortium (Agrogene, France) were used. Details for each of the 99 SSRs can be found online (http://www.cimmyt.org/english/webp/support/publications/support_materials/ssr_mw1.htm; verified 23 September 2003).

PCR reactions were performed in a model PTC225 thermocycler (MJ Research, Inc., Waltham, MA). Each 20-μL reaction mixture contained 25 ng template DNA, 150 nM of each primer, 250 μM dNTPs, 200 μM MgCl$_2$, 1× PCR buffer and 2.5 U of *Taq*-polymerase. Forward primers were labeled at the 5′ end with either one of three phosphoramidite fluores-

cent dyes 6-carboxyfluoresein, tetrachloro-6-carboxyfluoresein, or hexachloro-6- carboxyfluoresein. PCR was performed with the following standard temperature profile: 29 cycles with a 1 min denaturing step at 94°C, 2 min annealing temperatures between 50 and 64°C depending on the different primer combinations, and 2 min extension at 72°C. The 1-min time spread of the standard profile cycle was modified in some cases to fully optimize amplification conditions.

Amplification products were separated on an ABI™377 Sequencer (Perkin Elmer/Applied Biosystems, Foster City, CA) using 4.5% (w/v) polyacrylamide denaturing gels (acrylamide:bisacrylamide 29:1). Running conditions were 2400 V, 40 mA, 120 W electrophoresis power and 40 mW laser power. Products from up to five SSRs could be distinguished simultaneously because of the three different fluorescent dyes and migration distance differences. Fragment sizes were calculated semiautomatically by the computer software GeneScan 3.1 (Perkin Elmer/Applied Biosystems) by comparing fragments with an internal size standard (GeneScan 350 or 500) labeled with N,N,N,N,-tetramethyl-6-carboxyrhodamine. GeneScan fragments were assigned to alleles by the category function of the software Genotyper 2.1 (Perkin Elmer/Applied Biosystems). Sixty-four genotypes were run on each gel plus two wheat lines, Opata and Synthetic, as controls.

We could not optimize the amplification profile of nine SSRs for the scoring with Genotyper. These markers were optimized to run on small (16 by 20 cm) 6% (w/v), 19:1 acrylamide:bis-acrylamide denaturing gels (*ATTO*[8] AE-6220). The gels were run for 2 h at about 350 V, with a 100-bp ladder as a standard. For fragment visualization, silver staining was applied according to Applied Biotechnology Center's Manual of Laboratory Protocols. The fragment length of each SSR was determined with the scientific image system Kodak ID 2.02 (Kodak, New Haven, CT).

Statistical Analyses

Reproducibility of SSR amplification and scoring was determined on the basis of the percentage of disagreements in the fragment size of the two standard lines, Opata and Synthetic, for gels loaded with the same markers. Allele frequencies at the 99 loci, total gene diversity (H_T), gene diversity within MEs (H_S), and the proportion of diversity resulting from gene differentiation between MEs (G_{ST}) were calculated according to Nei (1987). The measures were considered separately for the two SSR sources to examine the influence of the genome location of the markers (genomic and EST-SSRs).

The COP for all pairwise combinations of wheat lines was calculated on the basis of fully expanded genealogical information extracted from CIMMYT's International Wheat Information System (Payne et al., 2002). Calculations of COP were based on the assumptions described by St. Martin (1982), except for sister lines, which were assigned a COP of 0.56 instead of 1.0 following Cox et al. (1985). Genetic similarity (GS) assigned as $1 -$ Rogers' distance (Rogers, 1972) was estimated to compare SSR-based and COP estimates. Pearson's correlation coefficient (r) between GS and COP values was calculated for related pairs of lines (COP ≥ 0.05). Standard errors of genetic similarity estimates were obtained by a bootstrap procedure with resampling over markers (Weir, 1996). Furthermore, modified Rogers' distance (Wright, 1978) was calculated among all possible pairs of lines as a basis for the application of multivariate methods, because it represents a Euclidean distance.

An analysis of molecular variance (AMOVA) on the basis of SSR data was computed to test the differentiation of the 68 genotypes according to the five MEs. The hierarchical analysis

Table 1. Pedigrees of the 68 CIMMYT spring bread wheat lines classified by megaenvironments (ME).

No.	Pedigree†
ME1IR (irrigated zones)‡	
1	KAUZ//ALTAR 84/AOS/3/KAUZ
2	ATTILA/3/HUI/CARC//CHEN/CHTO/4/ATTILA
3	PRINIA/WEAVER//STAR/3/WEAVER
4	OASIS/4*BORL95
5	WEAVER/WL3926//SW89.3064
6	RABE/2*MO88
7	CNDO/R143//ENTE/MEXI_2/3/AEGILOPS SQUARROSA (TAUS)/4/WEAVER/5/2*KAUZ
8	CHEN/AEGILOPS SQUARROSA (TAUS)//FCT/3/2*WEAVER
9	CHEN/AEGILOPS SQUARROSA (TAUS)//FCT/3/STAR
10	CHUM18/5*BCN
11	P1.861/RDWG
12	CMH80A.542/CNO79
13	SUPER SERI #2
14	PVN//CAR422/ANA/5/BOW/CROW//BUC/PVN/3/YR/4/TRAP#1
15	BABAX/AMADINA//BABAX (WEEBILL1)
ME1HT (irrigated hot zones)	
16	VEE/PJN//2*TUI
17	PFAU/WEAVER
18	CAZO/KAUZ//KAUZ
19	MNCH/3*BCN
20	W462//VEE/KOEL/3/PEG//MRL/BUC
21	XIANG82.2661/2*KAUZ
22	SW89-5124*2/FASAN
23	KEA/TAN/4/TSH/3/KAL/BB//TQFN/5/PAVON/6/SW89.3064
24	LAJ3302/2*MO88
25	CROC_1/AE.SQUARROSA (205)//2*BCN
26	PICUS/4/CS5A/5RL-1//BUC/BJY/3/ALD/PVN/5/LAJ3302
27	PFAU/MILAN
28	TAM200/TUI
29	SABUF/7/ALTAR 84/AE.SQUARROSA (224)//YACO/6/CROC_1/AE.SQUARROSA (205)/5/BR12*3/4/IAS55*4/CI14123/3/IAS55*4/EG,AUS//IAS55*4/ALD
30	[KASYON]//PVN/SPRW
ME2 (high rainfall zones)	
31	ALUCAN/DUCULA
32	IAS58/4/KAL/BB//CJ71/3/ALD/5/CNR/6/THB/CEP7780
33	R37/GHL121//KAL/BB/3/JUP/MUS/4/2*YMI #6/5/CBRD
34	TNMU/6/CEP80111/CEP81165/5/MRNG/4/YKT406/3/AG/ASN//ATR
35	TNMU/MILAN
36	TNMU/6/PEL74144/4/KVZ//ANE/MY64/3/PF70354/5/BR14/7/BR35
37	DUCULA/TNMU
38	TRAP#1/BOW//VEE#5/SARA/3/ZHE JIANG 4/4/DUCULA
39	PASTOR//MUNIA/ALTAR 84
40	OR791432/VEE#3.2//MILAN
41	MUNIA/ALTAR 84//AMSEL
42	TNMU/MUNIA
43	TNMU/ATTILA
44	HXL8088/DUCULA
45	PF869107/CEP8825//MILAN
ME3 (high rainfall, acid soil zones)	
46	ALUCAN/DUCULA
47	TNMU/TUI
48	TNMU/OCEP17
49	OCEP15/KAUZ//TNMU
50	BR14*2/SUM3/TNMU
51	TNMU/BR35//THB/CEP7780
52	KVZ/3/TOB/CTFN//BB/4/BLO/5/TAN/6/PRL/7/MILAN
53	PF869107/CEP8825//MILAN
ME4 (semiarid zones)	
54	DUCULA//VEE/MYNA
55	SRMA/TUI
56	CROC_1/AE.SQUARROSA (224)//OPATA
57	PIOS/DUCULA
58	LAJ3302/3/GZ156/NAC//PSN/URES/4/WEAVER
59	3VASKAR/G303.1M.1.3.2.2.2//KAUZ/3/SKAUZ/4/KAUZ
60	PASA/SAET
61	NL456/VEE#5//CHIL/3/MUNIA
62	TZPP*2/ANE//INIA/3/CNO67/JAR//KVZ/4/MN72252/5/SHI#4414/CROW
63	TSI/VEE#5//KAUZ
64	KAUZ/5/PAT10/ALD//PAT72300/3/PVN/4/BOW
65	KA/NAC
66	ALTAR 84/AE.SQ//2*OPATA
67	FRET2
68	BABAX/AMADINA//BABAX (WEEBILL1)

† Nomenclature according to Purdy et al. (1968): The initial cross is indicated by a single slash (e.g. A/B), the second cross by a double slash (e.g. A/B// C), and subsequent crosses in numerical order by flanked single slashes (e.g. A/B//C/3/D). Backcrosses are designated with an asterisk (*) and a number indicating the dosage of the recurrent parent.
‡ Refers to the five MEs described in Table 2.

Table 2. Characterization of important spring bread wheat megaenvironments (ME) defined by CIMMYT (Rajaram et al., 1994).

ME	Moisture regime	Temperature regime	Breeding objectives in addition to yield†	Year‡	Yield trials	Contribution of wide crosses§
ME1IR	Low rainfall, irrigated	Temperate	Resistance to lodging, SR, YR and LR, end-use quality	1945	Obregon: 700 mm by irrigations	China, India, Synthetic, Durum wheats
ME1HT	Low rainfall, irrigated	Hot	As for ME1IR plus tolerance to heat	1945	Obregon: 700 mm by irrigations, late planting	China, Argentina, India, Synthetic, Durum wheats
ME2	High rainfall	Temperate	Resistance to SR, YR, LR, *Septoria* spp., FHB, BYDV, waterlogging, pre-harvest sprouting. end-use quality	1972	Toluca: high rainfall (800 mm)	China, Brazil, Durum wheats
ME3	High rainfall	Temperate	As for ME2 plus Al and Mn tolerance, P-use efficiency	1974	Toluca: high rainfall (800 mm), seedling test for low pH and Al toxicity	China, Brazil
ME4	Low rainfall	Temperate or hot	Resistance to SR, YR, LR plus tolerance to drought, end-use quality	1970	Obregon: one pre-seeding irrigation, 300 mm available	Argentina, Nepal, Synthetic wheats

† SY = Stem rust, YR = Yellow rust, LR = Leaf rust, FHB = Fusarium head blight, BYDV = Barley Yellow Dwarf Virus.
‡ Refers to the year in which breeding for the respective ME began at CIMMYT.
§ Refers to the 68 lines included in this study.

divides the total variance into variance components due to intra- or inter-ME differences and tests their significance. Principal coordinate analysis was performed on the basis of the modified Rogers' distances to visualize the dispersion of the genotypes (Gower, 1966). The K-means clustering algorithm was used to identify groups of similar lines, on the basis of a least-squares partitioning method, which divides a collection of objects into k clusters depending on minimum distances to the centers of the clusters (MacQueen, 1967). COP values were calculated to the six progenitors most frequently used in the crosses and averaged within each ME and K-means cluster.

All analyses were performed with the Plabsim software (Frisch et al., 2000), which is implemented as an extension of the statistical software R (Ihaka and Gentleman, 1996). The AMOVA was performed by the software package Arlequin 2.0 (Schneider et al., 2000).

RESULTS

For the 68 CIMMYT advanced lines analyzed with 99 SSRs, a total of 425 alleles was detected with an average of 4.3 alleles per locus. The average number of alleles was considerably lower in EST-SSRs (3.3) than in genomic SSRs (5.4), with seven out of the 52 EST-SSRs being monomorphic (Table 3). However, an important feature of EST-SSRs was the high-quality fragment patterns obtained, which were devoid of stutter bands, resulting in a higher reproducibility (98.8%) and lower residual heterozygosity (1.4%) than with genomic SSRs (89.5 and 3.7%, respectively).

The total gene diversity (H_T) varied widely among loci from 0.01 at *DuPw138* to 0.83 at *Xgwm437*, with an average of 0.47 (Table 3). Considering the two different SSR sources, the average H_T and H_S values were lower for EST-SSRs than for genomic SSRs. The correspond-

ing G_{ST} value for all loci was 0.09 and for EST-SSRs (0.10) just slightly higher than for genomic SSRs (0.09).

The mean COP value over the 68 wheat genotypes was 0.14 and ranged from 0.01 to 0.87 for closely related pairs (Table 4). The mean COP values within MEs did not substantially differ between the five MEs. Thirty-nine percent of the COP values were smaller than 0.10, indicating that theoretically less than 10% of the genetic material segregating in ancestral populations was identical by descent in any two cultivars.

GS for all pairs of lines ranged from 0.39 to 0.91 with an average of 0.59 for all genotypes (Table 4). Similar to the COP values, MEs were not significantly different in their mean GS, but with an equal range of values. In the specific cases in which two progenies of the crosses Alucan/Duluca, PF869107/CEP8825//Milan and Babax/Amadina//Babax were selected for different MEs, GS values were high (0.88, 0.86, and 0.72, respectively), as expected.

The mean COP and GS values between MEs were of similar size as the means within MEs (Table 5). ME1HT was most distant to ME2 and ME3 on the basis of COP, and ME1IR most distant to ME3 on the basis of GS values. The AMOVA confirmed these results in that 92% of the total variation was found within MEs and just 8% between MEs (data not shown). The correlation between GS and COP values was $r = 0.43$.

The principal coordinate analysis based on modified Rogers' distances did not separate the genotypes according to their targeted MEs (Fig. 1). Fourteen of the chosen genotypes cluster somewhat together because of their resistance to acid soil. The K-means cluster algorithm identified more than one solution, the most frequent (90%, 1000 repetitions) comprising three defi-

Table 3. Average number of alleles per locus, residual heterozygosity, reproducibility and gene diversity estimated over the two different sources of markers used in this study.

SSR source	No. of SSR	Avg. no. of alleles/locus	Heterozygosity	Reproducibility	Gene diversity†		
					H_T	H_S	G_{ST}
Genomic SSRs	47	5.4	3.7	89.5	0.57	0.52	0.09
EST derived SSRs	52	3.3	1.4	98.8	0.37	0.33	0.10
Total	99	4.3	2.5	95.7	0.47	0.43	0.09

† H_T = total gene diversity, H_S = diversity within megaenvironments (MEs), G_{ST} = diversity between MEs.

Table 4. Total and unique number of alleles, number of monomorphic loci, mean genetic similarities (GS), and mean coefficient of parentage (COP) within each megaenvironment (ME).

ME	No. of lines	No. of alleles	No. of unique alleles†	No. of mono-morphic loci	GS			COP		
					Mean	Min.	Max.	Mean	Min.	Max.
ME1IR	15	267	20	21	0.59 ± 0.06	0.44	0.85	0.18 ± 0.11‡	0.02	0.59
ME1HT	15	289	33	17	0.59 ± 0.07	0.45	0.91	0.14 ± 0.11	0.01	0.48
ME2	15	253	12	22	0.64 ± 0.06	0.52	0.80	0.11 ± 0.08	0.02	0.37
ME3	8	232	18	24	0.60 ± 0.06	0.47	0.73	0.13 ± 0.06	0.04	0.24
ME4	15	272	23	17	0.63 ± 0.06	0.48	0.85	0.18 ± 0.08	0.04	0.50
Total	68	425	–	7	0.59 ± 0.06	0.39	0.91	0.14 ± 0.09	0.01	0.87

† Alleles occurring only in one ME.
‡ Standard deviation.

nite centers. K-means tended to form the clusters on the basis of common progenitors used in the crosses made during 1989 and 1996. Seven of the 11 lines with Kauz in the pedigree were included in cluster K1 (Table 6). Four lines containing Kauz did not group into this cluster. These lines had Chinese wheats in their pedigree or were selected in later segregating generations under different environmental conditions by the International Center of Agricultural Research in Dry Areas in Syria. On the basis of COP, Weaver holds the highest parental contribution to cluster K2 and Milan to cluster K3. Progenitor Tinamou contributed about equally to clusters K2 and K3. Twenty of the 68 genotypes had durum (*T. durum* Desf.) wheat, Chinese wheat, or synthetic hexaploid wheat in their pedigree. These lines were scattered all over the principal coordinate analysis plot because of different sources of Chinese lines and *Aegilops squarrosa* L. in the synthetic hexaploid wheats used as crossing parents.

DISCUSSION

Use of Genomic and EST-SSRs in Breeding Programs

Genomic SSR markers have been intensively used to detect the variability between bread wheat genotypes, but the large genome size of wheat is a challenge in identifying sufficiently robust and informative SSRs for fingerprinting. EST-SSRs present a novel source of SSRs and have some intrinsic advantages over genomic SSRs. They can be developed from available EST databases and their frequency is abundantly high in transcribed regions (Morgante et al., 2002). A concern is that the coding character of EST-SSRs limits their level of polymorphism.

Our results agree with other studies in rice (Cho et al., 2000), grape (*Vitis* ssp., Scott et al., 2000), and wheat (Eujayl et al., 2002) in that the overall level of polymorphism for genomic SSRs was higher than for EST-SSRs. However, compared with the latter study with 64 durum

wheat lines, where on average 5.5 alleles per locus for genomic SSRs and 4.1 for EST-SSRs were found, we detected slightly lower average numbers of alleles for both genomic (5.2) and EST-SSRs (3.2). The somewhat higher level of diversity reported by Eujayl et al. (2002) may be attributable to the more variable material in their study, which comprised a sample of various lines and landraces with different genetic backgrounds. Here, we studied advanced breeding lines ready for international dissemination to developing country breeding programs.

Higher-order repeat motifs were applied in our sample of EST-SSRs. In comparison to dinucleotide motifs, they are generally less polymorphic and insensitive to single-nucleotide polymorphisms in the flanking regions of the SSRs, which facilitate the designation of allele sizes (Chakraborty et al., 1997; Song et al., 2002; Mogg et al., 2002). This explains the higher reproducibility and the lower degree of heterozygosity or heterogeneity for EST-SSRs compared with genomic SSRs in our study.

EST-SSRs are assumed to reflect more accurately the effects of selection under which the germplasm has been developed. However, only a slightly better differentiation of the allele frequencies between MEs was observed for EST-SSRs than for genomic SSRs. Wheat is highly autogamous and, therefore, differential selection of fitness-related target loci will also affect genomic SSRs linked to them. Further evidence shows a functional importance of genomic SSR structures, which may cause some form of balancing selection (Innan et al., 1997; Li et al., 2000; Li et al., 2002).

Future opportunities to combine markers and phenotypic data in association studies may improve the application of EST-SSRs in the evaluation of germplasm, as exemplified in maize by Thornsberry et al. (2001). A specific marker could then be used to examine the functional diversity at a certain locus. We speculate that EST-SSRs from genes contributing to specific ME adap-

Table 5. Mean coefficient of parentage (above diagonal) and genetic similarity estimates (below diagonal) between five CIMMYT megaenvironments (ME).

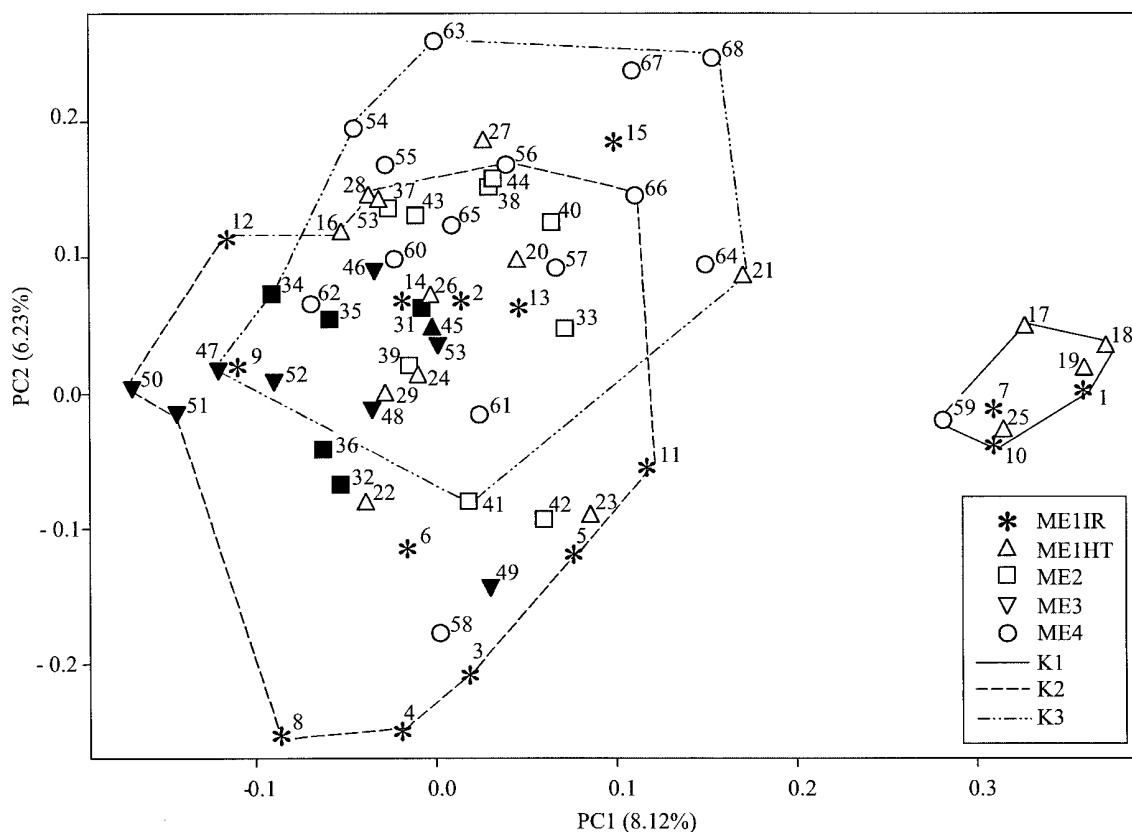|	ME1IR	ME1HT	ME2	ME3	ME4
ME1IRR		0.16 ± 0.12†	0.11 ± 0.06	0.11 ± 0.05	0.18 ± 0.10
ME1HT	0.57 ± 0.05		0.10 ± 0.05	0.10 ± 0.07	0.16 ± 0.10
ME2	0.62 ± 0.06	0.59 ± 0.05		0.14 ± 0.11	0.12 ± 0.04
ME3	0.56 ± 0.05	0.58 ± 0.08	0.59 ± 0.06		0.12 ± 0.06
ME4	0.58 ± 0.06	0.59 ± 0.07	0.62 ± 0.07	0.60 ± 0.07	

† Standard deviation.

Fig. 1. Associations among 68 CIMMYT wheat lines revealed by principal coordinate analysis performed with modified Rogers' distance estimates calculated from 99 SSR loci. Numbers refer to the 68 wheat lines, megaenvironments (ME) are designated by circles, triangles, and squares. K-mean clusters are indicated by solid and dashed lines, respectively.

tation might have revealed more clearly the effects of selection in our germplasm.

Correlation between Pedigree and SSR-Based Distance Estimates

In agreement with previous studies in wheat, the correlation between GS and COP estimates reported here was low but significant ($r = 0.43$). This low correlation can be explained by the unrealistic assumptions in calculating COP values and the substantial variation in GS estimates of unrelated lines (Graner et al., 1994). With increasing relatedness of lines, the association between COP and their corresponding GS values should become tighter. In fact, a higher correlation ($r = 0.55$) between COP (COP ≥ 0.01) and GS estimates was found in the study of Plaschke et al. (1995), where the average COP among a set of European wheats was 0.29. However, selection and drift are presumably also important factors reducing the correlation between GS and COP due to a shift in the genomic contribution of parental lines, particularly during the early selfing generations.

Variation of Genetic Diversity in Megaenvironments

Provided molecular markers represent an accurate picture for the genetic diversity at functional genes, the

Table 6. Average coefficient of parentage (COP) of predominant progenitors used (i) in crosses of this study within megaenvironments (ME) and (ii) in clusters revealed by the K-means algorithm performed on modified Rogers' distance.

Groups	Advanced lines	Predominant progenitor					
		Kauz	Parula	Weaver	Tinamou	Milan	Ducula
ME							
ME1IR	1-15	16.5	12.1	18.8	0.0	0.0	0.0
ME1HT	16-30	20.8	4.2	3.3	0.0	3.3	0.0
ME2	31-45	0.0	0.0	0.0	18.3	16.7	13.3
ME3	46-53	3.1	0.0	0.0	28.1	12.5	6.3
ME4	54-68	12.5	1.7	3.3	0.0	0.0	6.7
Mean		11.4	4.0	5.6	7.4	3.7	5.2
K-means cluster performed on MRD							
K1	1, 7, 10, 17, 18, 19, 25, 59	71.5	4.7	9.4	0.0	0.0	0.0
K2	3-6, 8, 9, 11, 12, 16, 22, 23, 26, 28, 29, 32, 36, 42, 45, 48-53, 56, 58, 62, 66	0.8	6.9	11.0	8.9	0.0	5.4
K3	2, 13-15, 20, 21, 24, 27, 30, 31, 33-35, 37-41, 43, 44, 46, 47, 54, 55, 57, 60, 61, 63-66, 67, 68	5.2	1.1	0.0	7.4	10.6	2.9

true genetic differences between the germplasm targeted for different MEs are small. Several reasons might explain this observed absence of genetic differentiation: (i) selection based on ME adaptation has not been practiced long enough to differentiate the germplasm, (ii) genes conferring fitness to one ME are not unique to that ME and may confer fitness to several MEs, and (iii) adaptation to MEs is not based on an accretion of random genes but rather a limited set of specific genes.

CIMMYT's concept of breeding for different MEs was implemented in the 1980s. Thus, selection history for ME adaptation has presumably been too short to result in a detectable genetic differentiation. In addition, shuttle breeding between two environmentally contrasting sites in Mexico (Cd. Obregon and Tuluca) during the selfing generations and intermating of germplasm adapted to different ME may have leveled the genetic differentiation between MEs (Rajaram and Van Ginkel, 2001).

The uniform level of relatedness between genotypes targeted to the five MEs (Table 4) and the principal coordinate analysis (Fig. 1) suggest the presence of a single core germplasm providing genes or gene combinations conferring fitness to several MEs. This is not surprising, because some MEs differ only in few of the classification criteria. For example, the high rainfall environment ME3 can be characterized as a specific subenvironment of ME2 and, consequently, selection to both environments depends on the same major abiotic and biotic stresses (Table 2).

Although many genetic diversity studies confirmed the use of SSRs for germplasm identification, it is an open question whether they are able to reveal functional diversity. Adaptation to different ME might be attributable to only a small number of genes regulating the underlying physiological processes, which were not reflected by the applied SSRs (Nevo, 2001). An exception was the cluster of acid soil resistant lines in Fig. 1, but this could also be explained by the fact that 8 of the 14 lines had the highly acid soil resistant line Tinamou as a parent in their pedigree.

High levels of genetic diversity were found within the germplasm targeted to each ME. To warrant diversity in every breeding cycle, on average about 25% of the crossing block entries are replaced with outstanding new introductions (Van Ginkel et al., 2002). In our study too, at least 20 wheat lines were included with ancestors from nonconventional sources such as Chinese, durum, or synthetic hexaploid wheats.

The high genetic diversity observed within each ME is desired in the CIMMYT wheat program for two reasons. First, the wheat producing areas combined in each ME are fairly diverse and often transcontinental. Diseases and especially races of fungal pests may vary considerably between regions as well as quality demands due to different wheat processing techniques and utilization of the end product. Furthermore, environmental conditions in the target areas are fluctuating based on highly variable seasons (Trethowan et al., 2002). Newly developed CIMMYT lines should reach the cooperating wheat research programs in the respective MEs with still sufficient inherent genetic variation remaining among

them for many traits. This enables cooperators to reselect and release cultivars adapted to their own local needs. Therefore, a sufficient number of diverse CIMMYT breeding lines for each ME reduces the risk of genetic vulnerability.

Second, breeding conditions at the test sites in Mexico and the targeted ME are not always closely associated. Trethowan et al. (2001) examined the relationship among various locations using yield data from CIMMYT's Semi-Arid Wheat Yield Trials corresponding to ME4. They found that the yield performance under residual moisture stress at the test location in Cd. Obregon (Mexico) was a good predictor of yield performance at other locations experiencing equivalent late drought stress, but this was not true when drought patterns were different, e.g., when early drought stress occurred.

The common genetic base of our germplasm targeted to different ME did not reflect CIMMYT's efforts to breed for a large number of environments. We speculate that assembling more diversified germplasm pools for different MEs and yield evaluation at ME-specific key sites could achieve higher genetic differentiation. However, increased genetic diversity does not necessarily lead to higher productivity or adaptation. While SSRs are a powerful tool for genotype identification, their usefulness for revealing genotype differentiation regarding specific traits such as adaptation to certain MEs could not be proven in the present study and warrants further research. In particular, the augmented use of EST-SSRs from genes with known functions should be very powerful for this purpose and, thus, could give further directions in the management of wheat breeding programs aiming to optimally serve their clients.

ACKNOWLEDGMENTS

REFERENCES

Adams, M.D., M. Dubnick, A.R. Kerlavage, R. Moreno, J.M. Kelley, T.R. Utterback, J.W. Nagle, C. Fields, and J.C. Venter. 1992. Sequence identification of 2375 human brain genes. Nature 355: 632–634.

Chakraborty, R., M. Kimmel, D.N. Stivers, L.J. Davison, and R. Deka. 1997. Relative mutation at di-, tri-, and tetranucleotide microsatellite loci. Proc. Natl. Acad. Sci. (USA) 94:1041–1046.

Cho, Y.G., T. Ishii, S. Temnykh, X. Chen, L. Lipovich, S.R. McCouch, W.D. Park, N. Ayres, and S. Cartinhour. 2000. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor. Appl. Genet. 100:713–722.

Christiansen, M.J., S.B. Anderson, and R. Ortiz. 2002. Diversity changes in an intensively bread wheat germplasm during the 20th century. Mol. Breed. 9:1–11.

Cox, T.S., Y.T. Hkiang, M.B. Gorman, and D.M. Rogers. 1985. Relationships between coefficient of parentage and genetic indices in the soybean. Crop Sci. 25:529–532.

Cowen, N.M., and K.J. Frey. 1987. Relationships between genealogical

distance and breeding behavior in oat (*Avena sativa* L.). Euphytica 36:413–424.

Devos, K.M., G.J. Bryan, A.J. Collins, P. Stephenson, and M.D. Gale. 1995. Application of two microsatellite sequences in wheat storage proteins as molecular markers. Theor. Appl. Genet. 90:247–252.

Donini, P., P. Stephenson, G.J. Bryan, and R.M.D. Koebner. 2000. The potential of microsatellites for high throughput genetic diversity assessment in wheat and barley. Genet. Resour. Crop Evol. 45:415–421.

Eujayl, I., M.E. Sorrells, M. Braun, P. Wolters, and W. Powell. 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor. Appl. Genet. 104:399–407.

Frisch, M., M. Bohn, and A.E. Melchinger. 2000. Plabsim: Software for simulation of marker-assisted backcrossing. J. Hered. 91:86–87.

Gilbert, J.E., R.V. Lewis, M.J. Wilkinson, and P.D.S. Caligari. 1999. Developing an appropriate strategy to assess genetic variability in plant germplasm collections. Theor. Appl. Genet. 98:1125–1131.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–388.

Graner, A., W.F. Ludwig, and A.E. Melchinger. 1994. Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. Crop Sci. 34:1199–1205.

Hoisington, D., M. Khairallah, and D. Gonzalez-de-Leon. 1994. Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory. 2nd ed. (Available online with updates at (http://www.cimmyt.cgiar.org/ABC/Protocols/manualABC.html; verified 22 September 2003). CIMMYT, Mexico, D.F.

Huang, X.Q., A. Börner, M.S. Röder, and M.W. Ganal. 2002. Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. Theor. Appl. Genet. 105:699–707.

Ihaka, R., and R. Gentleman. 1996. R: A Language for data analysis and graphics. J. Comput. Graph. Stat. 5:299–314.

Innan, H., R. Terauchi, and T. Miyashita. 1997. Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146:1441–1452.

Kota, R., R.K. Varshney, T. Thiel, K.J. Dehmer, and K.J. Graner. 2001. Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum vulgare* L.). Hereditas 135:145–151.

Lantican, M.A., P.L. Pingali, and S. Rajaram. 2002. *In* J. Ekboir (ed.) World wheat overview and outlook 2000–2001: Developing no-till packages for small scale farmers. CIMMYT, Mexico, D.F.

Li, Y.-C., T. Fahima, J.H. Peng, M.S. Röder, V.M. Kirzhner, A. Beiles, A.B. Korol, and E. Nevo. 2000. Edaphitic microsatellite DNA divergence in wild emmer wheat, *Triticum dicoccoides*, at a microsite: Tabigha, Israel. Theor. Appl. Genet. 101:1029–1038.

Li, Y.-C., A.B. Korol, T. Fahima, A. Beiles, and E. Nevo. 2002. Mircosatellites: Genomic distribution, putative functions and mutational mechanisms: A review. Mol. Ecol. 11:2453–2465.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. p. 281–297. *In* Proc. of the Berkeley Symposium on Mathematical Statistics and Probability, 5th. Vol. 1. University of California Press, Berkeley, CA.

Mogg, R., J. Batley, S. Hanley, D. Edwards, H. O'Sullivan, and K.J. Edwards. 2002. Characterization of the flanking regions of *Zea mays* microsatellites reveals a large number of useful sequence polymorphisms. Theor. Appl. Genet. 105:532–543.

Morgante, M., M. Hanafey, and W. Powell. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat. Genet. 30:194–200.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University press, New York.

Nevo, E. 2001. Evolution of genome-phenome diversity under environmental stresses. Proc. Natl. Acad. Sci. (USA) 98:6233–6240.

Parker, G.D., P.N. Fox, P. Langridge, K. Chalmers, B. Whan, and P.F. Ganter. 2002. Genetic diversity within Australian wheat breeding programs based on molecular and pedigree data. Euphytica 124:293–306.

Payne, T.S., B. Skovmand, C.G. Lopez, E. Brandon, and A. McNab. (ed.) 2002. The international wheat information system (IWIS™), Version 4. 2001. On compact disk. CIMMYT, Mexico, D.F.

Plaschke, J., M.W. Ganal, and M.S. Röder. 1995. Detection of genetic diversity in closely related brad wheat using microsatellite markers. Theor. Appl. Genet. 91:1001–1007.

Purdy, L.H., W.Q. Loegering, C.F. Konzak, C.J. Perterson, and R.E. Allan. 1968. A proposed standard method for illustrating pedigrees of small grain varieties. Crop Sci. 8:405–406.

Rajaram, S., and M. Van Ginkel. 2001. Mexico, 50 years of International Wheat Breeding. p. 579–608. *In* A.P. Bonjean and W.J. Angus (ed.) The world wheat book, a history of wheat breeding. Lavoisier Publishing, Paris.

Rajaram, S., M. Van Ginkel, and R.A. Fischer. 1994. CIMMYT's wheat breeding mega-envionments (ME). p. 1101–1106. *In* Z.S. Li and Z.Y. Xin (ed.) Proc. of the International Wheat Genetics Syposium, 8th, Beijing. 19–24 July. Agricultural Scientech Press, Beijing.

Röder, M.S., V. Korzun, K. Wendehake, J. Plaschke, M.-H. Tixier, P. Leroy, and M.W. Ganal. 1998. A microsatellite map of wheat. Genetics 149:2007–2023.

Rogers, J.S. 1972. Measures of similarities and genetic distances. p.145–153. *In* M.R. Wheeler (ed.) Studies in genetics VII. University of Texas, Austin, TX.

Saghai-Maroof, M.A., K. Soliman, R.A. Jorgensen, and R.W. Allard. 1984. Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc. Natl. Acad. Sci. (USA) 81:8014–8018.

Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin version 2.000: A software for population genetics data analysis. Genetics and biometry laboratory, University of Geneva, Switzerland (see also http://lgb.unige.ch/arlequin/; verified 22 September 2003).

Scott, K.D., P. Eggler, G. Seaton, M. Rossetto, E.M. Ablett, L.S. Lee, and R.J. Henry. 2000. Analysis of SSRs derived from grape ESTs. Theor. Appl. Genet. 100:723–726.

Song, Q.J., E.W. Fickus, and P.B. Cregan. 2002. Characterization of trinucleotide SSR motifs wheat. Theor. Appl. Genet. 104:286–293.

Stachel, M., T. Lelley, H. Grausgruber, and J. Vollmann. 2000. Application of microsatellites in wheat (*Triticum aestivum* L.) for studying genetic differentiation caused by selection for adaptation and use. Theor. Appl. Genet. 100:242–248.

St. Martin, S.K. 1982. Effective population size for the soybean improvement program in maturity groups 00 to IV. Crop Sci. 22:151–152.

Thornsberry, J.M., M.M. Goodman, J. Kresovich, S. Nielsen, and D. Buckler. 2001. Dwarf8 polymorphisms associate with variation in flowering time. Nat. Genet. 28:286–289.

Trethowan, R.M., J. Crossa, M. van Ginkel, and S. Rajaram. 2001. Relationships among bread wheat international yield testing location in dry areas. Crop Sci. 41:1461–1469.

Trethowan, R.M., M. van Ginkel, and S. Rajaram. 2002. Progress in breeding wheat for yield and adaptation in global drought affected environments. Crop Sci. 42:1441–1446.

Van Ginkel, M., R.M. Trethowan, K. Ammar, J. Wang, and M. Lillemo. 2002. A guide to bread wheat breeding at CIMMYT. Wheat special report No. 5 (Rev. ed.). CIMMYT, Mexico, D.F.

Weir, B.S. 1996. Genetic data analysis II. Sinauer Associates, Sunderland, MA.

Wright, S. 1978. Evolution and genetics of populations. University Chicago Press, Chicago.

Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheats and their backcross-derived lines using SSR markers

Pingzhi Zhang[1], S. Dreisigacker[2], A.E. Melchinger[2], J.C. Reif[2], A. Mujeeb Kazi[1], M. Van Ginkel[1], D. Hoisington[1] and M.L. Warburton[1,*]

[1]*International Maize and Wheat Improvement Center (CIMMYT, Int). Apdo Postal 6-641, 06600 Mexico, D.F., Mexio; [2]Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany; *Author for correspondence (E-mail: mwarburton@cgiar.org; Tel.: ++52 555 804 2004 ext. 1381; Fax: ++52 555 804 7567)*

Abstract

Synthetic hexaploid wheats (SHWs) and synthetic backcross-derived lines (SBLs) obtained from them are novel sources of useful traits for broadening the diversity in breeding germplasm of hexaploid bread wheat (*Triticum aestivum*). Fifty-one EST-derived and 39 genomic-derived microsatellite markers (SSRs) covering the A, B, and D genomes were used to assess the genetic diversity present in 11 SHWs, their backcross derived families, and their durum and bread wheat parents; and to test for the selective advantage of SHW alleles in SBL families after several generations of selection. The 90 SSR markers amplified 91 loci with 474 alleles across all genotypes. In many of the SHWs, novel alleles were observed which were stably inherited in the SBL families. Gene diversity, the average number of alleles per locus, cluster analysis, and principal coordinate analysis revealed a high level of genetic diversity in the *Aegilops tauschii* and durum parents of the SHWs, and also in the SBLs. In the latter, alleles from the SHW parent had a selective advantage for six SSR markers. This indicates that SHWs and SBLs are a valuable source for broadening the genetic base of elite wheat breeding germplasm. Fingerprinting of SBLs and their corresponding SHW and bread wheat parents, and testing for selective advantage of SHWs alleles promises to be a useful method for detecting chromosomal regions of interest for bread wheat improvement.

Introduction

Only a small proportion of the available genetic variation of the primary gene pool for most crop species has been exploited for crop improvement (Tanksley and McCouch 1997). Furthermore, modern plant breeding and intensive selection may reduce the genetic diversity in the elite germplasm pool, which can increase the risk of genetic vulnerability and threaten future breeding progress. A loss of diversity caused by plant breeding was demonstrated by genotyping recently released barley cultivars, their key progeni-

tors, and a group of 19 landraces with 28 simple sequence repeat (SSR) markers (Russell et al. 2000). A reduction in diversity during portions of the breeding process was also found in a survey of CIMMYT and CIMMYT-related bread wheat genotypes fingerprinted with 90 SSRs (unpublished data).

Introduction of valuable genes from exotic donors via wide crosses has been proposed to broaden the genetic base of many crop plants with known and closely related wild relatives, such as the diploid and tetraploid progenitors of bread wheat (BW) (*Triticum aestivum*; $2n = 6x = 42$, AABBDD) (Zohary et al.

1969). The wide crossing method is well suited for transfer of simply inherited traits. It is considerably more difficult for quantitative traits and only a few successful examples have been reported for grain yield (Villareal et al. 1995; Hoisington et al. 1999). One successful method for introducing variation from the progenitors of BW into the elite breeding germplasm is the creation of synthetic hexaploid wheats (SHWs). SHWs are produced by artificially crossing tetraploid forms such as modern durum wheat (*Triticum turgidum*, 2n=4x=28 AABB), donor of the A and B wheat genomes, with *A. tauschii* (2n=2x=14 DD), donor of the D genome. The resulting hybrid is haploid and carries the A, B, and D genomes (2n = 3x = 21, ABD). It is converted to a true hexaploid by using artificial chromosome doubling methods based on colchicine treatment (Mujeeb-Kazi et al. 1996).

Over a thousand new SHWs have been produced from more than 600 *A. tauschii* accessions at CIMMYT. SHWs possess favorable qualitative traits (Kema et al. 1995; Ma et al. 1995; Lage et al. 2001; Mujeeb-Kazi et al. 2001a, 2001b; Lage et al. 2002) and desirable quantitative traits (Villareal et al. 1994a, 1994b; Villareal et al. 2001; Rajaram 2001). However, in general, SHWs carry a large number of unfavorable alleles and are typically backcrossed to elite BW cultivars to produce agronomically acceptable synthetic backcross-derived lines (SBLs). Field trials have demonstrated that SBLs contain alleles contributing positive effects on yield components such as kernel weight (del Blanco et al. 2000, 2001; Gororo et al. 2002). The impact of SBLs is best illustrated by the fact that in 2003, about 25% of the wheat lines sent from CIMMYT to more than 100 collaborators in 50 countries for field trials in International Bread Wheat Screening Nurseries were SBLs (M. van Ginkel, unpublished data).

Molecular markers provide a powerful tool to assess the diversity within and among germplasm and to monitor the flux of diversity over time. In particular, SSRs show potential for large-scale DNA fingerprinting of wheat genotypes due to the high level of polymorphisms detected (Russell et al. 2000; Christiansen et al. 2002; Eujayl et al. 2002); their ability to be analyzed using automated systems (Sharon et al. 1997); and their high accuracy and repeatability (Heckenberger et al. 2002). Detailed knowledge about the molecular genetic diversity among SHWs, their parents, SBLs derived from them, and current elite BWs could promote the use of SHWs in breeding programs. Information about the flux of molecular diversity from SHWs to SBLs during subsequent backcrossing generations with BW could aid in detecting favorable chromosomal segments for important agronomic traits under selection. In a backcrossing program, only chromosomal regions carrying genes encoding the traits of interest, and closely linked markers to these genes, will be retained from the donor parent after several generations of backcrossing to the recurrent parent (with the possible exception of non-targeted introgressions remaining despite several generations of backcrossing; these should be minimal). These segments, including the genes and any DNA segments linked to these genes, will occur in the progeny in non-Mendelian proportions due to positive selection (Diaby and Casler 2003).

The objectives of this study were to (1) assess the genetic diversity within and among SHWs, their parents, SBLs, and BWs, and (2) test for a selective advantage of SHWs alleles in SBL families after several generations of selection during backcrossing and selfing.

Materials and methods

Plant materials

Four groups of wheat were fingerprinted in this study: (i) eight durum wheats (DWs), used as parents for the synthetic hexaploid crosses; (ii) 11 SHWs obtained by crossing the DW parents with *A. tauschii* accessions and doubling the chromosomes of the resulting hybrid plants; (iii) seven BWs, used as recurrent parents in backcrossing to the SHWs to generate the SBLs; and (iv) 15 SBL families (Table 1), with between 1 and 44 individuals per family, resulting in a total of 136 lines. Only 13 SBL families (105 lines) were considered in the final data analysis, because two SBL families showed an extremely high proportion (26.2% and 15.5%) of novel alleles (those not present in any of the parents; see explanation in the discussion). The *A. tauschii* genotypes used as parents for SHWs could not be analyzed individually because bulked pollen of heterogeneous *A. tauschii* accessions was used to pollinate the durum parents. Therefore, the *A. tauschii* genotypes were inferred using the fingerprint of the D genome of the SHWs.

Seeds of the 153 wheat genotypes were obtained from the CIMMYT Wheat Breeding Program or Plant Genetic Resources Center. The original synthetic

Table 1. Wheat germplasm fingerprinted with SSR markers: (i) durum wheats (DWs), (ii) synthetic hexaploid wheats (SHWs), (iii) recurrent bread wheat parents (BWs), and (iv) synthetic backcross-derived lines (SBLs) produced by backcrossing the SHW to BW for one (BC_1) or two (BC_2) generations. Novel alleles (NA) in the SHW not present in the DW parent, and alleles from the DW parent not recovered (NR) in the SHW, as well as the percentage of segregating loci (SL) in the SHW are given for the 60 SSR loci mapping to the A+B genomes. Selfing generation (SG) of SHW and backcross generations (BC) of SBL are also presented.

SHW						SG	SL	NA	NR	SBL			
Abbr.	Parents									BW Parents		Family	BC
	Abbr.	DW	Abbr.	A. tauschii/Origin						Abbr.	BW		
D1T1	D1	ALTAR 84	T1	219/Iran		S_4	7.8	6	2	B1	ESDA	A	BC_2
D1T1	D1	ALTAR 84	T1	219/Iran						B2	SERI	B	BC_1
D1T2	D1	ALTAR 84	T2	221/Iran		S_3	5.9	15		B4	BORL95	C	BC_2
D1T3	D1	ALTAR 84	T3	191/unknown		S_3	2.0	2.0		B6	YACO	D	BC_2
D2T4	D2	CROC_1	T4	205/China		S_4	3.9	6	3	B3	BCN	E	BC_1
D2T4	D2	CROC_1		205/China						B4	BORL95	F	BC_1
D2T5	D2	CROC_1	T5	224/unknown		S_5	5.9	23		B5	OPATA	G	BC_1
D3T6	D3	DVERD_2	T6	214/unknown		S_4	3.9	7	6	B3	BCN	H	BC_1
D4T7	D4	TK SN1081	T7	222/Iran		S_1	2.0	10	10	B3	BCN	I	BC_1
D5T8	D5	SORA	T8	323/Iran		S_1	2.0	9	9	B5	OPATA	J	BC_1
D6T9	D6	68.111/RGB-U/WARD	T9	325/Iran		S_1	5.9	24	22	B3	BCN	K	BC_1
D7T10	D7	CHEN	T10	429/Iran		S_1	5.9	13	12	B3 B7	BCN, KAUZ[1]	L	BC_1
D8T11	D8	CETA	T11	895/Iran		S_1	3.9	18	14	B3	BCN	M	BC_1

[1]BCN and KAUZ used to produce F_1, BC_1, respectively. Both are sister lines.

hexaploids were produced at CIMMYT in the Wheat Wide Crosses Program (Mujeeb-Kazi et al. 1996).

Marker analyses

Genomic DNA was extracted from bulked leaves harvested from 7 to 10 young plants according to Saghai-Maroof et al. (1984) and modified according to CIMMYT protocols (http:/www.cimmyt.cgiar.org/ ABC/Protocols/manual ABC.html). Quality and quantity of the extracted DNA was determined on 1% agarose gels by visually comparing extracted DNA bands to known concentration of a standard lambda DNA cut with *Hin*dIII. The SSR information was obtained from the Institut fur Pflanzengenetik und Kulturpflanzenforschung (IPK) (Gatersleben) and DuPont (Wilmington). In addition, the SSR markers Taglgap, Taglut (Devos et al. 1995), and *WMC56* developed by the Wheat Microsatellite Consortium (Agrogene, France) were used. Details for each of the 90 SSR (51 EST- and 39 genomic-derived) are available on the web at http://www.cimmyt.org/english/ webp/support/publications/support_materials/ ssr_mw1.htm.

PCR reactions for SSR analysis were performed according to Dreisigacker et al. (2004). Briefly, reactions containing a final volume of 20 ul were amplified in a 96-well Peltier Thermal cycler (MJ Research, Inc., Watertown, MA) following a standard temperature profile: 29 cycles consisting of 1 min denaturation at 94 °C, 2 min annealing using temperatures between 50 and 64°C (depending on primer combination), and 2 min extension at 72 °C. Primers were labeled with 6-FAM, HEX or TET fluorescent dyes. PCR products were amplified separately and run on an ABI Prism 377 DNA Sequencer (Perkin Elmer/ Applied Biosystems). When possible, multi-loading was applied to increase efficiency, with two to six loci with non-overlapping allele sizes loaded together in one well of the gel. Fragments were sized using Genescan 3.1 and assigned to allele categories using the software package Genotyper 2.1 (Perkin Elmer/ Applied Biosystems Biotechnologies, Foster City, USA). Forty-eight genotypes were run on each gel including two wheat lines as controls in every gel.

Statistical analyses

Alleles from the DW parents which were absent in the SHWs, novel alleles in the SHWs (from the A and B genomes but not contributed by the DW parents), and the percentage of segregating loci per polymorphic SSR in the SHWs were calculated for the 60 SSRs mapping to the A+B genomes. Gene diversity (H_S)

Table 2. Average number of alleles per locus (N_a), gene diversity within types (H_S) and proportion of polymorphic SSR loci (P) observed for the A+B, D genome and all 91 SSR loci in four types of wheat materials.

	A+B				D			All SSRs		
Statistic	DW[1]	SHW	BW	SBL	SHW	BW	SBL	SHW	BW	SBL
N_a	2.27	2.67	2.17	3.82	4.61	2.25	4.50	3.27	2.21	4.07
H_S	0.40	0.40	0.37	0.46	0.70	0.37	0.50	0.50	0.37	0.47
P	0.75	0.85	0.67	0.97	0.90	0.68	0.97	0.90	0.67	0.97

[1]DW: durum wheats; SHW: synthetic hexaploid wheats; BW: bread wheats; SBL: synthetic backcross-derived lines.

within DWs, SHWs, BWs, and SBLs were calculated for all loci according to Nei (1987). Modified Rogers' distances were calculated among the wheat groups or individuals (Wright 1987). Standard errors were estimated via bootstrapping over markers using 5000 repetitions. Cluster analysis with the unweighted paired group method with arithmetic averages (UPGMA) and principal coordinate analysis (PCoA; Gower 1966) were performed based on modified Rogers' distances.

Selective advantage of alleles was evaluated for each marker based on the parameter *p* of a binomial distribution by an exact test with Sidak correction (Dufner et al. 2002). Only BC[1] families G and H were used for the test because they consisted of more than 18 lines per family, whereas all other families had fewer than 12 lines and, consequently, the power of the test was too low. Monomorphic alleles, novel alleles, missing parental data and missing data in the SBLs were excluded from the calculation. Selective advantage of alleles in the SBLs was tested using the null hypothesis that the SHW allele does not have a selective advantage (H_o: $p \le 0.25$) versus the alternative hypothesis that the SHW allele has a selective advantage (H_A: $p > 0.25$). All analyses were carried out with version 2.0 of the Plabsim software (Frisch et al 2000), which is implemented as an extension of the statistical software R (Ihaka and Gentleman 1996).

Results

Genetic diversity within the four germplasm groups

The 90 SSR markers amplified a total of 91 loci, with 474 alleles across all genotypes. One marker (*DuPw23*) was monomorphic in all genotypes, while the most polymorphic marker (*Xgdm98*) generated 13

alleles. The proportion of polymorphic SSR loci, H_S and the average number of alleles per locus (N_a) of the A, B and D genomes were higher for SHWs and SBLs than for BWs (Table 2). H^S and N^a were higher for the D genome than for the A+B genome, except in the BWs. UPGMA cluster analyses, separately performed for the SSRs on the A+B and D genomes, revealed that both parents of the SHWs were highly diverse (Figure 1, Figure 2).

Genetic diversity among the four germplasm groups

The modified Rogers' distances between groups were high except for between D>s and SHWs and between SBLs and BW (Table 3). UPGMA cluster analysis of DWs and SHWs, based on SSRs mapping to the A+B genomes, showed that SHWs and their corresponding DW parents clustered closely together in all cases except entry D2T5 (Figure 2). PCoA based on all 90 SSRs of all genotypes except the DWs revealed that the SHWs clustered separately from the SBL families and BWs (Figure 3). Members of the same SBL family clustered together, (with the exception of one individual from family C) and were closely associated with their corresponding BW parents.

The SHWs and their corresponding DW parents should have had no allele differences for the A and B genomes, but differences caused by novel alleles present in the SHWs but absent in DW, and by alleles from the DWs that were not recovered in the SHWs (Table 1) caused a low genetic distance between the two groups. The number of alleles from the DW parents absent in the SHWs ranged from 2 to 22 per individual and averaged 9.8. The number of novel alleles in the SHWs absent in the DW parents ranged from 6 to 24 per individual, with an average of 13.7. For the 105 SBLs included in the final analysis, three lines had no novel alleles, 87 lines contained 1 to 12
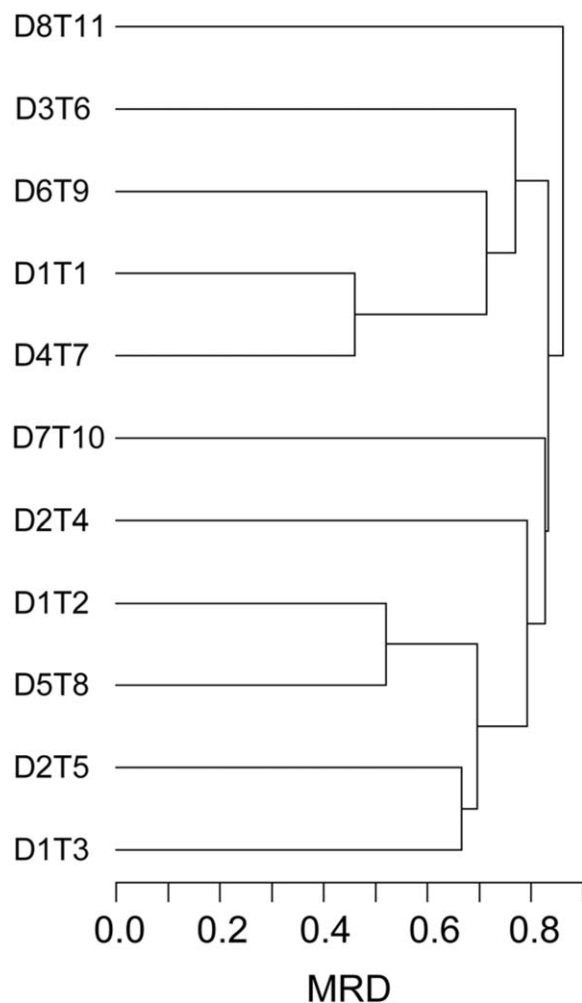
Figure 1. UPGMA cluster of the synthetic hexaploid wheats based on modified Rogers' distances. The 28 SSRs mapping to the D genome were used for the analysis (see Table 1 for abbreviations of synthetic hexaploid wheats).

Figure 2. UPGMA cluster of durum wheat parents and synthetic hexaploid wheats (SHWs) based on modified Rogers' distance (MRD). The 60 SSRs mapping to the A and B genomes were used for the analysis (see Table 1 for abbreviations of durum wheat parents and synthetic hexaploid wheats).

novel alleles and 15 lines possessed 13 to 26 novel alleles.

Selective advantage of SHWs alleles inherited by the SBLs

Allele frequencies in SBL families G and H were consistent with the null hypothesis at most SSRs, as they showed either expected Mendelian ratios or more BW alleles than expected (Table 4). However, four SSRs in family G and five SSRs in family H led to a rejection of the null hypothesis and showed a selective advantage in favor of alleles from the SHWs.
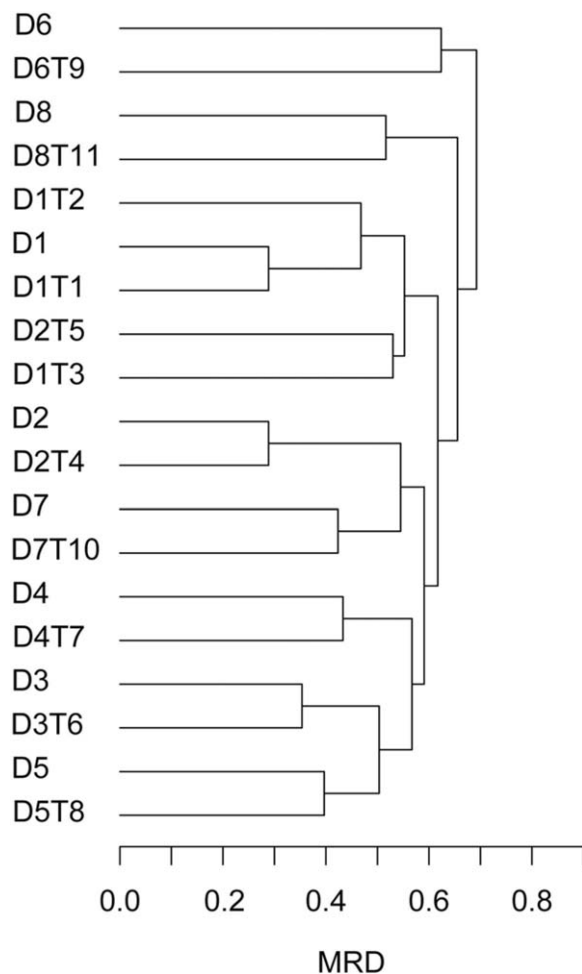
Discussion

Bread wheat was domesticated 12,000 years ago in the Fertile Crescent (Salamini et al. 2002). Unlike the other major grains, bread wheat is the result of two consecutive hybridizations between species and therefore the size of the founder population of bread wheat was most likely limited, causing a domestication bottleneck (Dvorak et al. 1988). Following domestication, genetic variation may have been further reduced due to selection by early farmers and current plant breeding efforts. CIMMYT's germplasm bank maintains the largest collection of wheat genetic resources in the world, consisting of more than 150,000

Table 3. Mean modified Rogers' distances (above diagonal) and associated standard errors (below diagonal) between four types of wheat germplasm based on SSR allele frequencies for the wheat genomes A+B and D as well as all 91 SSR loci.

Materials[1]	A+B				D			All SSRs		
	DW	SHW	BW	SBL	SHW	BW	SBL	SHW	BW	SBL
DW		0.189	0.652	0.522	–	–	–	–	–	–
SHW	0.043[2]		0.615	0.481		0.560	0.435		0.603	0.472
BW	0.036	0.039		0.203	0.054		0.216	0.036		0.207
SBL	0.030	0.035	0.036		0.053	0.039		0.035	0.034	

[1]For abbreviations, see footnote Table 2; [2]Standard errors were calculated by bootstrapping using 5000 repetitions.
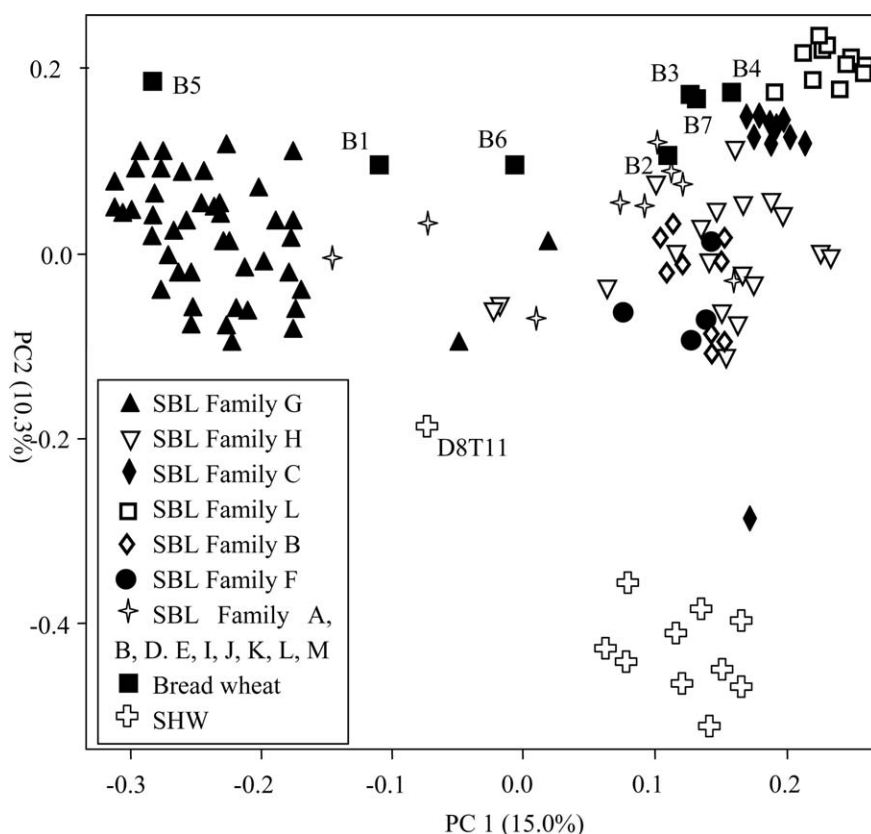

Figure 3. Principal coordinate analysis of synthetic backcross-derived lines, their recurrent bread wheat and synthetic hexaploid wheat (SHW) parents based on modified Rogers' distances calculated from all 90 SSRs. PC1 and PC2 are the first and second principal coordinates, respectively (see Table 1 for abbreviations and details).

accessions of landraces and wild relatives. CIMMYT therefore holds a key position in conserving the genetic diversity of wheat but also in making it available for breeding. The creation of SHWs and SBLs presents a promising method to unlock the diversity conserved in the wild relatives of wheat.

Genomic changes in SHWs and SBLs

We detected alleles from the SHW parents that were not recovered in the SHWs, some of which could be explained by a certain level of residual polymorphism within the DW parents. Although wheat cultivars are produced via self-pollination, a low percentage of outcrossing does occur in wheat. However, we also observed a high percentage of novel alleles in the

Table 4. SSR markers showing significant (P < 0.05) selective advantage for the synthetic hexaploid wheats (SHW) alleles in synthetic backcross-derived line (SBL) families G and H. Furthermore, chromosomal position of the markers, number of alleles transmitted from SHW or bread wheat (BW) parents are given.

SSR marker	Chr.	Family G			Family H		
		SHW	BW	Significant	SHW	BW	Significant
*DuPw*532	1A	17	29	No			
Taglgap	1B				12	4	Yes
*DuPw*431	1D				9	9	No
*DuPw*344	2B	17	27	No			
*Xgwm*705	3B	22	23	Yes	19	0	Yes
*Xgwm*161	3D	17	27	No			
*DuPw*122	3D	19	24	No			
*DuPw*221	3D				8	9	No
*DuPw*41	4D	28	16	Yes	12	7	Yes
*DuPw*278	4D	29	16	Yes	14	7	Yes
*DuPw*395	5A	21	24	No			
*DuPw*402	5A				11	8	No
*DuPw*137	6A	15	24	No	10	8	No
*DuPw*145	7B	17	26	No			
*DuPw*533	7B				17	0	Yes
*Xgwm*577	7B	26	18	Yes			

SHWs, which were not detected in their corresponding DW and BW parents and which were stably inherited in the SBLs (Table 1). These observations are in agreement with previous reports in *Triticum* and *Brassica* based on other marker types. Sasakuma et al. (1995) found in the ribosomal DNA spacer of synthetic amphiploids (AABBDD) hybridization alleles in addition to those inherited from their two parents. In newly synthesized amphiploids of *Triticum* and *Aegilops*, Liu et al. (1998a, 1998b) observed a rapid, high frequency elimination of low-copy DNA sequences. Likewise, Song et al. (1995) detected genomic changes using RFLPs in *Brassica* synthetic polyploids of *B. rapa* (A genome), *B. nigra* (B genome), and *B. oleracea* (C genome) and their reciprocal interspecific hybridizations.

Possible explanations for the genomic changes in our study are (i) genetic instabilities in the newly synthesized SHWs, (ii) seed handling errors, (iii) outcrossing and (iv) experimental errors during the laboratory assays for SSRs. The genetic instabilities in the new polyploids could have resulted from several different processes, such as chromosome rearrangements or mutations. Seed handling errors are unlikely in all but one case, since UPGMA cluster analysis revealed that the SHWs and their corresponding DW parents clustered very closely together 'except entry D2T5, Figure 2). D2T5 clustered close to D1, although its corresponding DW parent was D2. This may be due to the extremely high number (23) of novel alleles in D2T5, which made it appear more similar to D1 simply by coincidence (identical in state rather than by descent), but the occurrence of a simple seed handling error cannot be excluded in this case, either. Another possible explanation lies in partial male sterility which often occurs in the newly synthesized amphiploids, which increases the outcrossing rate in some of the SHWs, resulting in alleles actually coming from nearby BWs (M. van Ginkel, unpublished data). However, novel alleles due to outcrossing would lead to an increased heterozygosity in the SHWs, but only a low residual heterozygosity of 4.3% was observed in the SHWs. Consequently, outcrossing seems to be a rather unlikely explanation for the novel alleles. The error rate of misclassification of SSR alleles, estimated with the two controls, was on average less than 5%, whereas the average percentage of novel alleles detected in SHWs was 16%. Thus, experimental errors during the laboratory assays could be at best only a partial explanation for the genomic changes in the SHWs. It would therefore appear that genetic instabilities in the SHWs are the major cause of the novel alleles seen in this study. Further investigation, such as sequencing the novel SSR alleles, would be required to verify these results and to characterize the genomic changes in more detail.

Genetic diversity within and among SHWs, SBLs, DWs, and BWs

The structure of the UPGMA cluster of the SHWs, based on the D genome markers (Figure 1), can be explained by the pedigree information on the *A. tauschii* parents (Table 1). The two SHW entries D1T2 and D5T8 were derived from the same original *A. tauschii* accession. The *A. tauschii* parents of D1T1 and D4T7 originated from geographically very close regions in Iran. The origins of the *A. tauschii* parents of D1T3 and D2T5 are unknown, but since they clustered closely together, one may assume a similar geographic origin.

H^S for the D genome markers in the SHWs was high (0.70) and similar to the gene diversity ($H_S = 0.68$) reported in a previous study of 60 accessions of *A. tauschii* fingerprinted with 14 SSRs (Lelley et al. 2000). Average number of alleles of the D genome SSRs for the *A. tauschii* parents (4.6) was also higher than any other sample in this study, but was smaller than N_a values reported for genomic-derived SSRs in surveys by Lelley et al. (2000) (6.5) and Pestsova et al. (2000) (18.8). The differences may be due to a smaller sample size of *A. tauschii* in our study as well as by the fact that the EST-derived SSRs used in our study generally show a much lower degree of polymorphism than do the genomic-derived SSRs (Eujayl et al. 2002). The UPGMA cluster of the DW parents of the SHWs based on SSRs from the A+B genome shown in Figure 2, as well as N_a and N_s values (Table 2) indicates a high level of genetic diversity in this group of wheats as well.

PCoA revealed a clear separation of the SHWs from the BWs, with the exception of the SHW line D8T11, which clustered more closely to the BWs (Figure 3). This result is also reflected in the high genetic distance between these two groups (Table 3). It confirms that the SHWs form an entirely different pool of genetic variation than the CIMMYT elite BW recurrent parents and thus represent an interesting source for broadening the genetic base of the BW germplasm for certain traits. Nevertheless, they also contain many undesirable characteristics such as low yield and lack of uniformity for flowering time. Therefore, plants displaying unfavorable traits must be eliminated in a selection program during the development of SBLs.

The SBLs were clearly separated from the SHWs by the second principal coordinate (PC2), which explained 10.3% of the total variation (Figure 3). In general, BW parents grouped most closely to their corresponding SBL families, suggesting that the recurrent BW parents influenced the genetic constitution of the SBLs more than the SHWs did, as would be expected in a backcrossing program. Nevertheless, N_a, H_S, and the proportion of polymorphic SSR loci were higher for SBLs than for BWs (Table 2), indicating that SBLs have also received new genetic variation from their SHW parents. The usefulness of SBLs for increasing the genetic diversity in BW was also supported by the modified Rogers' distance between these two germplasm groups (MRD = 0.2, Table 3). Thus, some diversity from the SHWs is lost during the backcrossing process but considerable variation is transferred to the BWs. Some of this variation will provide novel alleles of traits of interest to wheat breeders which have not been tapped in the past, as this is the first time it has been present in the primary gene pool of wheat.

Selective Advantage of SHWs alleles

Due to selection during backcrossing in the formation of the SBLs, allele frequencies in the SBL families are expected to deviate from expected segregation ratios if the SSRs are linked to chromosomal regions controlling the traits under selection. The SBL families were grown in nurseries where they were exposed to biotic and abiotic stresses including (among others) karnal bunt, fusarium head blight and low levels of irrigation. Significant deviations of the expected allele frequencies in the direction of SHW parents can, therefore, be used to detect chromosomal regions of interest coming from the SHWs. A precondition to map these chromosomal regions is an adequately large family size to provide a sufficient power of the test. For that reason, we analyzed only families G and H with 44 and 19 individuals, respectively.

Most of the SSRs in SBL families G and H confirmed the null hypothesis, i.e., alleles did not deviate from expected ratios or the BW alleles had a selective advantage over SHW alleles. Consequently, more BW alleles were found than expected by Mendelian inheritance (Table 4) and can be explained by a period of approximately 12,000 years of adaptation and selection of BW for agronomically favorable alleles. In addition, chromosomal blocks are selected rather than individual genes, because of the limited number of meiosis between the BW and SHW haplotypes. Reduced recombination may also cause unfavorable alleles to be retained in some linkage blocks. Conse-

quently, the net effect of all genes linked in a chromosomal block has to be considered. It is very likely that the sum of genetic effects of such a chromosomal block from SHWs is not significantly higher than the corresponding sum from the BW black, except for SHW chromosomal blocks with major positive effects.

The null hypothesis was rejected for four SSRs in family G and five in family H, indicating a selective advantage in favor al alleles from the SHWs (Table 4). Selective advantage in favor of SHWs alleles was found simultaneously in both families for *Xgwm705*, *DuPw41* and *DuPw278*. Two of the three markers, *DuPw41* and *DuPw278*, map to the D genome, indicating a positive effect of traits from the *A. tauschii* parent. Marker *Xgwm705* indicates a linkage to a favourable trait coming from the DW parent of the SHWs. Furthermore, the SHW alleles of *Xgwm577* in family G and the SHW allele of *DuPw533* in family H showed a selective advantage and both these markers map to chromosome 7B, again indicating an association in this region with desired traits from the DW parents. Consequently, fingerprinting SBLs and their corresponding SHW and BW parents and testing for selection advantage of SHWs over BWs alleles seems to be a promising method to detect chromosomal regions of interest for bread wheat improvement. Narrowing these genomic regions and associating them with a trait of interest in the progeny is the next step in the identification of useful functional diversity from the progenitors of bread wheat.

Acknowledgements

References

Christiansen M.J., Anderson S.B. and Ortiz R. 2002. Diversity changes in an intensively bread wheat germplasm during the 20th century. Mol. Breed. 9: 1–11.

Del Blanco I.A., Rajaram S., Kronstad W.E. and Reynolds M.P. 2000. Physiological performance of synthetic hexaploid wheat-derived populations. Crop Sci. 40: 1257–1263.

Del Blanco I.A., Rajaram S. and Kronstad W.E. 2001. Agronomic potential of synthetic hexaploid wheat-derived populations. Crop Sci 41: 670–676.

Devos K.M., Bryan G.J., Stephenson P. and Gale M.D. 1995. Application of two microsatellite sequences in wheat storage proteins as molecular markers. Theor. Appl. Genet. 90: 247–252.

Diaby M. and Casler M.D. 2003. RAPD marker variation among smooth bromegrass cultivars. Crop Sci. 43: 1538–1547.

Dreisigacker S., Zhang P., Warburton M.L., van Ginkel M., Hoisington D., Bohn M., Melchinger A.E. 2004. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different mega-environments. Crop Sci. 44: 381–388.

Dufner J., Jensen U. and Schumacher E. 2002. Statistik mit SAS. B.G. Teubner, Stuttgart, Germany, pp. 212.

Dvorak J., Luo M.C., Yang Z.L. and Zhang H.B. 1998. The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. Theor. Appl. Genet. 97: 657–670.

Eujayl I., Sorrells M., Baum M., Wolters P and Powell W. 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor. Appl. Genet. 104: 399–407.

Eujayl I., Sorrells M., Baum M., Wolters P. and Powell W. 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor. Appl. Genet. 104: 399–407.

Frisch M.M., Bohn M. and Melchinger A.E. 2000. Plabsim: Software for simulation of marker-assisted backcrossing. J. of Heredity 91: 86–87.

Gororo N.N., Eagles H.A., Eastwood R.F., Nicolas M.E. and Flood R.G. 2002. Use of *Triticum tauschii* to improve yield of wheat in low-yielding environments. Euphytica 123(2): 241–254.

Gower J.C. 1966. Some distance properties of latent root and in multivariate analysis. Biometricka 53: 325–388.

Heckenberg M., Melchinger A.E., Ziegle J.S., Joe L.K., Hauser J.D., Hutton M. and Bohn M. 2002. Variation of DNA fingerprints among accessions within maize inbred lines with regard to the identification of essentially derived varieties. I. Genetic and technical sources of variation in SSR data. Mol. Breed. 10: 181–191.

Hoisington D., Khairallah M., Reeves T., Ribaut J.M., Skovmand B., Taba S. and Warburton M. 1999. Plant genetic resources: what can they contribute toward increased crop productivity? Proc. Natl. Acid. Sci. USA 96: 5937–5943.

Ihaka R. and Gentleman R. 1996. A language for data analysis and graphics. J. of Computational and Graphical Statistics 5: 299–314.

Kema G.H.J., Lange W. and Silfhout C.H. 1995. Differential suppression of stripe rust resistance in synthetic wheat hexaploids derived from *Triticum turgidum* subsp. *dicoccoides* and *Aegilops squarrosa*. Phytopathology 85: 425–429.

Lage J., Skovmand B. and Andersen S.B. 2001. Field evaluation of Russian wheat aphid (*Homoptera: Aphididae*) resistance in

Triticum dicoccum derived synthetic hexaploid wheats. Abstract, Annual meeting, ESA, San Diego, USA, December 2001.

Lage J., Skovmand B. and Andersen S.B. 2002. Expression and suppression of resistance to greenbug (*Homoptera: Aphididae*) in synthetic hexaploid wheats derived from *Triticum dicoccum x Aegilops tauschii* crosses. J. of Economic Entomology 96: 202–206.

Lelley T., Stachel M., Grausgruber H. and Vollmann J. 2000. Analysis of relationships between *Aegilops tauschii* and the D genome of wheat utilizing microsatellites. Genome 43: 661–668.

Liu B., Vega J.M., Segal G., Abbo S., Rodova M. and Feldman M. 1998a. Rapid genomic changes in newly synthesized polyploids of *Triticum* and *Aegilops*. I. Changes in low-copy noncoding DNA sequences. Genome 41: 272–277.

Liu B., Vega J.M. and Feldman M. 1998b. Rapid genomic changes in newly synthesized polyploids of *Triticum* and *Aegilops*. II. Changes in low-copy coding DNA sequences. Genome 41: 535–542.

Ma H., Singh R.P. and Mujeeb-Kazi A. 1995. Resistance to stripe rust in *Triticum turgidum*, *A. tauschii* and their synthetic hexaploids. Euphytica 82: 117–124.

Mujeeb-Kazi A., Rosas V. and Roldan S. 1996. Conservation of the genetic variation of *Triticum tauschii* (Coss.) Schmalh. (*Aegilops squarrosa* auct. non L.) in synthetic hexaploid wheats (*T. turgidum* L. x *A. tauschii*; 2n = 6x =42, AABBDD) and its potential utilization for wheat improvement. Genet. Res. Crop. Evol. 43: 129–134.

Mujeeb-Kazi A., Fuentes-Davila G., Villareal R.L., Cortes A., Roasas V. and Delgado R. 2001a. Registration of 10 synthetic hexaploid wheat and six bread wheat germplasms resistant to karnal bunt. Crop Sci. 41: 1652–1653.

Mujeeb-Kazi A., Gilchrist L.I., Villareal R.L. and Delgado R. 2001b. Registration of ten wheat germplasms resistant to *Septoria tritici* leaf blotch. Crop Sci 40: 590–591.

Nei M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, USA, pp. 164.

Nei M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, USA, . pp. 164.

Pestsova E., Korzun V., Goncharov N.P., Hammer K., Ganal M.W. and Röder M.S. 2000. Microsatellite analysis of *Aegilops tauschii* germplasm. Theor. Appl. Genet. 101: 100–106.

Rajaram S. 2001. The human right to food and livelihoods: the role of global wheat research. Presented at the national seminar, building on past achievements to advance the future direction of Australian agriculture, ASTE Crawford Fund and Australian Institute of Agricultural Science and Technology, 2001, University House, Canberra, Australia.

Russell J.R., Ellis R.P., Thomas W.T.B., Waugh R., Provan J., Booth A., Fuller J., Lawrence P., Young G. and Powell W. 2000. A retrospective analysis of spring barley germplasm development from ′foundation genotypes′ to currently successful cultivars. Mol. Breed. 6: 553–568.

Saghai-Maroof M.A., Soliman K., Jorgensen R.A. and Allard R.W. 1984. Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc. Natl. Acad. Sci. USA 81: 8014–8018.

Salamini F., Özkan H., Brandolini A., Schäfer-Pregl R. and Martin W. 2002. Genetics and geography of wild cereal domestication in the Near East. Nature Rev. Genet. 3: 429–441.

Sasakuma T., Ogihara Y. and Tsujimoto H. 1995. Genome rearrangement of repetitive sequence in processes of hybridization and amphiploidization in Triticinae.. In: Li Z.S. and Xin Z.Y. (eds), Proc. of the 8th Intl. Wheat Genet. Symp. Held at Beijing, China, 20–25 July 1993. China Agricultural Scientech Press, Beijing, China, pp. 563–566.

Sharon E.M., Kresovich S., Jester C.A., Hernandez C.J. and Szewc-McFadden A.K. 1997. Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop Sci. 37: 617–624.

Song K.M., Lu P., Tang K.L. and Osborn T.C. 1995. Rapid genome changes in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc. Natl. Acad. Sci. USA 92: 7719–7723.

Tanksley S.D. and McCouch S.R. 1997. Seed banks and molecular maps: Unblocking genetic potential from the wild. Science 277: 1063–1066.

Villareal R.L., Mujeeb-Kazi A., Del Toto E., Crossa J. and Rajaram S. 1994a. Agronomic variability in selected Triticum *turgidum* x *A. tauschii* synthetic hexaploid wheats. J. Agron and Crop Sci. 173: 307–317.

Villareal R.L., Mujeeb-Kazi A., Rajaram S. and Del Toto E. 1994b. Morphological variability in some synthetic hexaploid wheats derived from *Triticum turgidum* x *A. tauschii*. J. Genet. and Breed. 48: 7–16.

Villareal R.L., Del Toro E., Mujeeb-Kazi A. and Rajaram S. 1995. The 1BL/1RS chromosome translocation effect on yield characteristics in a *Triticum aestivum* L. cross. Plant Breed. 114: 497–500.

Villareal R.L., Sayre K., Banuelos O. and Mujeeb-Kazi A. 2001. Registration of four synthetic hexaploid wheat (*Triticum turgidum/Aegilops tauschii*) germplasm lines tolerant to waterlogging. Crop Sci. 41: 274.

Wright S. 1978. Evolution and the genetics of populations. University Chicago Press, Chicago III, USA.

Zohary D., Harlan J.R. and Vardi A. 1969. The wild diploid progenitors of wheat and their breeding value. Euphytica 18: 58–65.

Genetic Diversity among and within CIMMYT Wheat Landrace Accessions Investigated with SSRs and Implications for Plant Genetic Resources Management

S. Dreisigacker, P. Zhang, M. L. Warburton, B. Skovmand, D. Hoisington, and A. E. Melchinger*

ABSTRACT

Many wheat (*Triticum aestivum* L.) landrace cultivars (LCs) conserved in seed banks are not sufficiently characterized to inspire breeders' interest for their efficient exploitation. Patterns of genetic variation within and among wheat LCs are usually unknown. Two sets of wheat LCs stored in CIMMYT's plant genetic resources center were assessed for genetic diversity by means of 76 (Set 1) and 44 simple sequence repeat (SSR) markers (Set 2). Set 1 included 36 LC accessions originating from different countries, either collected as bulks, composed of a single LC subline, or an unknown collection method. Set 2 consisted of three to 25 sublines of five Mexican and four Turkish LCs already included in Set 1. In a principal coordinate analysis based on modified Rogers' distance (MRD), only three Turkish LC accessions formed a distinct cluster in Set 1. The Mexican accessions clustered together with a Spanish accession and a close relationship between a Chilean and Nigerian accession was observed. In Set 2, gene diversity (H_e) among the Turkish LCs (0.43) was higher than among the Mexican LCs (0.35). Analyses of molecular variance (AMOVA) revealed considerable genetic diversity within Mexican (52.7%) and within Turkish (67.6%) LCs. Pairwise fixation indices (F_{ST}) were significant, except between two Turkish LCs. Results were discussed in relation to the most suitable collection method of wheat LCs (bulk or individual sublines) as well as to the use of SSRs as a tool for seed bank management.

Wheat landraces are genetically diverse and dynamic populations but are still morphologically recognizable because of a certain integrity (Harlan, 1975). Thousands of landrace cultivars (LCs) in wheat are stored in seed banks worldwide but the majority is inadequately described for an efficient exploitation in plant breeding. High costs and time-lags associated with the extensive search for useful characteristics lead to the fact that breeders rarely resort to these genetic resources (Gollin et al., 2000). Subsequently, intensive prebreeding approaches are required to transfer desired genes from an unimproved LC material into advanced breeding lines (Skovmand and Rajaram, 1990).

Landrace cultivars undoubtedly represent an important source of genetic variation in wheat. One of the prime examples is the use of *Rht* dwarfing genes that became available through the Japanese wheat 'Norin 10', derived from the LC Shiro Daruma (Kihara, 1982). Two important genes, *Rht1* and *Rht2*, were observed to directly effect yield because of reduced lodging. Moreover, a considerable LC diversity was found for resistance to pests such as stem rust (caused by *Puccinia graminis* Pers.: Pers. f. sp. *tritici* Eriks. & E. Henn.), leaf rust [*P. recondita* Roberge ex Desmaz. f. sp. *tritici* (Eriks. & E. Henn.) D.M. Henderson], or Russian wheat aphid (*Diuraphis noxia* Mordv.) (Skovmand and Rajaram, 1990; Skovmand et al., 1994), and for tolerance to abiotic stresses like heat (Hede et al., 1999; Skovmand et al., 2001).

With a few exceptions, all evaluations for desired traits in wheat LCs were done in ex situ collections. Examinations included either a random bulk of LC genotypes or the collections of LC sublines. Preliminary evaluation data were usually recorded during the first seed multiplication and consisted of observations that were highly heritable, easily detectable, and expressed in different environments (DeLacy et al., 2000). However, little information about the genetic variation within LCs and associations among LC accessions is available. It is also still questionable which strategy is the best to ensure an appropriate maintenance of this variation for future generations.

Molecular markers can support a more detailed characterization of genetic resources. A vast potential lies in their ability to identify the structure of genetic diversity within and among accessions, which can be of great importance for the optimization of collections, the planning of seed regeneration, and the successful implementation of prebreeding approaches. Molecular markers provide a direct measure of genetic diversity and go beyond indirect diversity measures based on agronomic traits or geographic origin. Simple sequence repeats are highly polymorphic in wheat and, therefore, suitable for the discrimination of genotypes. They are generally genome specific, abundant, codominant, and cover all 21 wheat chromosomes. They have been successfully employed to characterize genetic diversity in seed bank collections of improved wheat germplasm (Börner et al., 2000; Huang et al., 2002) and wild relatives (Li et al., 2000; Hammer, 2000).

The objectives of our study were to (i) determine SSR-based genetic diversity among and within two sets of hexaploid wheat LCs stored in the plant genetic resources center of CIMMYT, (ii) compare the form of conservation in bulks and individual plant collections, which were applied to maintain these LCs, and (iii) evaluate the use of SSRs as a tool to improve the management of wheat genetic resources.

Abbreviations: AMOVA, analysis of molecular variance; F_{ST}, pairwise fixation indices; H_e, gene diversity; LC, landrace cultivar; MB, multiple bands; MRD, modified Rogers' distance; PC, principal coordinate; PCoA, principal coordinate analysis; SSR, simple sequence repeat.

Table 1. Name, CIMMYT accession number, country of origin, form and year of collection, and storage of 36 landrace cultivar (LC) accessions of wheat.

Number	LC	CIMMYT accession	Country	Collection form	Collection year	Year of storage†
			Africa			
1	Tchere	CWI32617	Chad	unknown	unknown	1990
2	Alkana	CWI32616	Chad	unknown	unknown	1990
3	Aethiopicum 1B.18.16	CWI21968	Ethiopia	unknown	unknown	1987
4	Aethiopicum 400	CWI21966	Ethiopia	unknown	unknown	1987
5	Abyssinia 1	CWI9819	Ethiopia	unknown	1962	1989
6	Dikwa 1	CWI74748	Nigeria	bulk	1970s	1999
			Asia			
7	Pissi Khawri	CWI65257	India	LC subline	unknown	1990
8	LPG 1	CWI28879	Nepal	bulk	1950s	1990
9	Kharkovskaya 2	CWI51805	Russia	unknown	1989	1993
10	Shorewaki	BW20313	Pakistan	unknown	unknown	1995
11	86PK1317	CWI28659	Pakistan	LC subline	1986	1990
12	86PK1271	CWI28683	Pakistan	LC subline	1986	1990
			West Europe			
13	Gentil Bianco	CWI42611	Italy	unknown	unknown	1992
14	Barbela	CWI10618	Portugal	unknown	1967	1989
15	Barbela 0248	CWI7874	Spain	unknown	unknown	1990
16	Cologgne Abastrado 11660	CWI17849	Spain	unknown	unknown	1990
17	Barbilla	CWI17538	Spain	unknown	unknown	1990
18	Blanquillo-de-Badajoz	CWI17542	Spain	unknown	unknown	1990
			Turkey			
19	AK Bugday	CWI11215	Turkey	unknown	1969	1989
20	Yayla 305	CWI41983	Turkey	bulk	1985	1992
21	Yilmaz 1	CWI32653	Turkey	LC subline	1985	1991
22	Yilmaz 11	CWI32659	Turkey	LC subline	1985	1991
23	AK 702	CWI11164	Turkey	bulk	unknown	1989
24	84TK520.001.01	CWI28416	Turkey	LC subline	1984	1984
25	84TK523.006.02	CWI28421	Turkey	LC subline	1984	1984
26	84TK538.002.02	CWI28427	Turkey	LC subline	1984	1984
27	84TK567.001	CWI28013	Turkey	LC subline	1984	1984
			Central America			
28	Pillon	CWI31398	Mexico	LC subline	1990	1990
29	Barbon	CWI31424	Mexico	LC subline	1990	1990
30	Quartito	CWI31470	Mexico	LC subline	1990	1990
31	Caña Morado	CWI31499	Mexico	LC subline	1990	1990
32	Tzumutaro	CWI31604	Mexico	LC subline	1990	1990
33	Crillo GTM National V	CWI74755	Guatemala	unknown	unknown	
			South America			
34	Trigo Blanco	CWI59547	Chile	unknown	unknown	1995
35	Trigo Africano	CWI12244	Chile	unknown	unknown	1989
36	Trigo Azul	CWI27062	Chile	unknown	unknown	1990

† Year since the accessions were placed in CIMMYT's plant genetic resources center for storage.

MATERIALS AND METHODS

Plant Materials

The collection and maintenance of wheat LCs in seed banks is conducted either in bulks or as individual plant collections. Bulks are usually created as a random sample of spikes per LC, harvested and threshed together in one bag. Individual plant collections are composed of a number of LC sublines, whose seeds are kept separately.

Two sets of germplasm were used to analyze the genetic variation of hexaploid wheat LCs stored in CIMMYT's plant genetic resource center. Set 1 included 36 LCs accessions, either collected as a bulk, composed of a single LC subline, or of an unknown collection method (Table 1). Set 2 consisted of supplementary individual plant collections of five Mexican and four Turkish LCs already included in Set 1 (refer to Table 2 for the names and available number of sublines per LC).

The LC accessions in Set 1 were chosen because they expressed several characteristics of particular interest to breeders (e.g., salt tolerance, zinc, or flooding tolerance). The individual plant collections of the Mexican LCs in Set 2 were collected by B. Skovmand, in Michoacan, Mexico in 1989 in cooperation with the Mexican Organization for the Study of Biodiversity (Skovmand et al., 1992). The collections were performed within the framework of a larger collection mission at 219 Mexican sites. It was assumed that the LCs, still commercially grown at the time of collection, were introduced from Spain in about 1550. The individual plant collections of the Turkish LCs in Set 2 were collected in 1984 by R. Metzger, together with researchers from the Turkish Ministry of Agriculture. Collection sites were located in the mountain regions of Hakkari, in southeast Turkey (Skovmand et al., 1994). Since the beginning of their storage at CIMMYT, all wheat LC accessions have been regenerated once, by sowing 100 seeds per accession.

SSR Analyses

Genomic DNA of each LC accession in Set 1 was extracted from fresh leaves of 10 to 12 randomly selected seedlings by a modified CTAB (cetyltrimethylammonium bromide) method (Hoisington et al., 1994). Quality and quantity of the isolated DNA was determined on 1% (w/v) agarose gels by comparing bands to known concentrations of lambda DNA. Equal quantities of eight DNA samples per LC accession were bulked together. For Set 2, genomic DNA was extracted from each LC subline, applying the same method.

Table 2. Number of sublines per landrace cultivar (LC), average number of alleles per locus, percentage of heterozygosity, number of unique alleles, monomorphic loci, and gene diversity (H_e) in each of five Mexican and four Turkish LCs.

LC	Number of sublines	Average number of alleles per locus†	Heterozygosity	Unique alleles†	Monomorphic loci	H_e
			Mexican LCs			
Pillon	20	1.9	2.6	27.3	15	0.31
Barbon	24	2.1	4.9	24.5	10	0.37
Quartito	17	1.9	1.6	37.0	12	0.27
Caña Morado	25	2.3	1.9	41.3	8	0.41
Tzumutaro	13	2.3	1.7	45.9	14	0.41
Total/mean	99	4.6	2.5	–	5	0.35
			Turkish LCs			
84TK523.006.02	4	1.4	2.5	46.9	20	0.49
84TK538.002.02	7	1.4	5.0	33.5	10	0.55
84TK567.001	6	1.4	3.1	19.8	14	0.44
84TK567.002	3	1.5	0.0	7.2	32	0.20
Total/mean	20	3.7	2.7	–	6	0.43

† Standardized values calculated by resampling ten sublines per Mexican and two per Turkish LC without replacement. The mean was then calculated from 5000 repetitions.

A total of 76 SSRs was applied to fingerprint the LC accessions in Set 1. On the basis of these results the 44 most polymorphic SSRs, equally distributed over the entire genome, were selected for the analyses of Set 2. Simple sequence repeat information was provided by the Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany (Röder et al., 1998; Röder, unpublished data, 2000) and DuPont (Wilmington, DE) (Eujayl et al., 2002; DuPont, unpublished data, 2001). In addition, the marker *WMC56* developed by the Wheat Microsatellite Consortium (Agrogene, France) was used. Information on map location, repeat type, annealing temperature, fragment sizes, number of alleles, as well as polymorphic information content for each SSR is available at http://www.cimmyt.org/english/webp/support/publications/support_materials/ssr_mw1.htm (verified 14 Nov. 2004). PCR amplification and allele detection were performed with an ABI-Prism Sequencer 377 in combination with the computer software GeneScan 3.1 and Genotyper 2.1 (PerkinElmer Biotechnologies, Foster City, CA), as described in detail by Dreisigacker et al. (2004).

Statistical Analyses

The proportion of SSRs showing multiple bands (MB) was determined to estimate the genetic variation of each accession in Set 1. The presence of MB indicates that for a given SSR more than one allele was observed, which may reflect residual heterozygosity and/or segregation at the respective SSR marker. Ordinary *t* tests were calculated to compare the observed genetic variation of LC accessions composed of bulks or single LC sublines (SAS Institute, 1990).

For the comparison of the LCs in Set 2, which was based on different numbers of LC sublines, standardized average numbers of observed alleles per locus and standardized numbers of unique alleles were calculated. Standardized values were computed by resampling 10 sublines per Mexican and two per Turkish LC and taking means over 5000 repetitions. Gene diversity of each Mexican and Turkish LC was calculated according to Nei (1973).

Analyses of molecular variance (AMOVA) were conducted on the basis of SSR data to divide the genetic variation in Set 2 into components attributable to variance components among and within LCs. Pairwise fixation indices were determined to estimate the extent of LC isolation by distance within the two countries, Mexico and Turkey. Significance levels were computed by permuting sublines between LCs.

Modified Rogers' distance was calculated for each pairwise combination in Set 1 and Set 2 according to the following equation (Wright, 1978):

$$MRD = \sqrt{\frac{1}{2m} \sum_{i=1}^{m} \sum_{k=1}^{a_i} \left(p_{ij} - q_{ij} \right)^2},$$

where p_{ij} and q_{ij} are the allele frequencies of the *j*th allele at the *i*th marker; a_i refers to the number of alleles at the *i*th marker; and *m* is the number of SSRs. The allele frequencies of the accessions in Set 1 were estimated on the basis of the peak area and height of each band in the electrophoresis detected by GeneScan 3.1. Standard errors of the MRD estimates were obtained by a bootstrap procedure with resampling 1000 times over markers (Weir, 1996). Principal coordinate analyses (PCoA) were performed on the basis of the MRDs to visualize the dispersion of genotypes in Set 1 and Set 2 (Gower, 1966).

The AMOVA and pairwise F_{ST} values were calculated by the software package Arlequin (Schneider et al., 2000). All other analyses were performed by applying the Plabsim software (Frisch et al., 2000), which is implemented as an extension of the statistical software R (Ihaka and Gentleman, 1996).

RESULTS

Genetic Diversity among 36 Wheat Landrace Cultivar Accessions

The 76 SSRs assayed in Set 1 resulted in a total of 419 alleles, with 11 SSRs detecting monomorphic bands. The average number of alleles per locus accounted for 6.0 alleles with a minor variation among the three genomes (Table 3). Most of the SSR loci of the LC accessions were homozygous. On average 10.0% of the SSRs showed MB.

In Set 1, SSRs amplifying more than two distinct al-

Table 3. Number of SSRs and alleles per locus, as well as percentage of SSRs with multiple bands (MB) per accession determined for the three genomes in 36 landrace cultivar (LC) accessions of wheat.

Genome	No. of SSRs	Alleles per locus		SSRs with MB per accession†	
		Average	Range	Average	Range
A	20	5.9	1–13	11.5	0.0–47.1
B	27	6.8	2–16	9.0	0.0–48.0
D	24	5.8	1–17	9.5	0.0–45.8
Total	76‡	6.0	1–17	10.0	0.0–44.9

† MB for a given SSR in one accession reflect to residual heterozygosity and/or segregation.
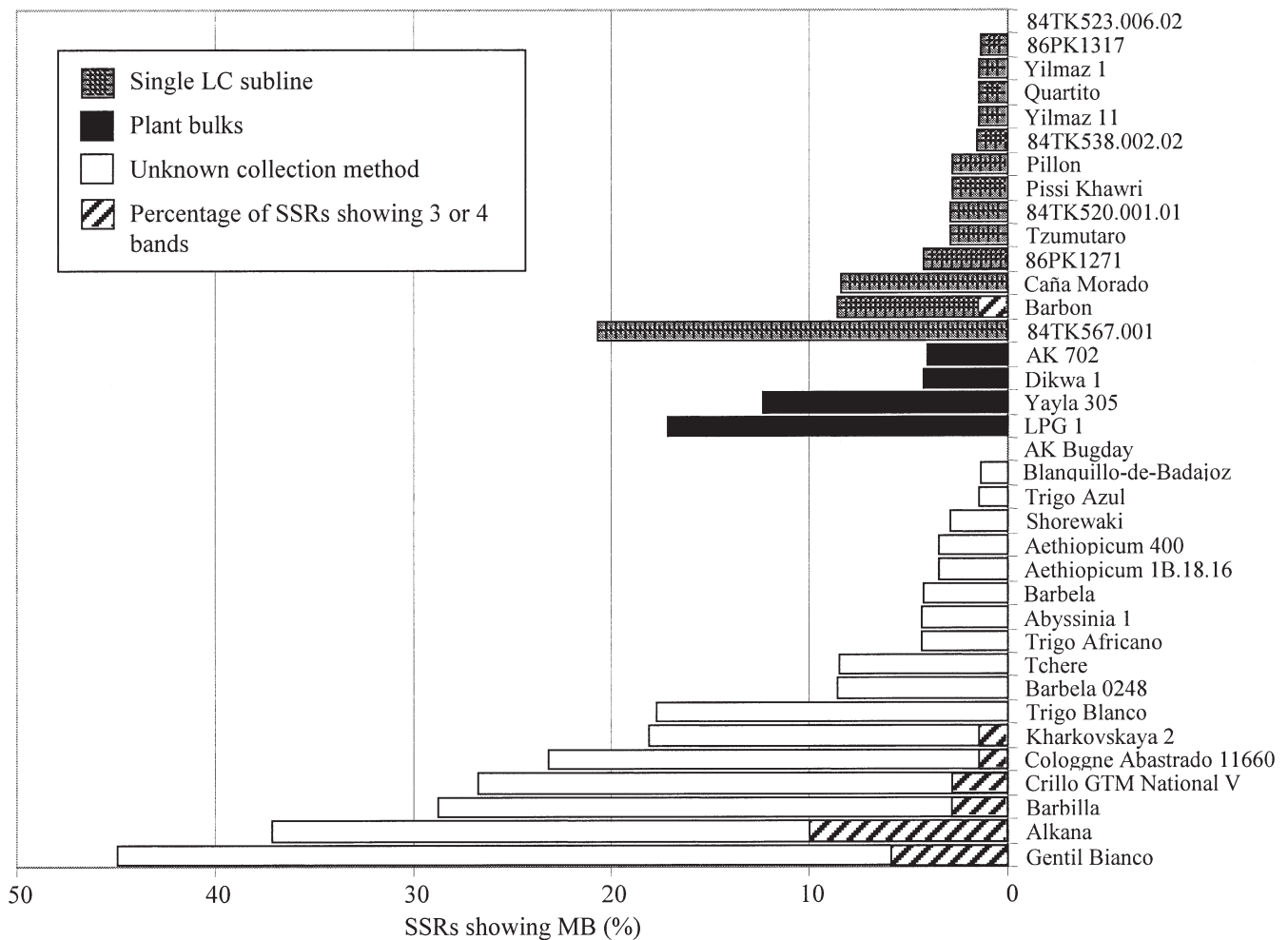‡ Genome location of 5 SSRs was unknown.

Fig. 1. **Proportion of SSR loci with multiple bands (MB) determined in each of 36 landrace cultivar (LC) accessions of wheat.**

leles per SSR were found in LC accessions with generally high allelic variation. Accession 27 (84TK567.001) showed a fairly high proportion of SSRs with MB (20.6%), although it was based on a single LC subline (Fig. 1). The mean proportion of SSRs with MB did not significantly differ ($P < 0.05$) between accessions composed of bulks and single sublines.

The MRD between LC accessions of Set 1 averaged 0.69. The lowest MRD value (0.16) was observed between the LC accessions Barbela and Barbela 0248 and the highest value (0.82) between the LC accessions Aethiopicum 400 and Yilmaz 1. Standard errors of MRD estimates ranged from 0.02 to 0.06. In the PCoA based on MRD estimates, the first three principal coordinates (PC) explained 8.7, 7.8, and 6.9% of the total variation, respectively (Fig. 2). The accessions did not group according to their continent or country of origin for the most part. Three Turkish accessions (84TK520.001.01, 84TK523.006.02, and 84TK567.001) formed a distinct cluster. The accessions from Mexico and Guatemala, were separated together with accession 18 (Blanquillo-de-Badajoz) from Spain and accession 5 (Abyssinia 1) from Ethiopia on the basis of PC3. A close relationship was revealed between accession 35 (Trigo Africano)

originating from Chile and accession 6 (Dikwa 1) originating from Nigeria (MRD = 0.18).

Genetic Diversity within and between Mexican and Turkish Landrace Cultivars

Tzumutaro and Caña Morado were the most diverse ($H_e = 0.41$) Mexican LCs in Set 2 on the basis of the high average number of alleles per locus and the number of unique alleles (Table 2). The lowest number of unique alleles was observed in Barbon, which was still highly diverse ($H_e = 0.37$) becaue of heterozygosity. Among the Turkish LCs in Set 2, gene diversity was highest ($H_e = 0.55$) in 84TK538.002.02. Only three LC sublines were available from 84TK567.002, which revealed 32 monomorphic and no segregating loci.

In the AMOVA, 18.4% of the total variance was found between the combined populations of Mexican vs. Turkish LCs in Set 2. Considering exclusively Mexican LCs, the variance within the populations accounted for 52.3%. All Mexican LCs were significantly ($P < 0.05$) different from each other, whereas corresponding pairwise F_{ST} values ranged from 0.37 to 0.68 (Table 4). Variance within was twice as large (67.6%) than between the Turkish LCs. The highest F_{ST} value (0.62) was found between
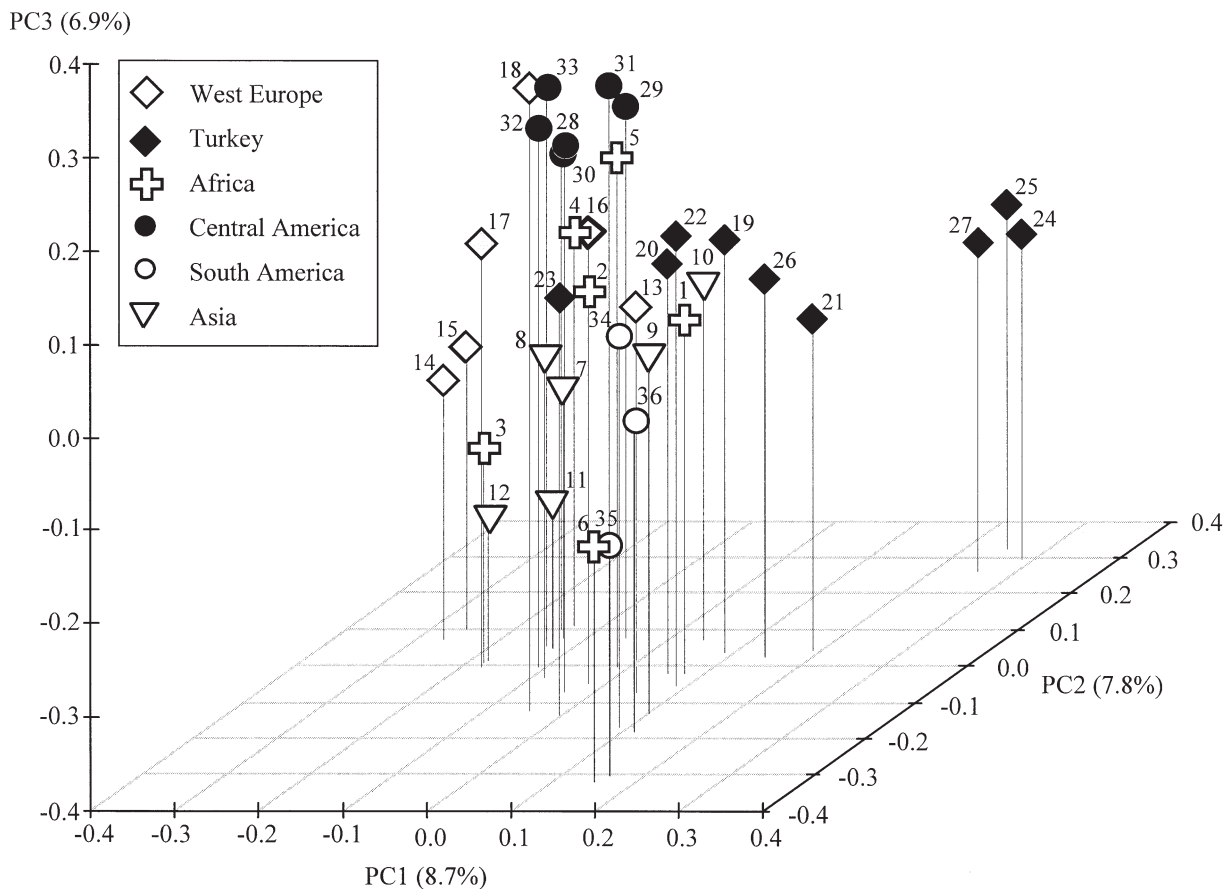
Fig. 2. Associations among 36 landrace cultivar (LC) accessions of wheat revealed by principal coordinate analysis (PCoA) performed with modified Rogers' distances (MRD) calculated from 76 SSRs. Numbers refer to the list in Table 1. Geographic origin is designated by symbols (see legend).

84TK523.001.01 and 84TK567.002 by far the smallest value (0.07) between 84TK.567.001 and 84TK.567.002. A considerably higher range of F_{ST} values was observed for the Mexican LCs.

Principal coordinate analysis revealed a clear grouping among the Mexican LCs with the exception of two sublines of Quartito that were located outside its main cluster and closer to sublines from other LCs (Fig. 3A). However, mean genetic distances of these two sublines to the LC main cluster were smaller (MRD = 0.72 and 0.82) than the maximum genetic distance within the main cluster (MRD = 0.83). Groupings of the Turkish LCs were less clear (Fig. 3B) as reflected by the larger variation within the Turkish LCs in the AMOVA. Sublines of 84TK538.002.02 and 84TK567.001 were widely dispersed and did not form a single main cluster.

DISCUSSION

Wheat Landrace Cultivar Diversity

The genetic variability of LCs has been affected by various factors throughout their evolutionary history. In autogamous crops, outcrossing and fitness-relevant mutations generate an intrapopulation diversity, whereas directed natural or human selection and bottleneck effects lead to an increase in interpopulation diversity (Ennos, 1983).

In our study, a considerable genetic diversity was revealed within rather than between Mexican and Turkish LCs. A surprisingly high intrapopulation diversity seemed to be in contrast to the high selfing rate of wheat but was consistent with previous results found in Pakistani wheat LCs analyzed with protein markers (Tahir et al., 1996) and Italian LCs of emmer [*Triticum turgidum* L. subsp. *dicoccum* (Schrank ex Schübl.) Thell.] analyzed with RAPDs (Barcaccia et al., 2001). The higher diversity observed within the Turkish than within the Mexican LCs can be explained by a much longer evolutionary history of wheat in Turkey. Furthermore, wheat LCs

Table 4. Pairwise fixation index (F_{ST}) for five Mexican and four Turkish landrace cultivars (LCs).

	LC			
LC	**1**	**2**	**3**	**4**
		Mexican LCs		
1: Pillon				
2: Barbon	0.50*			
3: Quartito	0.68*	0.42*		
4: Caña Morado	0.50*	0.37*	0.50*	
5: Tzumutaro	0.47*	0.37*	0.54*	0.39*
		Turkish LCs		
1: 84TK523.001.01				
2: 84TK538.002.02	0.31*			
3: 84TK567.001	0.41*	0.17*		
4: 84TK567.002	0.62*	0.38*	0.07	

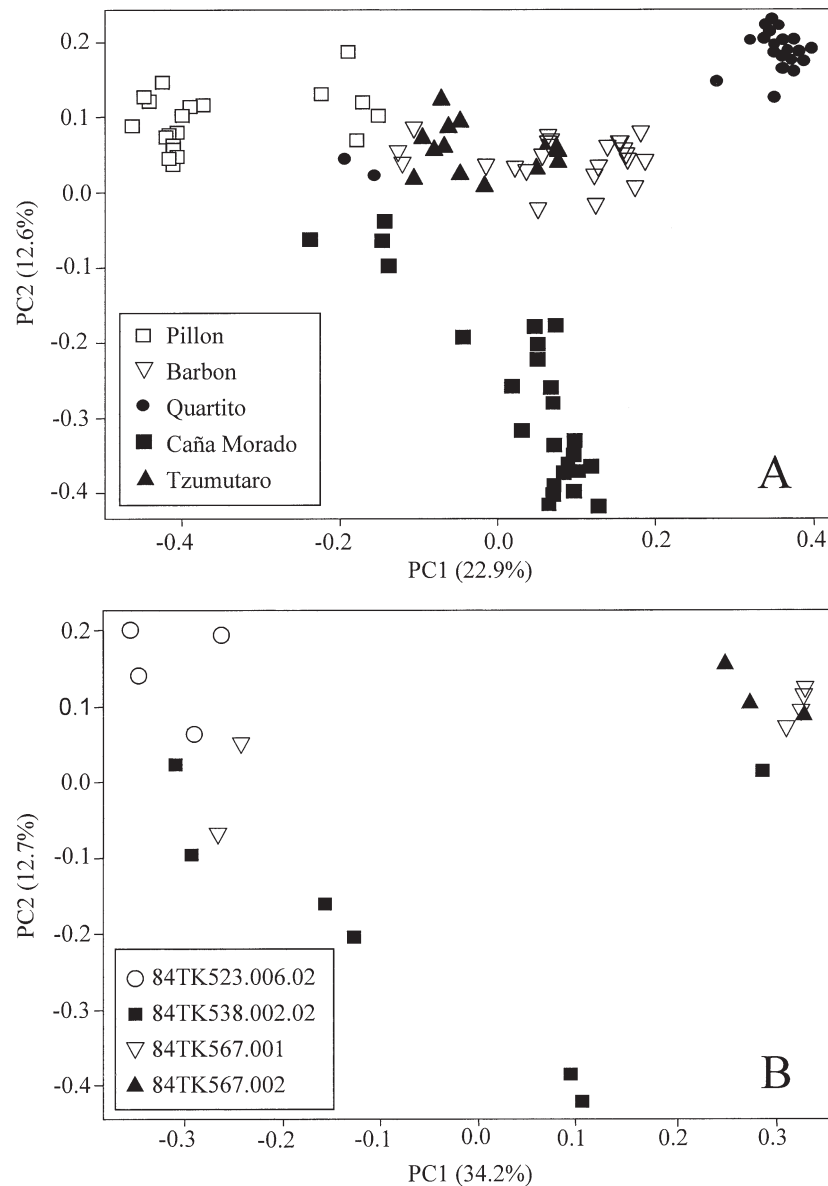* Significant at the 0.05 probability level.

Fig. 3. Associations among five Mexican (A) and four Turkish (B) landrace cultivars (LC) of wheat revealed by principal coordinate analyses (PCoA) performed with modified Rogers' distances (MRD) calculated with 44 SSRs. The LCs consisted of 3 to 25 sublines.

or varieties mainly transferred from Spain to the New World were presumably limited in population size, thus resulting in a founder effect.

The LCs in Set 1 were not grouping according to their continent or country of origin in the PCoA (Fig. 2). We speculate that many LCs analyzed in our study were relatively late in history transferred from the Near East or Europe to other parts of the world and/or environmental adaptation changed their genetic composition only little. The Turkish LCs, which formed a distinct cluster, were collected in the primary center of diversity of wheat in proximal locations in Hakkari, Turkey. All three LCs show resistance to Russian wheat aphid. As expected, the two LCs Barbela and Barbela 0248 were closely related, the latter being considered as a subrace of Barbela, a very old Portuguese LC showing impressively wide adaptation to different environments, in par-

ticular high acid soil and drought tolerance. The Chilean LC Trigo Africano clustered together with the African LC Dikwa 1, which was collected in a small homonymous region in the northeast of Nigeria (Zeven, 1974). The Spanish name Trigo Africano directly refers to the continent of origin, Africa, but not necessarily to the Nigerian region. In view of the lack of historical records, a larger number of accessions per country should be fingerprinted before drawing any firm conclusions about the evolutionary relationships of accessions.

Diversity within wheat LCs rests more on the allelic variation between individual plants than on heterozygous individuals. In our study, this was reflected by a low mean of heterozygosity (2.6%) observed in the sublines of Mexican and Turkish LCs, which was similar to the mean (2.5%) reported for improved lines from CIMMYT's wheat breeding program (Dreisigacker et

al., 2004). An exception was the Turkish LC accession 27 (84TK567.001), which was supposedly composed of a single subline, but showed an extremely high percentage of SSRs (20.6%) with MB (Fig. 2). This high variation might be due to outcrossing, seed contamination, or experimental errors. Employing the formula of Crow and Kimura (1970, p. 93), our estimate of mean heterozygosity corresponds to 1.3% outcrossing rate and is thus slightly higher than reported in the literature (Martin, 1990; Hucl, 1996). This outcrossing rate is sufficient to generate off-types by contamination with foreign pollen. Outcrossing might also explain why some sublines of the Mexican and Turkish LCs were positioned separately from their main clusters in the PCoA (Fig. 3). Furthermore, the Turkish LCs could be intercrossed with wild species, such as goatgrass [*Triticum tauschii* (Coss.) Schmal.], which are still widely grown in the mountain regions of Hakkari (Braun et al., 2001).

Bulk versus Individual Plant Conservation

In early expeditions of genetic resources acquisition, collections of bulks were preferred since the prime focus was to collect as much material as possible in a short time and to cover widely diverse geographic regions. Collecting individual plants separately was first advocated by Bennett (1970) and later reinforced by Ford-Lloyd and Jackson (1986). On one hand, the conservation in bulks offers the advantage of including seed of many different plants, which prevents a dramatic reduction in the original population size and simplifies the procedure of sampling and conservation (Frankel, 1977; Marshall, 1990). On the other hand, the presence of different genotypes makes a precise characterization of bulks difficult. Bulk accessions must therefore be "de-bulked" or evaluated on a larger scale before the best individuals are identified and used in prebreeding programs.

We observed a low molecular variation in LC accessions conserved as bulks. In general, the variance of genetic diversity measures increases with reduced numbers of examined genotypes (Weir, 1996). The variance of gene diversity in the Turkish LCs of Set 2 was higher than in Mexican LCs, the former being composed of only three to seven LC sublines (Table 2). The regeneration procedure at CIMMYT, where only 100 seeds are sown per accession, could be an additional reason for the low molecular variation observed in the bulk accessions. Small effective population sizes lead to the risk of losing molecular variation during seed regeneration. Major threats are genetic drift and selection as shown in previous studies on barley (*Hordeum vulgare* L.) (Parzies et al., 2000), rice (*Oryza sativa* L.) (Gao et al., 2000), and rye (*Secale cereale* L.) (Chwedorzewska et al., 2002). Therefore, larger samples for seed regeneration are recommended in the literature. Assuming a population with 20 000 polymorphic loci and two alleles per locus, Lawrence et al. (1995) concluded that about 172 plants are sufficient to conserve nearly all alleles with frequencies not lower than 0.05. According to Crossa and Vencovsky (1999), for 5 to 100 loci and 2 to 20 alleles per locus, between 105 and 335 plants per population are required

to maintain alleles at a 5% frequency. Some of the bulk accessions in our study were probably subsamples received from or shared with cooperators. These samples usually contain only 100 to 200 seeds, which could be another reason for a loss of variation.

In individual plant collections, alleles are usually fixed in each accession. Because of their uniformity, the accessions can be more precisely characterized and, hence, exploitation by breeders may proceed more rapidly (DeLacy et al., 2000). Its disadvantages are extensive space and labor costs essential for conservation and seed regeneration. The genetic variation within individual plant collections directly depends on the number of collected sublines. The Mexican and Turkish LCs in Set 2 might therefore represent only a part of the variation present in the original LCs. Indigenous knowledge of LCs would be extremely useful for the optimization of the sampling of sublines of each particular LC (for a review see Zeven, 2002).

Both ex situ conservation methods maintain only a part of the original LCs genetic variation and disregard their integrity. For instance, low input agriculture relies on the buffering effect of LCs, which is responsible for their broad adaptation but requires the intact original level of diversity. Thus, a combination of both conservation forms could be a reasonable solution: the storage of (i) a large bulk to preserve the natural state of the LC variation in a simple manner, and (ii) separate LC sublines representing potentially useful variants for breeding programs.

Implications of SSR-Based Genetic Diversity for Seed Bank Management

Currently some of the limiting factors in the use of LC ex situ collections are (i) missing or incomplete passport data, and (ii) the precise characterization of the collections. Passport data were not available for half of the 36 CIMMYT wheat LC accessions used in our study. Most collecting expeditions were of such a short duration that it was difficult to locate and interview all relevant landowners at the collection sites. Additionally, personnel, management, and political changes in seed banks may have contributed to the incompleteness of the records. Molecular markers may provide new and reliable information for the description and optimization of LC collections in seed banks.

The increasing costs to efficiently manage large ex situ collections encourage curators to identify redundant germplasm accessions. Verifying duplications is complex, because their definition can vary from "accessions with similar passport data" to "identical genotypes" (Hintum, 2000). Suspected duplicates were identified in collections of sorghum [*Sorghum bicolor* (L.) Moench] and barley by means of 15 and 35 SSRs in combination with passport data and AMOVA as a biometrical tool (Dean et al., 1999; Lund et al., 2003). In our study, 84TK567.001 and 84TK567.002 were assumed to be closely related, because of adjacent collection sites. Applying AMOVA, these LCs showed nonsignificant differences, further strengthening the notion to manage these two individual

collections as one LC. However, in wheat, which is particularly suitable for seed storage, conserving either new or existing accessions in perpetuity (including regeneration in 25-yr intervals, germination tests, etc.) is currently still more cost-effective than DNA fingerprinting even with a relatively small number of SSRs (Dreher et al., 2000; Pardey et al., 2001). Thus, the identification and removal of suspected duplicates should not be considered as the main role of molecular screenings in seed bank collections.

The assessment of LC variability in seed banks demands large-scale screenings of collections. According to Zhang et al. (2002), 300 to 400 alleles are required to reflect stable relationships between wheat accessions and effectively establish core collections. In our study, 256 alleles detected with 44 SSRs (two SSRs per chromosome) were sufficient to differentiate individual genotypes of Mexican and Turkish LCs. Moreover, half of the SSRs applied were developed from expressed sequence tags, which are generally less polymorphic but might reflect functional diversity more accurately. Future opportunities to combine these markers and phenotypic data in association studies may narrow the search for new alleles at loci of interest (Thornsberry et al., 2001).

In conclusion, SSRs provide important information about the genetic variation of wheat LCs and demonstrate a powerful tool for the future tasks of seed bank management. However, the high costs warrant the further optimization of SSR application. The standardization of molecular methods, for instance, would allow to coordinate collections of different seed banks and the incorporation of new technologies like GPS, to relate the molecular diversity with their geographic dispersion.

ACKNOWLEDGMENTS

REFERENCES

Barcaccia, G., L. Molinari, O. Porfiri, and F. Veronesi. 2001. Molecular characterization of emmer (*Triticum diccon* Schrank) Italian landraces. Genet. Resour. Crop Evol. 49:415–426.

Bennett, E. 1970. Tactics of plant exploration. p. 157–179. *In* O.H. Frankel and E. Bennett (ed.) Plant genetic resources in plants: Their exploration and conservation. F.A. Davis Co., Philadelphia.

Börner, A., S. Chebotar, and V. Korzun. 2000. Molecular characterization of the genetic integrity of wheat (*Triticum aestivum* L.) germplasm after long-term maintenance. Theor. Appl. Genet. 100:494–497.

Braun, H.J., N. Zencirci, F. Altay, A. Atli, M. Avci, V. Eser, M. Kambertay, and T.S. Payne. 2001. Turkish wheat pool. p. 851–879. *In* A.P Bonjean and W.J Agnus (ed.) The world wheat book: A history of wheat breeding. Lavosier Publishing, Paris.

Chwedorzewska, K.J., P.T. Bednarek, and J. Puchalski. 2002. Studies on changes in specific rye genome regions due to seed aging and regeneration. Cell. Mol. Biol. Lett. 7:569–576.

Crossa, J., and R. Vencovsky. 1999. Sample size and variance effective

population size for genetic resource conservation. FAO/IBPGR Plant Genet. Resour. Newsl. 119:15–25.

Crow, J.F., and M. Kimura. 1970. An introduction to population genetics theory. Harper & Row, New York.

Dean, R.E., J.A. Dahlberg, M.S. Hopkins, S.E. Mitchell, and S. Kresovich. 1999. Genetic diversity and redundancy among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci. 39:1215–1221.

DeLacy, I.H., B. Skovmand, and J. Huerta. 2000. Characterization of Mexican wheat landraces using agronomically useful attributes. Genet. Resour. Crop Evol. 47:591–602.

Dreher, K., M.L. Morris, M.M. Khairallah, J.M. Ribaut, S. Pandey, and G. Sirinivasan. 2000. Is marker-assisted selection cost-effective compared to conventional plant breeding methods? p. 203–236. *In* R.E. Evenson and D. Zilberman (ed.) Economic and social issues in agriculture biotechnology. CABI Publishing, CAB International, Wallingford, UK.

Dreisigacker, S., P. Zhang, M. van Ginkel, M. Warburton, D. Hoisington, M. Bohn, and A.E. Melchinger. 2004. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different mega-environments. Crop Sci. 44:381–388.

Ennos, R.A. 1983. Maintenance of genetic variation in plant populations. p. 129–155. *In* M.K. Hecht et al. (ed.) Evolutionary biology. Vol. 16. Plenum Press, New York.

Eujayl, I., M.E. Sorrells, M. Braun, P. Wolters, and W. Powell. 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor. Appl. Genet. 104:399–407.

Ford-Lloyd, B., and M. Jackson. 1986. Plant genetic resource: An introduction to their conservation and use. Eduard Arnold, London.

Frankel, O.H. 1977. Genetic resources. *In* P.R. Day (ed.) The genetic basis of epidemics in agriculture. Ann. N.Y. Acad. Sci. 287:332–334.

Frisch, M., M. Bohn, and A.E. Melchinger. 2000. Plabsim: Software for simulation of marker-assisted backcrossing. J. Hered. 91:86–87.

Gao, L.Z., A. Schaal, C.H. Zhang, J.Z. Jia, and Y.S. Dong. 2000. Assessment of population genetic structure in common wild rice *Oryza rufipogon* Griff. using microsatellite and allozyme markers. Theor. Appl. Genet. 106:173–180.

Gollin, D., M. Smale, and B. Skovmand. 2000. Searching an *ex situ* collection of wheat genetic resources. Am. J. Agric. Econ. 82:812–827.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–388.

Harlan, J.R. 1975. Our vanishing genetic resources. Science (Washington, D.C.) 188:618–621.

Hammer, K. 2000. Microsatellite markers—A new tool for distinguishing diploid wheat species. Genet. Resour. Crop Evol. 47:497–505.

Hede, A.R., B. Skovmand, M.P. Reynolds, J. Crossa, A.L. Vilhelmsen, and O. Stolen. 1999. Evaluating genetic diversity for heat tolerance traits in Mexican wheat landraces. Genet. Resour. Crop Evol. 46: 37–45.

Hintum, T.J.L. 2000. Duplications within and between germplasm collections. III. A quantitative model. Genet. Resour. Crop Evol. 47:507–513.

Hoisington, D., M. Khairallah, and D. Gonzalez-de-Leon. 1994. Laboratory protocols: CIMMYT Applied Molecular Genetics Laboratory. 2nd ed. (Available online at http://www.cimmyt.org/ABC/Protocols/manualABC.html; verified 14 Nov. 2004). CIMMYT, Mexico, D.F.

Hucl, P. 1996. Out-crossing rates for 10 Canadian spring wheat cultivars. Can. J. Plant Sci. 76:423–427.

Huang, X.Q., A. Börner, M.S. Röder, and M.W. Ganal. 2002. Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. Theor. Appl. Genet. 105:699–707.

Ihaka, R., and R. Gentleman. 1996. R: A language for data analysis and graphics. J. Comput. Graph. Stat. 5:299–314.

Kihara, H. 1982. Wheat studies: Retrospect and prospects. Elsevier Science Ltd., St. Louis, MO.

Lawrence, M.J., D.F. Marshall, and P. Davies. 1995. Genetics of genetic conservation. I. Sample size when collecting germplasm. Euphytica 84:89–99.

Li, Y.-C., T. Fahima, J.H. Peng, M.S. Röder, V.M. Kirzhner, A. Beiles, A.B. Korol, and E. Nevo. 2000. Edaphitic microsatellite DNA divergence in wild emmer wheat, *Triticum dicoccoides*, at a microsite: Tabigha, Israel. Theor. Appl. Genet. 101:1029–1038.

Lund, B., R. Ortiz, I.M. Skovgaard, R. Waugh, and S.B. Anderson.

2003. Analysis of potential duplicates in barley seed bank collections using re-sampling of microsatellite data. Theor. Appl. Genet. 106:1129–1138.

Marshall, D.R. 1990. Crop genetic resources: Current and emerging issues. *In* A.H.D. Brown et al. (ed.) Plant population genetics, breeding, and genetic resources. Sinauer Assoc. Inc., Sunderland, MA.

Martin, T.J. 1990. Outcrossing in twelve hard red winter wheat cultivars. Crop Sci. 30:59–62.

Nei, M. 1973. Analyses of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70:3321–3323.

Pardey, P.G., Bonwoo Koo, B.D. Wright, M.E. van Dusen, B. Skovmand, and S. Taba. 2001. Costing the conservation of genetic resources: CIMMYT's ex-situ maize and wheat collection. Crop Sci. 41:1286–1299.

Parzies, H.K., W. Spoor, and R.A. Ennos. 2000. Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp. *vulgare*) conserved for different lengths of time in *ex situ* gene banks. Heredity 84:476–486.

Röder, M.S., V. Korzun, K. Wendehake, J. Plaschke, M.H. Tixier, P. Leroy, and M.W. Ganal. 1998. A microsatellite map of wheat. Genetics 149:2007–2023.

SAS Institute. 1990. SAS Institute Inc., Cary, NC.

Schneider, S., and D. Roessli. amd L. Excoffier. 2000. Arlequin version 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, Univ. of Geneva, Switzerland. (Available online at http://lgb.unige.ch/arlequin/; verified 14 Nov. 2004).

Skovmand, B., and S. Rajaram. 1990. Utilization of genetic resources in the improvement of hexaploid wheat. *In* J.P. Srivastava and A.B. Damania (ed.) Wheat genetic resources: Meeting diverse needs. J. Wiley & Sons, Chichester, UK.

Skovmand, B., G. Varughese, and G.P. Hettel. 1992. Wheat genetic resources at CIMMYT: Their preservation, enrichment, and distribution. CIMMYT, Mexico, D.F.

Skovmand, B., L. Bertschinger, and J. Robinson. 1994. Quantifying heterogeneity for reactions to the Russian wheat aphid in accessions of Turkish hexaploid wheat landraces previously identified as resistant. p. 139–142. *In* Proc. of the Russian Wheat Aphid Workshop, 6, Fort Collins, CO. 23–25 Jan. 1994.

Skovmand, B., M.P. Reynolds, and I.H. DeLacy. 2001. Mining wheat germplasm collections for yield enhancing traits. Euphytica 119: 25–32.

Tahir, M., T. Turchetta, R. Anwar, and D. Lafiandra. 1996. Assessment of genetic variability in hexaploid wheat landraces of Pakistan based on polymorphisms for HMW glutein subunits. Genet. Resour. Crop Evol. 43: 221–220, 1996.

Thornsberry, J.M., M.M. Goodman, J. Kresovich, S. Nielsen, and E. Buckler. 2001. Dwarf8 polymorphisms associate with variation in flowering time. Nat. Genet. 28:286–289.

Weir, B.S. 1996. Genetic data analysis II. Sinauer Associates, Sunderland, MA.

Wright, S. 1978. Evolution and genetics of populations. University Chicago Press, Chicago.

Zeven, A.C. 1974. Indigenous bread wheat varieties from northern Nigeria. Acta Bot. Neerl. 23:137–144.

Zeven, A.C. 2002. Traditional maintenance breeding of landraces: 2. Practical and theoretical considerations on maintenance of variation of landraces by farmers and gardeners. Euphytica 123:147–158.

Zhang, X.Y., C.W. Li, L.F. Wang, H.M. Wang, G.X. You, and Y.S. Dong. 2002. An estimation of the minimum number of SSR alleles needed to reveal genetic relationships in wheat varieties. I. Information from large scale planted varieties and cornerstone breeding parents in Chinese wheat improvement and production. Theor. Appl. Genet. 106:112–117.

Hybrid performance and heterosis in spring bread wheat, and their relations to SSR-based genetic distances and coefficients of parentage[5]

S. Dreisigacker, A.E. Melchinger, P. Zhang, K. Ammar, C. Flachenecker, D. Hoisington, and M.L. Warburton*

S. Dreisigacker, A.E. Melchinger, C. Flachenecker, Inst. of Plant Breeding, Seed Science, and Population Genetics, Univ. of Hohenheim, 70593 Stuttgart, Germany; P. Zhang, K. Ammar, D. Hoisington, M.L. Warburton, CIMMYT, International Maize and Wheat Improvement Center, Apdo. Postal 6-641, 06600 Mexico D.F., Mexico, *Corresponding author (mwarburton@cgiar.org).

ABSTRACT

Development of hybrids is considered as a promising avenue to enhance the yield potential of crops. We investigated (i) the amount of heterosis, (ii) relative importance of general (GCA) *vs.* specific (SCA) combining ability, and (iii) the relationship between heterosis and genetic distance measures in four agronomic traits of spring bread wheat. Eight male and 14 female lines, as well as their 112 hybrids produced in a factorial design were grown in replicated trials at two environments in Mexico. Principal coordinate analysis based on Rogers' distance (RD) estimates calculated from 113 SSRs revealed three different groups of parents. Mid-parent heterosis (MPH) for grain yield averaged 0.02 t ha^{-1} (0.5%) and varied from -15.33% to 14.13%. MPH and hybrid performance (F1P) were higher for intra-group hybrids than for inter-group hybrids, with low values observed in inter-group crosses involving two non-adapted Chinese parents. Combined analyses of variance revealed significant differences among parents and among hybrids. Estimates of GCA variances were more important than SCA variances for all traits. Tight correlations of GCA with line *per se* performance, and mid-parent value with F1P were observed for all traits. In contrast, correlations of MPH with RD and coefficient of parentage were not significant. It was concluded that the level of heterosis was too low to warrant a

commercial exploitation in hybrids. SSRs proved to be a powerful tool for the identification of divergent groups in advanced wheat breeding materials.

Abreviations: COP, coefficient of parentage; EST, expressed sequence tag; GCA, general combining ability; LP, line *per se* performance; PC, principal coordinate; PCoA, principal coordinate analysis; PIC, polymorphism information content; RD, Rodgers distance, SCA, specific combining ability; SSR, simple sequence repeat.

INTRODUCTION

In view of the growing need for wheat production worldwide, the International Maize and Wheat Improvement Center (CIMMYT) considers the development of hybrids as a promising option in a multiproned strategy for breaking the yield barrier in wheat (Reynolds et al., 1996). Hybrid wheat has shown potential for enhanced yield performance and stability across diverse environments. For its successful implementation, three prerequisites are considered crucial: (i) a cost-effective system of seed production, (ii) adequate levels of heterosis, and (iii) the development of heterotic groups and patterns to ensure future progress through hybrid breeding (Lang, 1989). During the last four decades, hybrid wheat research has mainly focused on the introduction of male sterility and on acceptable cross-pollination characteristics (Pickett, 1993). The production of hybrids has been greatly enhanced by the discovery of effective chemical hybridizing agents (Pickett and Galwey, 1997). However, knowledge about hybrid performance, the relative importance of general (GCA) and specific (SCA) combining ability, and the genetic background of parental materials for maximum exploitation of heterosis in wheat, remains limited.

Quantitative genetic theory suggests that high heterosis can be expected in a hybrid if the source populations have (i) a high frequency of genes with partial or complete dominance and/or (ii) maximum differences in gene frequencies of overdominant loci (Hallauer et al., 1988). Consequently, for an optimum exploitation of heterosis, parents should be derived from genetically divergent germplasm pools, commonly referred to as heterotic groups (Melchinger and Gumber, 1998). Separate cultivation of populations of maize and other allogamous crops such as rye, facilitated their classification into heterotic groups according to their evolutionary history and geographic origin. In contrast, heterotic

groups are not available or easily discernable in wheat, owing to its past breeding history. As for other autogamous crops, breeding of pure line cultivars has relied on deriving transgressive segregants from crosses between complementary parents obtained through a relatively widespread exchange of germplasm (Heisey et al., 2002). Furthermore, introgression of genes for new resistance or tolerance to various biotic and abiotic stresses from wild relatives into modern cultivars contributed to a mix of germplasm from diverse genetic origins.

The coefficient of parentage (COP) as well as phenotypic and molecular data have been used to study diversity among subsets of wheat germplasm (Souza et al., 1994). A linear association between marker-based genetic distance and heterosis was determined both in theory (Charcosset and Essioux, 1994) and in numerous experiments with maize (Reif et al., 2003) and other crops (Brummer, 1999). In wheat, a few studies applying RFLP or RAPD markers were carried out, but no clear relationship between molecular diversity and heterosis could be observed (Liu et al., 1999; Corbellini et al., 2002). However, marker systems used in these studies were of limited use in wheat owing to a low degree of polymorphism or poor reproducibility. Currently, simple sequence repeats (SSRs) represent the most suitable marker system in wheat. They allow an even coverage of the genome, are abundant, genome specific, co-dominant in nature, and have been successfully used to characterize genetic diversity in advanced wheat breeding materials (Dreisigacker et al., 2004; Röder et al., 2002).

The objectives of our research were to (i) determine the genetic diversity among 22 spring bread wheat lines representing widely grown landmark cultivars or successful breeding lines from CIMMYT, (ii) investigate the level of heterosis in hybrids produced from these lines, (iii) assess the relative importance of general *vs.* specific combining effects in the hybrids, and (iv) examine the relationship between heterosis and genetic distances based on COP and SSR markers.

MATERIALS AND METHODS

Genetic materials

Twenty parents were chosen from a set of widely grown landmark cultivars and successful breeding lines from CIMMYT, to represent contrasting phenotypes and diverse genetic backgrounds for production of hybrids (Table 1). In addition, two Chinese lines were used as male parents. Fourteen hand- emasculated female lines were crossed with eight male parents to produce 112 hybrids in a factorial mating design during the 2001 and 2002 winter seasons at Ciudad. Obregon, Sonora, Mexico.

SSR analyses

Genomic DNA of the parental lines was extracted from bulked leaves harvested from seven to ten young plants using the modified CTAB procedure described in CIMMYT's manual of laboratory protocols (CIMMYT, 2001). A set of 113 SSRs (60 genomic- and 55 expressed sequence tag (EST) derived SSRs) was employed covering the three (A, B, and D) wheat genomes. Four to eight SSRs were located on each chromosome, whereas the map location of five SSRs was unknown. The genomic-derived SSRs were developed by M. Röder at the Institute of Plant Genetics and Crop Plant Research in Gatersleben, Germany, and by the Wheat Microsatellite Consortium Agrogene in France. The EST-derived SSRs were developed at DuPont, Wilmington, USA. In addition, the SSR marker 'Taglgap' developed by Devos et al. (1995) was used. The SSRs were multiplexed for maximum efficiency. Information on map location, repeat type, annealing temperature, fragment sizes, number of alleles, and polymorphic information content (PIC) for each SSR is available at http://www.cimmyt.org/english/web/support/publications/support_ materials/ssr_mwl.htm. PCR amplification and allele detection were performed using an ABI-Prism Sequencer[TM]377 in combination with computer software GeneScan 3.1 and Genotyper 2.1 (Perkin Elmer Biotechnologies), as described in detail by Dreisigacker et al. (2004).

Table 1. Name, abbreviation, year and country of release, as well as average grain yield evaluated at two environments in Mexico of 22 spring bread wheat lines used in the production of hybrids.

Wheat line	Abbr.	Parent	Year of release	Country of release	Grain yield (t ha^{-1})
Sonora 64	SN	female	1964	Mexico	4.21
Sonalinka	SKA	female	1973	Bangladesh	4.94
Jupateco F73	JUP	female	1973	Mexico	6.02
Yecora rojo 76	YRR	female	1975	USA	3.24
UP 262	UP	female	1978	India	4.51
Debeira	DEBA	female	1982	Sudan	5.68
HUW 234	HUW	female	1984	India	5.41
Prointa Federal	PIFED	female	1989	Argentina	5.37
Nesser	NESSER	female	1990	Jordan	5.43
BAW 898	BAW	female	1996	Bangladesh	5.29
Alucan/Ducula[†]	AL1	female	-	Mexico	4.44
Chum18/5*BCN[†]	AL2	female	-	Mexico	6.23
Tilhi	TIL	female	2003	Mexico	7.00
Heilo	HEI	female	2003	Mexico	5.55
Kalyansona	KAL	male	1967	India	5.41
Pavon F 76	PVN	male	1976	Mexico	6.18
HD2329	HD	male	1985	India	5.26
Inqalab 91	IQB	male	1991	Pakistan	5.40
Tobarito M97	TOB	male	1993	Mexico	5.89
Baviacora M92	BAV	male	1997	Mexico	7.38
SW89.5124*2/Fasan[†]	SW89	male	-	Mexico	4.06
SW90.1057	SW90	male	-	China	4.94

[†] Advanced breeding lines from CIMMYT.

Field trials

The 112 hybrids, duplicate entries of the parents, and two local commercial checks (RAYON F89 and KAMBARA 1) were sown in yield trials arranged as randomized latinized 16 × 10 α-lattice designs with two replications. The trials were conducted near Ciudad Obregon, Sonora, in Northern Mexico during the winter season of 2002-2003, and the CIMMYT headquarter's station at El Batan in the central Mexican highlands, State of Mexico, during the 2003 summer season. The experimental plots consisted of four rows of 3 m length and were sown at a constant plant density of 200 seeds m^{-2} (corresponding to an

average seeding rate of 50 kg ha^{-1}). Grain yield (t ha^{-1}), plant height (cm), days to flowering and maturity in days after sowing were recorded on a plot basis.

Statistical analyses

For the grouping of germplasm, COP values of the parents were determined with fully expanded genealogical information extracted from the CIMMYT database IWIS version 4 (Payne et al., 2002). Rogers' (1972) distances (RD) based on the 113 SSR markers were calculated for all pairwise combinations of lines. Standard errors for RD estimates were obtained by using a bootstrap procedure with re-sampling over markers. A principal coordinate analysis (PCoA) was performed to group the parents based on RD estimates (Gower, 1966). Classification of the parents revealed by PCoA was used to determine the effect of genetic distance on intra- and inter-group hybrids.

For each hybrid the mid-parent value (MP), absolute mid-parent heterosis (MPH), relative mid-parent heterosis (MPH%), and relative better parent heterosis (BPH) were calculated as follows: MP = (P1 + P2)/2; MPH = F1P – MP; MPH% = (MPH/MP) × 100; BPH% = (F1P – Pb)/Pb × 100, where P1 and P2 are the parents of the hybrid, F1P the hybrid performance, and Pb the higher yielding, earlier, or taller parent. Mid-parent heterosis was tested for significance by an ordinary t-test. General (GCA) and specific combining ability (SCA) effects were estimated for all traits according to established methods (Simmonds, 1979).

Combined analyses of variance were performed separately for the parents and hybrids for all traits, considering all effects as random. Sums of squares of hybrids were further partitioned into GCA and SCA effects. Because the classification of lines as male or female was arbitrary, GCA variance components were pooled with the following weights: $\sigma^2_{GCA} = 0.485\ \sigma^2_{GCA\ male} + 0.515\ \sigma^2_{GCA\ female}$. A corresponding subdivision was conducted on the hybrid × environment sums of squares. Parents and hybrids mean squares were tested for significance by *F*-tests using the corresponding interaction mean squares with environments. Parents × environment, hybrids × environment, GCA × environment, and SCA × environment mean squares were tested for significance by using the corresponding pooled error mean square in the denominator. From the genotypic and phenotypic

variances among parents and hybrids, broad-sense heritabilities on an entry-mean basis were calculated.

Pearson correlation coefficients (*r*) were calculated for F1P with MP, GCA with line *per se* performance (LP), RD with COP, as well as for MPH and F1P with RD and COP. Analyses of variance were performed with software packages SAS (1989) and PLABSTAT (Utz, 1993). The PCoA was conducted using the software PLABSIM (Frisch et al., 2000), which is implemented as an extension of the statistical software R (Ihaka and Gentleman, 1996).

RESULTS

SSR and COP data

The 113 SSRs amplified a total of 420 alleles across all parents. RD estimates ranged from 0.26 to 0.62, with a mean of 0.46. Pedigrees of the two Chinese parents SW89 and SW90 were not available and, therefore, could not be considered for the estimation of COPs. For the remaining 20 parents, COP values ranged from 0.06 to 0.35, with a mean of 0.17. The correlation between RD and COP values was significant ($P < 0.05$) but low ($r = 0.24$).

In the PCoA based on RD estimates, the first three principal coordinates (PC) explained 11.4%, 9.8%, and 7.6% of the total variation, respectively (Fig. 1). The parents clustered in three distinct groups. Group I comprised parents representing mainly elite cultivars developed during the last two decades, Group II was composed of cultivars mainly released in the 1960s and 1970s. Group III involved the two Chinese parents.

Hybrid performance and heterosis

Four hybrids produced with Chinese line SW90 expressed hybrid necrosis and were excluded from subsequent analyses. Hybrids showed a significantly ($P < 0.05$) higher grain yield and plant height than the corresponding MP, combined with earlier flowering and maturity. MPH% for grain yield was on average 0.5% and ranged from -15.33% for hybrid TIL × KAL to 14.13% for the top yielding hybrid BAW × BAV (Table 2). Eighteen of the
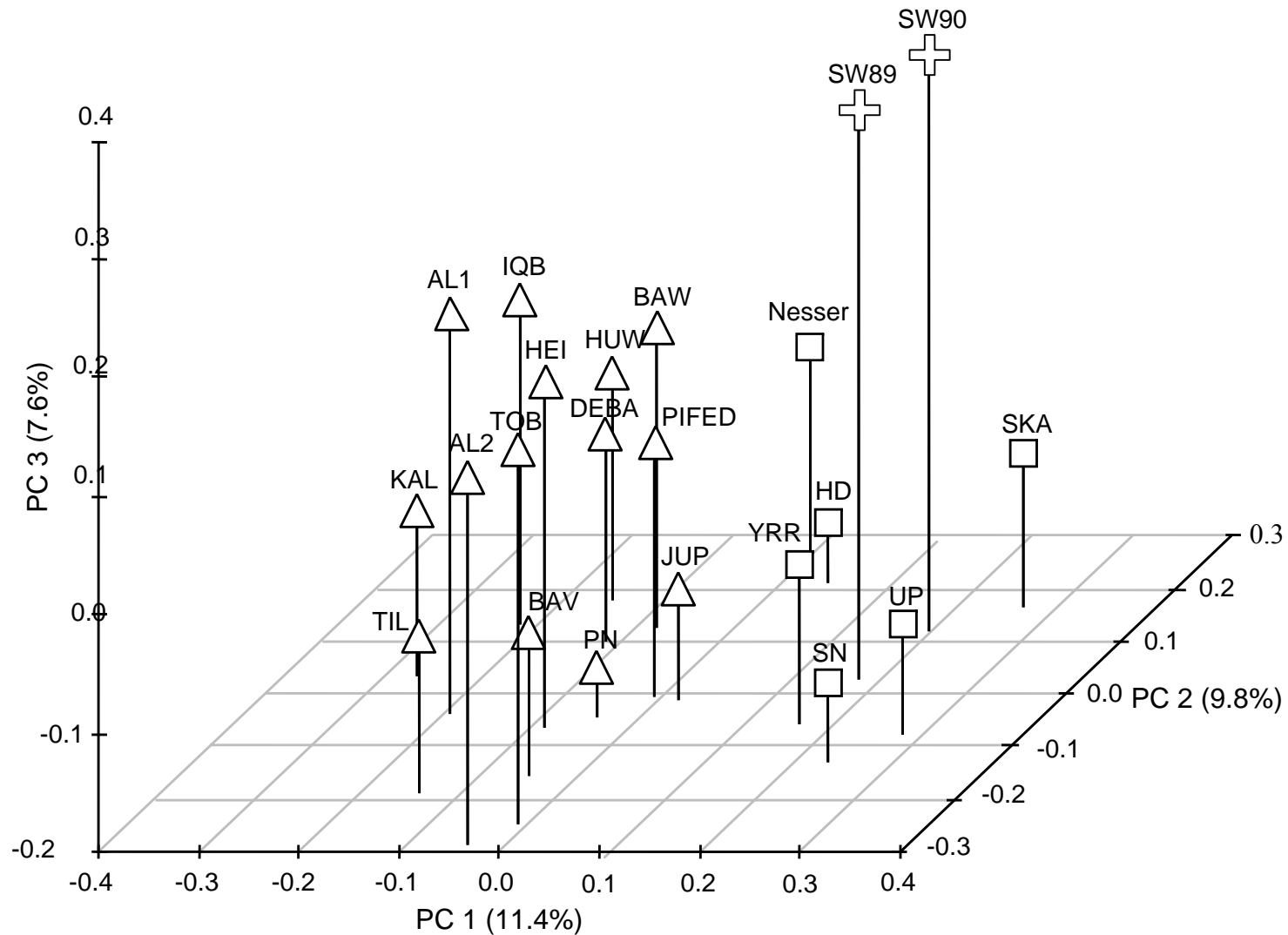
Figure 1. Principal coordinate analysis of 22 spring bread wheat lines performed with Rogers' distance estimates calculated from 113 SSRs. Abbreviation of the line names refer to Table 1. Different groups of germplasm are designated with triangles (Group I), squares (Group II), and crosses (Group III).

108 hybrids outyielded the better parent, but the differences between the hybrid and the better parent were not significant, and the average BPH was negative.

Table 2. Mid-parent value (MP), hybrid performance (F1P), absolute (MPH) and relative mid-parent heterosis (MPH%) and relative better-parent heterosis (BPH%) observed for four agronomic traits of 22 spring bread wheat lines and 108 wheat hybrids evaluated at two environments in Mexico.

Statistic	Grain yield t ha^{-1}	Plant height cm	Flowering date days	Maturity date days
MP				
Mean[†]	5.41	80.8	70.7	108.0
Min.	3.64	65.1	63.7	98.9
Max.	7.20	89.6	78.1	113.2
LSD$_{5\%}$	0.82	7.7	3.9	5.2
F1P				
Mean	5.44	81.7	69.3	107.5
Min.	3.44	60.7	61.5	100.8
Max.	7.32	95.7	77.7	112.9
LSD$_{5\%}$	0.98	2.4	3.9	1.7
MPH				
Mean	0.02*	0.9**	-1.4**	-0.4**
Min.	-0.80	-8.3	-7.6	-7.5
Max.	1.00	11.1	4.5	6.9
LSD$_{5\%}$	0.82	10.1	3.9	6.6
MPH (%)				
Mean	-0.02	0.8	-2.1	-0.4
Min.	-15.33	-13.7	-11.8	-7.5
Max.	14.13	12.21	6.1	6.5
BPH (%)				
Mean	-9.3	6.5	1.8	2.8
Min.	-37.26	-5.0	-7.8	-6.4
Max.	14.12	28.0	11.9	17.2

*,** Significant at the 0.05 and 0.01 probability levels, respectively.
[†] Values represent averages across environments.

Intra-group hybrids outperformed inter-group hybrids for grain yield (Fig. 2) and other traits (data not shown). On average, F1P for grain yield was largest in intra-group hybrids of Group I and smallest in inter-group hybrids involving the two Chinese parents of Group III.
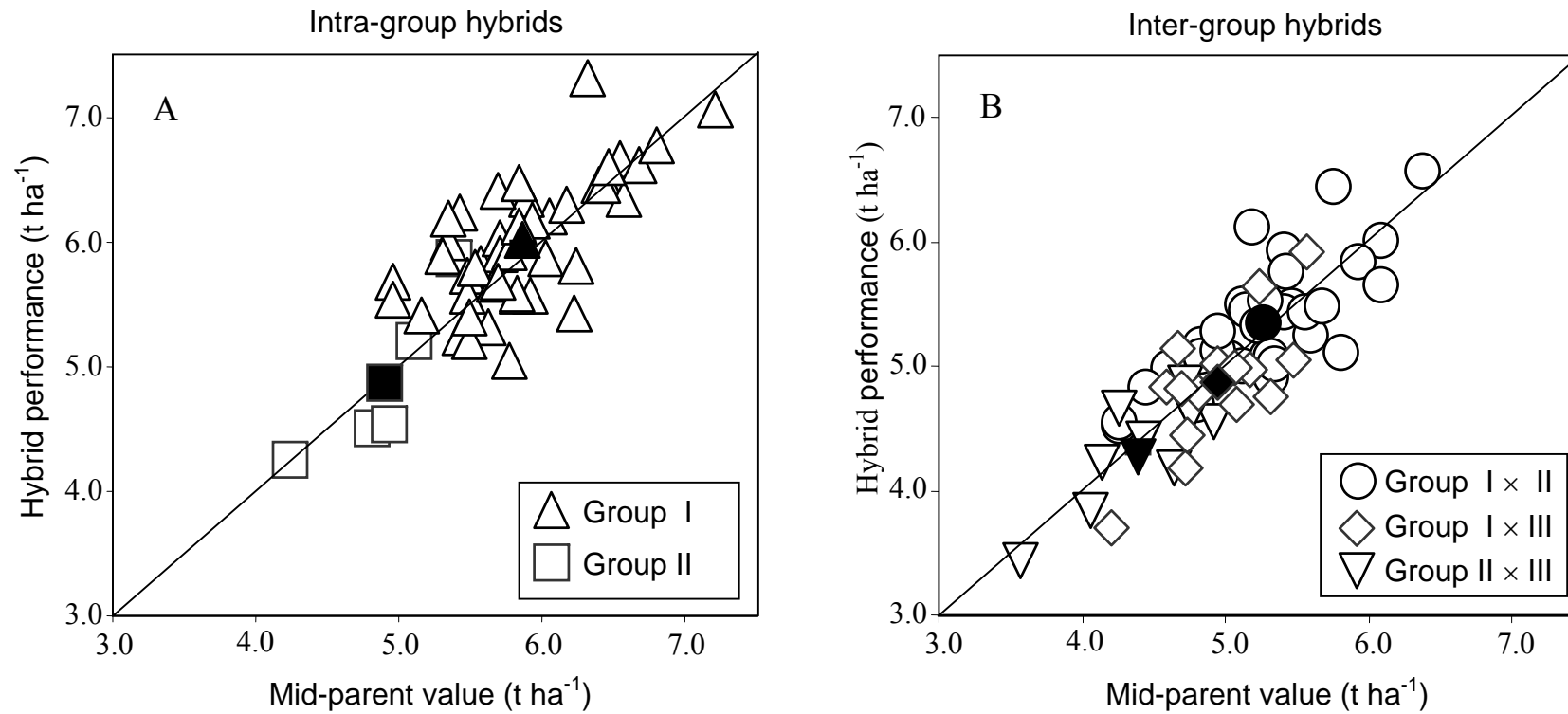
Figure 2. Hybrid performance *vs*. mid-parent value for grain yield of (A) intra-group and (B) inter-group hybrids evaluated in two environments in Mexico. Groups I to III refer to Fig. 1 of this study and are based on groupings of 22 spring bread wheat breeding lines determined by principal coordinate analysis. Group means are designated by the corresponding filled symbols.

The combined analyses of variance revealed significant (P < 0.05) genetic variation among parents and among hybrids for all traits (Table 3). Estimates of genotypic variance were twice as large among the parents than among the hybrids for all traits. Estimates of GCA and SCA variances were significant (P < 0.05) for all traits, with the former being three to four times greater than the latter. The largest GCA effect for grain yield was observed for the male parent BAV (0.92 t ha^{-1}), the lowest for female YRR (-0.96 t ha^{-1}).

Table 3. Estimates of variance components (σ^2) and heritability (h^2) for four agronomic traits determined from the combined analyses of variance of 22 spring bread wheat lines and 108 hybrids evaluated at two environments in Mexico.

Variance components[†]	df	Grain yield t ha^{-1}	Plant height cm	Flowering days	Maturity days
Parents					
σ^2_G	21	0.71**	74.1**	9.4*	18.5**
$\sigma^2_{G \times L}$	21	0.33**	13.6**	16.1**	7.7
h^2		0.81	0.92	0.54	0.83
Hybrids					
σ^2_{GCA}	20	0.20**	18.9**	4.5**	2.8**
σ^2_{SCA}	87	0.04**	3.0**	0.6*	0.9*
$\sigma^2_{GCA \times L}$	20	0.86**	2.3**	0.9	1.9**
$\sigma^2_{SCA \times L}$	87	0.11	6.8	2.3	2.8
h^2		0.78	0.82	0.87	0.67

*,** Significant at the 0.05 and 0.01 probability levels, respectively.
[†] G = genotype; L = environment, GCA = general combining ability, SCA = specific combining ability.

Relationships of MPH and F1P with genetic distance

Hybrid performance was significantly (P < 0.01) correlated with MP for all traits (Table 4). Correlation coefficients were highest for plant height and grain yield (*r* = 0.86) and lowest for days to maturity (*r* = 0.67). A tight association existed also between GCA effects and LP for all traits. The correlation between F1P and COP was not significant for grain yield, flowering and maturity. Between F1P and RD the correlation was significantly negative but of low magnitude. The correlations between MPH and COP or RD were generally of small magnitude and not significant for most traits. For Group I comprising

the most promising parents for hybrid production, the correlation between MPH and RD was also low and not significant ($r = 0.06$).

Table 4. Correlation coefficients of hybrid performance (F1P) with mid-parent performance (MP), general combining ability (GCA) with line *per se* performance (LP), as well as F1P and mid-parent heterosis (MPH) with coefficient of parentage (COP) and Rogers' distance (RD) based on 113 SSRs for four agronomic traits evaluated in wheat lines and hybrids grown at two environments in Mexico.

r (x, y)		Grain	Plant	Flowering	Maturity
x	y	yield	height	date	date
F1P	MP[†]	0.86**	0.86**	0.81**	0.67**
GCA	LP	0.91**	0.95**	0.78**	0.88**
F1P	COP	-0.05	-0.21*	-0.18	-0.20
F1P	RD	-0.38**	-0.21*	-0.32**	-0.45**
MPH	COP	-0.03	-0.31**	-0.14	-0.13
MPH	RD	-0.04	-0.07	-0.31**	-0.05

*,** Significant at the 0.05 and 0.01 probability levels, respectively.
[†] Correlation coefficients based on means across environments.

DISCUSSION

SSR diversity and grouping of germplasm

The degree of polymorphism observed across the 22 lines was similar to the findings of Ahmad (2002) and Bryan et al. (1997), who analyzed 13 and 10 wheat cultivars, respectively. We had expected a somewhat higher degree of allelic variation in the present study, given the broad genetic base of the genotypes evaluated and the marker system used. However, the EST-SSR markers employed generally revealed a much lower degree of polymorphism than the genomic- derived SSRs used in other studies (Eujayl et al., 2002).

Principal coordinate analysis based on RD estimates confirmed the power of SSR markers to group breeding germplasm according to common ancestry and geographic origin. The separate grouping of the two Chinese parents was expected because wheat breeding programs in China have operated independently and under specific environmental conditions until the mid-1980s, making their germplasm quite unique compared to materials from other countries (He et al., 2001). The grouping of the remaining lines

reflected the chronological evolution through breeding, as Group I was composed mostly of more recently bred, high yielding lines, and Group II of the earlier released semi-dwarf wheats.

No clear grouping of the parents based on COPs was observed. COP values represent only an indirect measure of genetic diversity. The diverse genetic background of the parents resulted in a small range of COP values, which complicated the classification of the parents, whereas a large variation in the RD was observed. Furthermore, calculating COPs is based on simplifying assumptions regarding the relatedness of ancestors, parental contribution to the offspring, and absence of selection and genetic drift, most of which are not met under breeding conditions (Cox et al., 1985).

Exploitation of heterosis in wheat

The major goal in hybrid breeding is the exploitation of heterosis. While a large number of studies found significant heterosis in wheat, there is considerable variation in its magnitude. High MPH% of 92% for grain yield was reported by Walton (1971) and up to 46% by Bailey et al. (1980). In recent papers, lower levels of MPH% for grain yield were measured, which rarely exceeded 10% (Liu et al., 1999; Oury et al., 2000). Cukadar et al. (1999) evaluated 260 hybrids including CIMMYT advanced lines and reported BPH% between 3.5% and 6.5%. In the present study, MPH% for grain yield was on generally small and average BPH% was even negative. It is striking that most of the earlier studies reporting high levels of heterosis employed low seeding densities. With small numbers of space-sown plants, heterosis estimates are most likely inflated due to a higher degree of tillering (Pickett, 1993). With the discovery of chemical hybridizing agents, it became feasible to produce sufficient quantities of F_1 seeds and to test a large number of hybrid combinations at normal seeding densities, which gives more realistic estimates of heterosis under common agricultural practices (Pickett and Galwey, 1997). Owing to hand emasculation, the seeding rate was reduced in our study, but it was still close enough to commercial seeding rates to avoid strong effects on the yield levels of either the hybrids, parental lines or checks.

If dominant, as opposed over dominant, gene action is the basis of heterosis in wheat, the advantage of heterozygosity can be eliminated by fixing favourable alleles in pure line

varieties. Furthermore, since wheat is an allopolyploid species with three ancestral genomes, substantial benefits from a high degree of intergenomic heterosis are expected. When we exclusively considered hybrids produced from parents released during the last decade, 50% outyielded the commercial pure line check RAYON F89, but none of them outperformed KAMBARA 1, the most recently released check. Thus, the hybrid advantage over lines applied only to the parents but not to the newest released line cultivar.

The relative amount of heterosis also depends on environmental factors. MPH% for grain yield has been found to be smaller under optimum than under stress conditions in maize and sorghum (Betrán, et al., 2003). This could be a further explanation for the low levels of MPH% observed in our study, because common agronomic practices were applied (e.g., full irrigation, complete fertilization, fungicide treatment) for achieving high yields.

Earliness is a desirable breeding goal, and hence the presence of negative heterosis for days to flowering and maturity, observed in this study, would favor hybrids. On the other hand MPH% for plant height was significant but on average only 0.9%. This would be of no concern for the release of wheat hybrids, because most of them fell within the range of their parents.

Combining ability of parental lines and implications for hybrid breeding

The choice of parental combinations yielding superior hybrids is the most important aspect in hybrid breeding. Analysis of the relative importance of GCA and SCA effects provides an indication of the type of gene action involved in the expression of traits and allows inferences about optimum allocation of resources in hybrid breeding. In accordance with earlier studies in wheat, GCA variances were more important than SCA variances, indicating the predominance of additive effects. Theoretical and experimental results in maize show that SCA effects are of primary significance in intra-group crosses, whereas GCA effects are predominate in inter-group crosses (Melchinger and Gumber, 1998). Our findings in wheat are in contrast with these results for unknown reasons.

The tight correlations of GCA with LP and MP with F1P (Table 4) suggest that the probability to obtain superior hybrids is greater by crossing the highest yielding parents. Thus, hybrid wheat breeding should be relatively efficient based only on selection for

parental performance and a relative small number of testcrosses involving outstanding parental lines. Although the correlations depend on the material studied and were often reported to be lower in other studies, the use of MP as a predictor for F1P was suggested earlier owing to its simple assessment and because reliable information about LP is readily available from line breeding programs (Oury et al., 2000).

Relationship between MPH and diversity measures

Quantitative genetic theory suggests a linear correlation between MPH and the squared modified Rogers' distance under certain simplifying assumptions (Falconer and Mackay, 1996). Because the parents were homozygous lines in our study, the squared modified Rogers' distance corresponds to the RD (Melchinger et al., 1993). Contrary to expectations, however, the observed correlation between MPH and RD was low for all traits. Four possible explanations are (Reif et al., 2003): (i) a poor association between heterozygosity estimated from the SSR data and heterozygosity at quantitative trait loci controlling the trait, (ii) a lack of association between heterozygosity and heterosis at quantitative trait loci in the crosses examined, (iii) existence of multiple alleles with similar effects on a given trait, and (iv) epistasis among the respective quantitative trait loci. Providing the latter two factors are absent, Melchinger (1999) pointed out that a correlation between MPH and molecular distance was more likely to be found in intra-group crosses than in inter-group crosses. This could not be confirmed in our study, presumably due to the lower levels of heterosis observed in inter-groups than in intra-group crosses.

A decrease in MPH and F1P in extremely wide crosses with large RDs, as observed in hybrids with the Chinese parents from Group III, was also found in crosses between tropical and U.S. maize populations (Moll et al., 1965). The authors attributed this nonlinear relationship between geographic distance and heterosis to the lack of co-adaptation between both allelic and non-allelic combinations from the two parental genomes, which resulted in negative dominance and negative epistatic effects, respectively (Falconer and Mackay, 1996).

Prospects of hybrid breeding in wheat

In light of the low level of MPH% and BPH% observed in the present and other studies, the successful dissemination of hybrids in wheat is rather questionable, especially if the costs of hybrid seed production remain high. Agnus (1997) concluded that a yield advantage of 5% over the best conventional variety is required to compensate for the higher seed costs associated with male emasculation and cross fertilization, and to justify the additional expenses in breeding of hybrid wheat. Pickett and Galwey (1997) argued that 6 to 34% of MPH% is necessary to make wheat hybrids commercially viable. Thus, our estimates of MPH% and BPH% as well as the F1P relative to the checks cannot be considered encouraging for large-scale development and global acceptance of hybrid wheats.

Owing to their better vigor, robustness and stress tolerance compared with pure line cultivars under more marginal conditions, the development of hybrids may be justified in such environments where, in addition, low seeding rates are used (Jordaan et al., 1999). Biotechnological approaches including the exploitation of apomixis might nurture hopes to facilitate seed production and make hybrid wheat an attractive alternative for niche environments in the future. Finally, genetic distances based on SSRs cannot be considered a promising tool to predict hybrid performance, but could be a powerful tool for identification of divergent groups in advanced wheat breeding materials.

ACKNOWLEDGEMENTS

REFERENCES

Agnus, J.W. 1997. Hybrid wheat – a dream or reality? Aspects of Applied Biology 50:15-22.

Ahmad, M. 2002. Assessment of genomic diversity among wheat genotypes as determined by simple sequence repeats. Genome 45:646-651.

Bailey T.B., C.O. Qualset, and D.F. Cox. 1980. Predicting heterosis in wheat. Crop Sci. 20:339-342.

Betrán, F. J., D. Beck, M. Bänziger, and G.O. Edmeades. 2003. Genetic analysis of inbred and hybrid grain yield under stress and nonstress environments in tropical maize. Crop Sci. 43:807-817.

Brummer, E.C. 1999. Capturing heterosis in forage crop cultivar development. Crop Sci. 39:943-954.

Bryan, G.J., A.J. Collins, P. Stephenson, A. Orry, J.B. Smith, and M.D. Gale. 1997. Isolation and characterization of microsatellites from hexaploid bread wheat. Theor. Appl. Genet. 94:557-563.

Chacrosset A. and L., Essioux. 1994. The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. Theor. Appl. Genet. 89:336-343.

CIMMYT. 2001. The Applied Biotechnology Center's Manual of Laboratory Protocols. 1st ed. (Available online at http://www.cimmyt.cgiar.org/ABC/Protocols/manual ABC.html), CIMMYT, Mexico, D.F, Mexico.

Corbellini M., M. Perenzin, M. Accerbi, P. Vaccio, and B. Borghi. 2002. Genetic diversity in bread wheat, as revealed by coefficient of parentage and molecular markers, and its relationship to hybrid performance. Euphytica 123:273-285.

Cox, T.S., Y.T. Hkiang, M.B. Gorman, and D.M. Rodgers. 1985. Relationships between coefficient of parentage and genetic indices in the soybean. Crop Sci. 25:529-532.

Cukadar, B., M. van Ginkel, D. Dunphy, and S. Rajaram. 1998. Hybrid wheat research at CIMMYT using genesis hybridizing agent. pp. 31-36. *In* Zhang, A., H. Tiecheng (eds.) Hybrid wheat. A new crop going to farmers. Proceedings of the International Workshop on Hybrid Wheat, China Agricultural University, China.

Cukadar, B. M. van Ginkel, S. Rajaram, and G. Ortiz Ferrara. 1999. Yield Increase through hybrid wheat. CIMMYT's on-going efforts. *In* Proc. of the Hybrid Wheat Discussion Group (DWR), 17-18. September 1999, Karnal, India.

Devos, K.M., G.J. Bryan, A.J. Collins, P. Stephenson, and M.D. Gale. 1995. Application of two microsatellite sequences in wheat storage proteins as molecular markers. Theor. Appl. Genet. 90:247-252.

Dreisigacker, S., P. Zhang, M. van Ginkel, M. Warburton, D. Hoisington, M. Bohn, and A.E. Melchinger. 2004. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different mega-environments. Crop Sci. In press.

Eujayl, I., M.E. Sorrells, M. Braun, P. Wolters, and W. Powell. 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor. Appl. Genet. 104:399-407.

Falconer, D.S and T.F. Mackay. 1996. Introduction to Quantitative Genetics. 4th ed. Longman Group, London.

Frisch, M., M. Bohn, and A.E. Melchinger. 2000. Plabsim. Software for simulation of marker-assisted backcrossing. J. Hered. 91:86-87.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325-388.

Hallauer, A.R., W.A. Russell, and K.R. Lamkey. 1988. Corn breeding. pp.532-540. *In* Spraque, G.F., J.W. Dudley (eds.) Corn and Corn Improvement. 3rd ed. Agron. Monogr. 18, ASA, CSSA, and SSSA, Madison, WI.

He, Z.H., S. Rajaram, Z.Y. Xin, and G.Z. Huang. 2001. pp. 5-8. A history of wheat breeding in China. CIMMYT, Mexico, D.F, Mexico.

Heisey, P.W., M.A. Lantican, and H.J. Dubin. 2002. pp. 4. Impacts of international wheat breeding research in developing countries, 1966-1997. CIMMYT, Mexico, D.F, Mexico.

Ihaka, R. and R. Gentleman. 1996. R. A language for data analysis and graphics. J. Comput. Graph. Stat. 5:299-314.

Jordaan, J.P., S.A. Engelbrecht, J.H. Maran, and H.A Knobel. 1999. Wheat and heterosis. pp. 441-421. *In* Coors J.G. and S. Pandey (eds.) The genetics and exploitation of heterosis in crops. ASA-CSSA, Madison, WI.

Lang, L. 1989. Hybrid wheat and its perspectives. Sveriges Utsadesforenings Tidskrift 99:121-124.

Liu, Z.-Q., Y.Pei, and Z.-J. Pu. 1999. Relationship between hybrid performance and genetic diversity based on RAPD markers in wheat, *Triticum aestivum L.* Plant Breeding 118:119-123.

Melchinger, A.E. 1993. Use of RFLP markers for analysis of genetic relationship among breeding materials and prediction of hybrid performance. pp 621-628. *In* Buxton, D.R., R. Shibles, R.A. Forsberg, B.L. Blad, K.H. Asay, G.M. Paulson, and R.F. Wilson (eds.) International Crop Science I. CSSA, Madison, WI.

Melchinger, A.E. 1999. Genetic diversity and heterosis. pp. 99-117. *In* Coors J.G. and S. Pandey (eds.) The genetics and exploitation of heterosis in crops. ASA-CSSA, Madison, WI.

Melchinger, A.E., and R.K. Gumber. 1998. Overview of heterosis and heterotic groups in agronomic crops. pp 29-43. *In* Lamkey, K.R., and J.E. Staub (eds.) Concepts and breeding of heterosis in crop plants. CSSA Special Publ. No. 25, CSSA, Madison, WI.

Moll, R.H., J.H. Longquist, J.V. Fortuna, and E.C. Johnson. 1965. The relationship of heterosis and genetic divergence in maize. Genetics 52:139-144.

Oury, F.X., P. Brabant, P. Berard, P. Pluchard. 2001. Predicting hybrid value in bread wheat. biometric modelling based on a "top-cross" design. Theor. Appl, Genet. 100:96-104.

Payne, T.S., B. Skovmand, C.G. Lopez, E. Brandon, and A. McNab (eds.). 2002. The international wheat information system (IWIS™), 4th ed., Compact disk., CIMMYT Mexico, D.F, Mexico.

Pickett, A.A. and N.W. Galwey. 1997. A further evaluation of hybrid wheat. Plant Varieties and Seeds 10:13-32.

Reif, J.C., A.E. Melchinger, X.C. Xia, M.L. Warburton, D.A. Hoisington, S.K. Vasal, G. Srinivasan, M. Bohn, and M. Frisch. 2003. Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. Crop Sci. 43:1275-1282.

Reynolds, M.P., J. van Beem, M. van Ginkel, and D. Hoisington. 1996. Breaking the yield barriers in wheat. A brief summery of the outcomes of an international consultation. pp. 1-10. *In* Reynolds, M.P., S. Rajaram, and A. NcNAb (eds). Increasing yield potential in wheat. Breaking the yield barriers. CIMMYT, Mexico, D.F, Mexico.

Röder, M., K. Wendehake, V. Korzun, G. Bredemeijer, D. Laborie, B. Bertrand, P. Isaac, S. Rendel, J. Jackson, R.J. Cooke, B. Vosman, and M.W. Ganal. 2002. Construction and analysis of a microsatellite-based database of European wheat varieties. Theor. Appl. Genet. 106:67-73.

Rogers, J.S. 1972. Measures of similarities and genetic distances. pp.145-153. *In* Wheeler M.R. (ed) Studies in genetics VII. University of Texas, Austin.

SAS Institute. 1988. SAS language guide for personal computers. Release 6.03. ed., SAS inst. Cary, NC.

Simmonds, N.W., 1979. Principles of crop improvement. Longman, London.

Souza, E., P.N. Fox, D. Byerlee, and B. Skovmand. 1994. Spring wheat diversity in irrigated areas of two developing countries. Crop Sci. 34:774-783.

Utz, H.F. 1993. PLABSTAT. A computer program for the statistical analysis of plant breeding experiments, Univ. of Hohenheim, Germany.

Walton, P.D. 1971. Heterosis in spring wheat. Crop Sci. 11:422-42.

6 GENERAL DISCUSSION

Trend of genetic diversity in CIMMYT wheat germplasm

Over the past century, the development and successful application of plant breeding methods has produced high-yielding crop varieties upon which modern agriculture is based. New varieties are usually bred by crossing a set of genetically related modern varieties, followed by an intensive selection in succeeding generations. Primitive ancestors and wild relatives are thereby rarely considered. Although they have a diverse genetic base, their agronomic performance is generally low. Consequently, the genetic variation of crop plants is continued to be reduced by plant breeding (Tanksley and McCouch, 1997). It is the plant breeding process itself that threatens the genetic base on which breeding depends. In the light of these developments, the main objective of our research was to examine the genetic diversity in modern wheat breeding materials and genetic resources provided by and stored at CIMMYT.

Several studies have determined the trend of genetic variation in plant breeding programs in wheat. Donini et al. (2000) and Manifesto et al. (2001) measured the genetic diversity based on SSR and AFLP markers in wheat germplasm of the UK and Argentina. They reported that the overall level of genetic variation in the material investigated has been maintained during the last half century. Smale et al. (2002) calculated COPs and found no decrease in the diversity of 27 CIMMYT wheat lines released between 1965 and 1990. A more extensive study of Reif et al. (2004) with SSRs and COPs observed a decline of genetic diversity in 123 CIMMYT major wheat cultivars from 1950 to 1989, followed by an increase of genetic diversity since the 1990s. The loss of genetic diversity in CIMMYT wheat germplasm after the 1950s can be explained by numerous releases of semi-dwarf cultivars, which were derived from a limited number of progenitors, and rapidly replaced tall cultivars and LCs (Dalrymphe, 1986). In the late 1970s, CIMMYT breeders became aware of the risk of genetic vulnerability caused by the effect of genetic erosion and the wide cultivation of a few favored cultivars. CIMMYT changed its breeding strategy and aimed to broaden the genetic base on a large scale (Reeves, 1999; Smale et al., 2002). The exploitation of a spring × winter wheat crosses is one example that has resulted in higher yield gains in the bread wheats and increased levels of diversity. Furthermore, the germplasm exchange of CIMMYT with national agricultural research programs from

Brazil and China enabled to incorporate important traits such as the resistance to aluminium toxicity, head scab and karnal bunt (Rajaram, 1995). When examining different categories of wheat germplasm in our research, our findings corroborated the success of CIMMYTs' strategies to enhance the genetic diversity of their breeding materials. We observed high levels of genetic diversity in CIMMYT advanced breeding lines as well as in newly developed SHWs and in wheat LCs, which are considered valuable sources for the introgression of genes into improved wheat materials.

Genetic variation via CIMMYT's ME concept

In a recent study, Evenson and Gollin (2003) evaluated the global impact of varieties bred by International Research Centers over two defined periods, an early (1960 to 1980) and late (1981 to 2000) Green Revolution period. The authors reported that the impact of varieties on productivity gains in developing counties was substantially higher in the late than early Green Revolution period owing to the expansion of breeding activities on new agroecological regions. Improved yield potential of varieties targeted to more "marginal" environments allowed to increased food production with only modest increases in area planted and with a relatively slow growth in the use of inputs.

CIMMYT introduced its concept to breed wheat specifically adapted to different areas in the early 1980s. Since that time, all crosses are directed to specific MEs, and ME relevant requirements are being taken in account during the selfing generations. The ME concept seems to be somehow in contrast to CIMMYTs methodology of shuttle breeding implemented by N. Borlaug in 1944. Breeding germplasm is thereby shuttled between two environmentally contrasting locations in Mexico, which permits pyramiding valuable genes for improved germplasm stability. However, according to our results, these two methods lead to the presence of a single core germplasm in which the same set of genes may confer fitness to several MEs. Genetic diversity was mainly observed within than among lines targeted to specific MEs. Although the diversity estimates might not be directly associated with yield potential, they demonstrate that in each ME the possibility for further selection on a very high breeding level is still given. Therefore, the ME concept can be considered as one contribution to increase the genetic diversity in CIMMYT wheat germplasm.

The progress in improving yield potential and rates of adaptation to marginal regions increased markedly in the recent years, as reflected by the current cultivation of CIMMYT advanced breeding lines in more than 60 different countries (Van Ginkel et al., 2002). The MEs could be classified in more detail using additional climate and crop production data provided by newly developed geographic information systems (GIS). These systems help breeders to more efficiently reach breeding objectives in any defined region.

Exploitation of genetic diversity via SHWs and LCs

The genetic enhancement of germplasm pools in breeding programs has largely resulted from the wider utilization of plant genetic resources (Rajaram and Van Ginkel, 1996). Bread wheat offers an enormous potential for the incorporation of genetic resources, because its allopolyploid nature enables the introgression of genes from wild species by recombination of the homoeologous chromosomes (Valkoun, 2001). Excellent sources represent particularly *T. durum* and *T. tauschii*, which are the most closely related wild ancestors. Due to the small number of independent crosses between *T. durum* and *T. tauschii* during the evolution of *T. aestivum*, the contribution of the total diversity from these species were small (Dvorak et al., 1998). Estimates of RFLP diversity at *RbcS* indicated that bread wheat contains perhaps 30% of the diversity levels found in its diploid relatives (Galili et al., 2000). With the development of SHWs, which repeat the natural hybridization event, genetic diversity of *T. durum* and *T. tauschii* can be incorporated into modern breeding lines.

Another category of useful genetic resources represent wheat LCs. Bread wheat has been cultivated since its hybridization. The early geographic dispersal of bread wheat in the evolution of cereals may have contributed to its adaptability across many different climates. Wheat LCs are therefore rich sources for valuable quantitative traits.

Important factors regarding the exploitation of genetic resources are (i) whether they offer beneficial alleles for increased yield potential and (ii) how these alleles can be transferred into improved materials. SHWs have shown to be a valuable source for resistance to several diseases (karnal bunt: Villareal et al., 1996; stripe rust: Ma et al., 1995), tolerance to abiotic stresses such as cold temperature and saline soils (Gorham, 1990; Limin and Fowler, 1993), and yield components such as kernel weight (Del Blanco

et al., 2001). SHWs can easily be crossed with elite materials. They serve as a bridge between cultivated bread wheat and its wild ancestors, which is a great advantage for the transfer of desired genes. A limited number of backcrosses to the recurrent elite bread wheat parent already gives acceptable performance of SBLs. Del Blanco et al. (2000, 2001) reported improved performance of BC_2F_2 derived SBLs compared to their recurrent parents for several important agronomical and physiological traits as well as positive transgressive segregation. Large genetic diversity and the separate grouping of the germplasm from elite materials make SHWs even more attractive in the future for the genetic enhancement of breeding pools.

Wheat LCs have proven to be tolerant to many abiotic and biotic stresses (Dubin et al., 1997; Hede et al., 1999). However, in comparison to SHWs, wheat LCs are more difficult to incorporate into elite and local adapted materials due to their widely differing phenotypes making the search for useful characteristics rather complex (Hoisington et al., 1999). This was confirmed by our results showing higher genetic diversity observed within rather than between LCs.

Diversity studies of genetic resources enable to determine unique genotypes that are likely to contain novel alleles, some of which may be of agronomic use. Therefore, these studies build the basis of molecular mapping techniques such as advanced backcross QTL analysis (Tanksley and Nelson 1996) or marker-trait associations (Buckler et al., 2002). Advanced backcross QTL analysis is a method for the simultaneous identification and transfer of valuable QTLs from genetic resources into advanced breeding germplasm. Unadapted materials are backcrossed to elite lines two to three generations before QTL for desired traits are identified and used for marker-assisted selection. Despite many successful examples reported in the literature (wheat: Huang et al., 2003, barley: Pillen et al., 2003, maize: Ho et al., 2002), the method is somehow limited because only those loci that have large effects on the quantitative variation can be identified. Furthermore, the resolution of QTL maps is generally limited to 5-10 cM, which might still contain hundreds of genes and subsequent fine mapping is very time consuming. In marker-trait association approaches, allele diversity is evaluated across natural populations, and polymorphisms that correlate with phenotypic variation are identified. Major advantages of this approach include the direct implementation of high resolution genetic maps as well as sequence information of selected candidate genes.

The management of genetic resources

CIMMYT's plant genetic resource center holds the worldwide largest wheat germplasm collection with about 168 000 accessions. Worldwide more than 1300 germplasm collections exist with more than 6 million accessions stored and conserved as seed (FAO, 1996). After this first period of collecting accessions, plant genetic resource centers today have to shift their operations towards improved evaluation and characterization strategies for the better exploitation of accessions. We have seen in our study that seed bank accessions particularly LCs may not be properly preserved and some may not even be catalogued. According to Hede et al. (1999) and DeLacy et al., (2000), data collected on unreplicated hill plots could be used to improve the phenotypic evaluation of accessions. Different agronomic attributes of accessions are thereby examined and provide useful information for potential users and seed bank curators. Genetic diversity studies based on small numbers of molecular markers could already enable a genotypic characterisation of germplasm collections. In combination with a global database, the observed information could be stored and allow the comparison of accessions even between different seed banks. However, lack of sufficient funds limit these activities in most genetic resource centers. In addition, there is still a real threat of genetic erosion to native species in several different countries. In Oman e.g., the total area cultivated with LCs has decreased from 1,000 ha in 1988 to 238 ha in 1996 (Al-Maskri et al., 2003). Large allelic variation has been observed in Omani wheat LCs (Zhang et al., 2003), which emphasizes the need of their maintenance via ex-situ conservation. Skovmand et al. (2003) mentioned critical collection needs for wheat landraces and other types of Triticeae in the western Mediterranean, Central and South America, Eritrea, and Iran. In Guatemala, Honduras, Peru, and Bolivia, wheats may still be grown, which descend from wheats introduced early on by the Spanish. In several other countries such as Eritrea and Iran, wheats and wild relatives are difficult to access at this time for political reasons.

Hybrid wheat

Inspired by the success of hybrids in the case of maize, the production of hybrid wheat has been a breeding objective over the past 40 years through many routes. Research initially centered on the development of cytoplasmatic male sterility (CMS) systems, in

particular those derived from *Triticum timopheevi* (Agnus, 1997). In the early 1980s, the discovery of chemical hybridizing agents (CHAs) stimulated several companies to participate in the production of hybrid wheat on a large scale. Hybrid wheat has been marketed in South Africa, Europe, Australia and the US (Edwards et al., 1994). At CIMMYT, the hybrid wheat breeding program was reinitiated in 1996 in collaboration with the Monsanto Company, producer of the CHA Genesis.

During the past years, however, much of the initial enthusiasm for hybrid wheat has been vanished. According to most recent studies as well as our own results, the difficulties that still remain are (i) the high developmental costs of hybrid wheat and (ii) the low levels of heterosis observed. Those CHAs, which have approached full commercial release since the 1980s, have finally been withdrawn from the market because of toxicity concerns, high costs for licensing, or relatively small volumes needed for the production of hybrid wheat. Thus, private and public programs involved in hybrid breeding research had to re-focus on the use of the more complex and laborious CMS systems. On the female side, several generations of backcrossing are required to incorporate suitable CMS factors. The maintenance of male sterile lines in wheat is additionally unsatisfactory due to the high seeding rate and low seed multiplication factor. On the male side, mostly more than one restorer gene has to be incorporated, but germplasm with good fertility restoration is rare (Pickett and Galwey, 1997).

Positive heterosis for grain yield in wheat has been reported in many studies in the literature. However, the observed levels of heterosis did not lead to significant gains of hybrid wheat over leading standard cultivars and were insufficient to compensate for the higher seed costs. Genetic distance between parents did not play a major role in predicting heterosis. Highest levels of heterosis were mostly observed by crossing parents with high *per se* performance. The economic success of hybrid wheat seems therefore questionable, as reflected by the closing down of the hybrid wheat program in many companies during the past decade.

The genetic gains that are attainable in hybrid wheat have to be evaluated in other factors than improved yield. It is expected that in 2004 about 110 000 ha of hybrid wheat will be cultivated in France due to its improved stability under specific stress conditions (http://www.saaten-union.de/). About 20 000 to 60 000 ha of hybrid wheat are currently grown in South Africa, because the low seeding rates give some advantages for cultivating

F1 plants with improved vegetative growth. The most important incentives for CIMMYT to consider hybrid wheat was its improved tolerance to abiotic stresses in marginal environments, the rapid response to new disease problems, or certain specific end use properties, which are required in specialty markets in developing countries.

The SSR marker system

In terms of genetic diversity studies, SSRs currently represent the most powerful marker technique in wheat. SSRs are relatively uncomplicated in their use because low amounts of DNA are required and the assay can be automated. They are dispersed throughout the genome, codominant and easily transferable between populations. In the literature, high levels of SSR polymorphism have been reported in wheat (Huang et al., 2002; Ahmad, 2002; Röder et al., 2002), which was confirmed in our study applying a set of 113 SSRs. Similar levels of polymorphisms in wheat were only observed with AFLPs (Bohn et al., 1999; Roy, 2004). However, AFLP loci appear not to be evenly distributed across the wheat genome. The choice of primer sets used in generating AFLPs directly influences the distribution of loci. Additionally, AFLPs are costly and complex to use.

The development of SNP markers is still underway in wheat. SNPs have several advantages over SSRs. Previous results in maize, soybean, and rice, have shown higher rates of polymorphisms for SNPs than for SSRs. In addition, many techniques are available to type SNPs in an automated fashion and yield outcomes are revealed (a positive or negative point mutations) that can be easily interpreted (Rafalski, 2002; Gut, 2001). SNPs positioned within or closely related to genes can be easily targeted using existing sequences or EST libraries. However, SNP markers are biallelic and their expected heterozygosity is low. Therefore, they are useful for marker-trait association approaches rather than large-scale genetic diversity studies. Similar to SSRs, the development of SNPs markers, specific for only one of the three genomes in wheat, remains difficult.

Currently developed EST libraries represent also a new valuable source for the development of SSRs. In our study, a total of 52 EST-SSRs were applied. Levels of polymorphism for EST-SSRs were lower than for genomic SSRs due to their high conservation in the genome. Although EST-SSRs are located within transcribed genes, we could not detect functional diversity. The EST-SSRs were probably not linked to genes

causing detectable phenotypic variation. Moreover, it is still unknown whether the polymorphisms of EST-SSR are directly related to differences in gene expression (Vigouroux et al., 2002).

Future directions

Our research demonstrated that SSR-based diversity studies provide many insights into population structures and breeding systems. Disadvantages are that only the genotypic variation of the considered populations could be described. Therefore, current and future research in crops should focus on identifying nucleotide diversity associated with interesting phenotypes to describe functional differences between and within populations.

In association approaches, alleles at a few selected candidate genes may be tested or the whole genome is scanned to identify associations with particular phenotypes (Buckler and Thornsberry, 2002). The minimum number of loci required for association mapping depends on the extent of linkage disequilibrium (LD) and, thus, on the history of recombination between the loci. Because LD is affected by many factors such as genetic drift, selection within populations, or population admixtures, its genomic structure in a particular crop plants must be determined before association tests can be applied. While some research has already been conducted in maize (Remington et al., 2001; Tenaillon et al., 2001), patterns of LD have still to be analyzed in wheat.

These new tools in the exploitation of genetic diversity hold great promise for further genetic improvement of crops. Together with innovative breeding strategies like the development of SHWs in wheat, it might then be possible to achieve a steady increase in yield potential. Main limitations for the application of these new methodologies are the high costs required for the technical equipment and well-trained staff. Consequently, for breeding institutions in developing countries it seems mandatory to establish central service labs in international centers like CIMMYT for carrying out these sophisticated analyses.

References

Ahmad, M. 2002. Assessment of genomic diversity among wheat genotypes as determined by simple sequence repeats. Genome 45:646-651.

Al-Maskri, A., M. Nagieb, K. Hammer, A.A. Filatenko, I. Khan, and A. Buerkert. 2003. A note about Triticum in Oman. Genet. Resour. Crop Evol. 50:83-87.

Agnus, W.J. 1997. Hybrid wheat – a dream or reality? Aspects Appl. Biol. 50:15-22.

Buckler, E.S. and J. M. Thornsberry. 2002. Plant molecular diversity and applications to genomics. Curr. Op. Plant Biol. 5:107-111.

Dalrymple, D.G. 1986. Development and spread of high-yielding wheat varieties in developing countries, 7th ed. US Agency for Int. Develop., WA.

DeLacy, I.H., B.Skovmand, and J.Huerta. 2000. Characterization of Mexican wheat landraces using agronomically useful attributes. Genet. Resour. Crop Evol. 47:591-602.

Del Blanco, I.A., S. Rajaram, W.E. Kronstad, and M.P. Reynolds. 2000. Physiological performance of synthetic hexaploid wheat-derived populations. Crop Sci. 40:1257-1263.

Del Blanco, I.A., S. Rajaram, and W.E. Kronstad. 2001. Agronomic potential of synthetic hexaploid wheat-derived populations. Crop Sci. 41:670-676.

Donini, P., P. Stephenson, G.J. Bryan, and R.M.D. Koebner. 2000. The potential of microsatellites for high throughput genetic diversity assessment in wheat and barley. Genet. Resour. Crop Evol. 45:415-421.

Dubin, H.J., R.A. Fischer, A. Mujeeb-Kazi, R.J. Pena, K.D. Sayre, B. Skovmand, and B. Valkoun. 1997. Wheat. pp 309-320. *In* Fucillio, D., L. Sears, and P. Stapleton (eds.) Biodiversity in Trust: Conservation and Use of Plant Genetic Resources in CGIAR Centers, Cambridge University Press, Cambridge.

Dvorak, J., M.C. Luo, Z.L. Yang, and H.B. Zhang. 1998. The structure of the *Aegilops tauschii* genpool and the evolution of hexaploid wheat. Theor. Appl. Genet. 97:657-670.

Edwards, I.B., and G. Dorlencourt. 1994. Hybrid wheat - current status and future role in European wheat production. Melhoramento 33:87-104.

Evenson, R.E., and D. Gollin. 2003. Assessing the impact of the Green Revolution, 1960 to 2000. Science 300:758–762.

FAO, 1996. FAO State of the World's Plant Genetic Resources for Food and Agriculture. Food and Agriculture Organization of the United Nations, Rome, Italy.

Galili, S., Y. Avivi, E. Millet, and M. Feldman. 2000. RFLP-based analysis of there *RbcS* subfamilies in diploid and polyploidy species of wheat. Mol. Gen. Genet. 263:674–680.

Gorham, J. 1990. Salt tolerance in the Triticeae: K/NA discrimination in synthetic derived hexaploid wheats. J. Exp. Bot. 41:623-627.

Gut, I.G. 2001. Automation in genotyping single nucleotide polymorphisms. Hum. Mutat. 17: 475-492.

Hede, A.R., B. Skovmand, M.P. Reynolds, J. Crossa,, A.L. Vilhelmsen, and O. Stolen. 1999. Evaluating genetic diversity for heat tolerance traits in Mexican wheat landraces. Genet. Resour. Crop Evol. 46:37-45.

Ho, J.C., S.R. McCouch, and M.E. Smith. 2002. Improvement of hybrid yield by advanced backcross QTL analysis in elite maize. Theor. Appl. Genet. 105:440-448.

Hoisington, D., M. Kahairallah, T. Reeves, J.-M. Ribaut, B. Skovmand, S. Taba, and M. Warburton. 1999. Plant genetic resources: What can they contribute toward increased crop productivity. Proc. Natl. Acad. Sci. 96:5937-5943.

Huang, X.Q., H. Coster, M.W. Ganal, and M.S. Röder. 2003. Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (Triticum aestivum L.). Theor. Appl. Genet. 106:1379-1389.

Huang, X.Q., A. Borner, M.S. Röder, and M.W. Ganal. 2002. Assessing genetic diversity of wheat (Triticum aestivum L.) germplasm using microsatellite markers. Theor. Appl. Genet. 105:699-707.

Lantican, M.A., P.L. Pingali, and S. Rajaram. 2002. *In* Ekboir, J. (ed.) World wheat overview and outlook 2000-2001: Developing no-till packages for small scale farmers. CIMMYT Mexico, D. F, Mexico.

Limin, A.E., and D.B. Fowler. 1993. Inheritance and cold hardiness in *Triticum aestivum ×* synthetic hexaploid wheat crosses. Plant Breed. 110:103-108.

Ma, H., R.P Singh, and A. Mujeeb-Kazi. 1995. Resistance to stripe rust in *Triticum turgidum*, *T. tauschii* and their synthetic hexaploids. Euphytica 82:117-124.

Manifesto, M.M., A.R. Schlatter, H.E. Hopp, E.Y. Suárez, and J. Dubcovsky. 2001. Quantitative evaluation of genetic diversity in wheat germplasm using molecular markers. Crop Sci. 41:682-690.

Pickett, A.A. and N.W. Galwey. 1997. A further evaluation of hybrid wheat. Plant Varieties and Seeds 10:13-32.

Pillen, K., a. Zacharias, and J. Leon. 2003. Advanced backcross QTS analysis in barley (Hordeum vulgare L.). Theor. Appl. Genet. 107:340-352.

Rafalski, A. 2002. Applications of single nucleotide polymorphisms in crop genetics. Curr. Op. Plant Biol. 5:94-100.

Rajaram, S. 1995. Wheat germplasm improvement: Historical perspectives, philosophy, objectives, and missions. pp 1-10. *In* Rajaram, S., and G.P. Hettel (eds.) Wheat Breeding at CIMMYT: Commemorating 50 years of research in Mexico for global wheat improvement. Wheat special report No. 29., CIMMYT, Mexico D.F., Mexico.

Rajaram, S., and M. van Ginkel. 1996. Wheat breeding methodology: International perspectives. 1996. *In:* Tanner, D.G., T.S Payne, and O.S. Abdalla, (eds.). CIMMYT regional wheat workshop for Eastern, Central and Southern Africa, 9, 2-6 Oct 1995, Addis Ababa, Ethiopia.

Reeves, T., S. Rajaram, M. van Ginkel, R. Trethowan, H. Braun, and K. Cassaday. 1999. New wheats for a secure, sustainable future. CIMMYT, Mexico, D.F., Mexico.

Remington, D.L. J.M. Thornsberry, Y. Matsuoka, M.L. Wilson, S.R. Whitt, J. Doebley, S. Kresovich, M.M. Goodman, and E.S. Buckler. 2001 Strukture of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. 98:11479-11484.

Reif, J., P. Zhang, S. Dreisigacker, M.L. Warburton, M. van Ginkel, D. Hoisington, M. Bohn, and A.E. Melchinger. 2004. Wheat genetic diversity trends during domestication and breeding. Proc. Natl. Acad. Sci., submitted.

Röder, M.S., K. Wendehake, V. Korzun, G. Bredermeijer, D. Laborie, L. Bertrand, P. Isaac, S. Rendell, J. Jackson, and R.J. Cooke. 2002. Construction and analysis of a

microsatellite-based database of European wheat varieties. Theor. Appl. Genet. 106:67-73.

Roy, K., M.S. Lakshmikumaran H.S. Balyan, and P.K. Gupta. 2004. AFLP-based genetic diversity and its comparison with diversity based on SSR, SAMPL, and phenotypic traits in bread wheat. Biochem. Genet. 42:43-59.

Skovmand, D., M. Warburton, S. Sulliwan and J. Lage. 2003. Managing and collecting genetic resources. *In:* Nachit, M., N. Di Fonzo, J.L. Araus, W. Pfeiffer, G.A. Slafer, and R. Conxita (eds.). Durum wheat breeding: current approaches and future strategies. In press.

Tanksley S. D. and S. R. McCouch. 1997. Seed banks and molecular maps: Unlocking genetic potential from the wild. Science 277:1063-1066.

Tanksley, S.D. and Nelson, J.C. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet. 92:191-203.

Tenaillon, M.I., Sawskins M.C., A.D. Long, R.I. Gaut, J.F. Doebley, and B.S. Gaut. 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (Zea mays ssp. mays L.) Proc. Natl. Acad. Sci. 98:9161-9166.

Valkoun, J.J. 2001. Wheat pre-breeding using wild progenitors. Euphytica 119:17-23.

Van Ginkel, M., R. M. Trethowan, K. Ammar, J. Wang, M. Lillemo. 2002. A guide to bread wheat breeding at CIMMYT. Wheat special report No. 5 (Rev. edition.). CIMMYT, Mexico, D. F., Mexico.

Villareal, R.L., and A. Mujeeb-Kazi. 1996. Exploitation of synthetic hexaploids (Triticum turgidum x T. Tauschii) for some abiotic resistances in wheat. *In:* Richards, R.A., C.W. Wrigley, H.M. Rawson, G.J. Rebetzke, J.L. Davidson, R.I.S. Brettell (eds.) Proceedings of the 8th wheat breeding society of Australia, Canberra, Australia, 29 Sep - 4 Oct 1996. Wheat Breeding Society of Australia, Tamworth, Australia.

Vigouroux, Y., J.S. Jaqueth, Y. Matsuoka, O.S. Smith, W.D. Beavis, J.S.C. Smith, and J. Doebley. 2002. Rate and pattern of mutation at microsatellite loci in maize. Mol. Biol. Evol. 19:1251-1260.

Zhang P., S. Dreisigacker, S. Alkhanjari, A. Buerkert, A. E. Melchinger, M. L. Warburton. 2004. Genetic diversity and relationships of wheat landrace cultivars from Oman investigated with SSR markers. Genet. Resour. Crop Evol., submitted.

7 SUMMARY

Wheat (*Triticum aestivum*) is one of the major cereals in the world. During the past years, the world consumption of wheat increased up to nearly 600 million tones, whereas wheat production continuously decreased. Due to land limitations, new production gains must be achieved from improved plant management systems as well as from the development of high yielding varieties. The International Maize and Wheat Improvement Center (CIMMYT) employs different strategies to enhance yield potential in wheat especially for developing countries. For instance, the wheat breeding program focuses on defined mega-environments (MEs), assuming similar growing conditions in certain countries. In the search for useful alleles, breeders often turn back to wild relatives of wheat stored in the CIMMYT seed bank. With the production of synthetic hexaploid bread wheat (SHWs), characteristics from *T. durum* and *T. tauschii* can be combined and via backcrossing incorporated into modern breeding materials. Wheat landraces (LCs) are an additional reservoir of resistances to pests and diseases as well as for environmental adaptation. The production of wheat hybrids is seen as a further option to improve yield potential. A considerable amount of genetic diversity among the materials is a prerequisite for all strategies. Due to the worldwide importance of CIMMYT wheat varieties, they represent a suitable source to examine different breeding strategies in wheat.

The main objective of our research was to determine the genetic diversity in modern wheat breeding materials and genetic resources at CIMMYT. Specific research questions were: (i) Is the systematic breeding targeted for different MEs reflected in the genetic diversity among breeding lines (Experiment 1)? (ii) Does the production of SHWs (Experiment 2) and the use of LCs (Experiment 3) enhance the genetic variation in modern breeding materials? (iii) Does the development of hybrids represent an option to improve yield potential in wheat (Experiment 4)? (iv) Is it possible to predict levels of heterosis with the determination of genetic distance (GD) among hybrid parents (Experiment 4)? (v) Do genomic and EST- derived SSRs differ in the measurement of genetic diversity (Experiments 1 to 3)? (vi) Are GD values based on SSRs correlated with the coefficient of parentage (COP) (Experiments 1 and 4)?

In Experiment 1, a total of 68 CIMMYT advanced breeding lines was analyzed with 99 SSRs, of which 51 were EST- and 46 genomic derived SSRs. A high level of genetic

diversity ($\overline{GD}$ = 0.41) was observed among the breeding lines. The majority of variation (91%) was detected among lines targeted to one specific ME, which indicates a broad genetic base of the current CIMMYT breeding materials. Principal coordinate analysis (PCoA) could clearly separate the lines, but they clustered independently from their target MEs. Main explanations are: (i) alleles were selected that provide fitness to several MEs, (ii) adaptation depends only on a small number of genes that were not detected with the SSRs applied, or (iii) too few cycles of selection were considered to separate the germplasm.

In Experiment 2, a total of 11 SHWs, 7 recurrent parent lines, and 13 families of backcross-derived lines (SBLs) were analyzed with 90 SSRs. The SHWs clustered far from the SBLs and the recurrent parents in the cluster analyses and PCoA, and formed a distinct germplasm pool with high allelic variation. Two families of SBLs were tested for a selective advantage of the SHW alleles. Six SSRs revealed non-Mendelian inheritance, indicating that the genomic region of SHWs was actively selected for. Thus, the production of SHWs provides a promising approach for the enhancement of genetic variation in modern breeding materials.

In Experiment 3, seed bank accessions of 36 LCs from different countries and a total of 119 accessions from nine LCs populations collected in Turkey and Mexico were analysed with 44 and 76 SSRs, respectively. Both LC materials revealed high allelic variation ($\overline{GD}$ = 0.69 and 0.54). The 36 LC accessions could not be separated according to their continent of origin. Furthermore, an unexpected relationship was observed between the Chilean LC "Trigo africano" and the Nigerian LCs "Dikwa". All of the nine LC populations could be discriminated except for two Turkish LCs collected from the same location. In accordance with previous studies, considerable genetic variation was observed within the LC populations. Our results contributed substantially to the characterization of the LCs and generated important knowledge for the management of seed bank accessions.

In Experiment 4, a total of 112 wheat hybrids and their 22 parental lines were evaluated at two locations in Mexico for grain yield, plant height, days to flowering and maturity. The level of heterosis varied between -15.3% and 14.1%, but was generally too low to compensate the high costs of hybrid seed production. The correlations between mid-parent values and hybrid performance, as well as between parental line *per se* performance and general combining ability were significant ($P < 0.01$) for all traits, and particularly

high for grain yield (r = 0.86 and 0.91). PCoA based on 113 SSR markers revealed three groups of parents. However, the correlations of GDs and COPs with the heterosis values were negative and not significant. Thus, the prospects of large-scale cultivation of hybrid wheat in developing countries are low.

The correlations between GDs and COP in Experiments 1 and 3 were generally significant but low. This can be explained by unrealistic assumptions in the calculation of COPs, which ignore the effects of selection and genetic drift. Similarly to genomic SSRs, EST-SSRs did not reflect functional diversity. The latter revealed a lower degree of polymorphism than genomic SSRs in all experiments, but the allele designation was simpler and more reliable.

Across all experiments, our study demonstrates that plant breeding does not inevitably lead to a loss of genetic diversity. We confirmed that CIMMYT's breeding strategies contributed to a successful increase in genetic variation. These results provide useful information to wheat breeders in CIMMYT and other national programs, regarding the use of wild relatives and landraces for the enhancement of the genetic base of wheat germplasm. In addition, our research provides a base of knowledge for future association studies, identification of useful alleles, and their use in marker-assisted selection.

8 ZUSAMMENFASSUNG

Weizen (*Triticum aestivum*) ist eine der wichtigsten Getreidearten der Welt. Während in den letzten Jahren der Weizenverbrauch weltweit auf knapp 600 Mio. t anstieg, ist die Produktion rückläufig. Ein erneuter Produktionszuwachs kann angesichts weltweit limitierter Anbauflächen nur über verbesserte Anbauverfahren und die Züchtung verbesserter Weizensorten erfolgen. Mit verschiedenen Strategien versucht das Internationale Institut für Mais und Weizenzüchtung (CIMMYT) in Mexiko das Ertragspotential des Weizens speziell in Entwicklungsländern zu erhöhen. Aufgrund der in vielen dieser Länder vergleichbaren Anbaubedingungen züchtet das Weizenprogramm u.a. für definierte Makro-Umwelten (MEs). Auf der Suche nach wertvollen Allelen greifen die CIMMYT-Züchter häufig auf Wildtypen in der Genbank zurück. Mit der Erzeugung von synthetischen Weizen (SHWs) werden Eigenschaften von *T. durum* und *T. tauschii* kombiniert und anschließend mittels Rückkreuzungen in modernes Zuchtmaterial von *T. aestivum* eingebracht. Landrassen (LCs) stellen ein zusätzliches Reservoir für Resistenzen und verbesserte Umweltanpassung dar. Die Züchtung von Hybridweizen wird als weitere Option zur Erhöhung des Ertragspotentials angesehen. Die genetische Diversität im Ausgangsmaterial ist gleichermaßen wichtig für alle diese Züchtungsstrategien. Aufgrund seiner weltweiten Bedeutung stellt das Weizenmaterial des CIMMYT eine ausgezeichnete Quelle zur Untersuchung der verschiedenen Optionen für die Weizenzüchtung dar.

Hauptzielsetzung der vorliegenden Studie war die Erforschung der genetischen Diversität in aktuellem Zuchtmaterial und genetischen Ressourcen des CIMMYT. Im Einzelnen wurden folgende Fragestellungen bearbeitet: (1) Spiegelt sich die systematische Züchtung für verschiedene MEs in der genetischen Diversität zwischen Weizenlinien wieder (Experiment 1)? (2) Kann durch die Erzeugung von SHWs (Experiment 2) bzw. die Nutzung von LCs (Experiment 3) eine Erweiterung der genetischen Variation in modernem Zuchtmaterial erreicht werden? (3) Welche Möglichkeiten zur Anhebung des Ertragspotentials bietet die Erzeugung von Hybriden (Experiment 4)? (4) Inwieweit kann mit der Bestimmung genetischer Distanzen (GD) zwischen Elternlinien die Heterosis in Weizenhybriden vorausgesagt werden (Experiment 4)? (5) Zeigen genomische und EST-abgeleitete SSRs Unterschiede in der Bestimmung genetischer Diversität (Experimente 1 bis 4)? (6) Ist die genetischer Diversität, berechnet anhand von SSRs, mit dem Abstammungskoeffizienten (COP) korreliert (Experimente 1 und 3)?

In Experiment 1 wurden 68 Hochleistungslinien mit 99 SSR Markern untersucht, von denen 51 EST- und 46 genomische SSRs waren. Es wurde ein hohes Ausmaß an genetischer Diversität ermittelt ($\overline{GD}$ = 0,41). Der Hauptteil der genetischen Variation (91%) wurde innerhalb der für MEs gezüchteten Linien gefunden, was auf eine außerordentliche Breite des CIMMYT-Zuchtmaterials hinweist. In der Hauptkoordinatenanalyse (PCoA) gruppierten die Linien unabhängig von ihrer Anpassung an fünf MEs. Mögliche Ursachen dafür sind: (a) Es wurden Allele selektiert, welche hohe Fitness in mehreren MEs bewirken; (b) die für die Anpassung verantwortlichen Allele wurden nicht mit den untersuchten SSRs erfasst; (c) die Selektion auf Anpassung an die MEs erstreckte sich über zu wenige Selektionszyklen.

In Experiment 2 wurden 11 SHWs, 7 rekurrente Elternlinien (BWs) und 13 davon abgeleitete Familien von Rückkreuzungslinien (SBLs) mit 90 SSRs untersucht. Anhand einer Clusteranalyse und PCoA konnten die SHWs eindeutig von den SBLs sowie BWs getrennt werden. Sie stellen somit einen eigenständigen Pool an Genmaterial mit hoher Allelvariation dar. Zwei SBL-Familien wurden auf einen selektiven Vorteil der SHW-Allele getestet. Sechs SSRs zeigten dabei eine signifikante Abweichung von den unter Mendelscher Vererbung erwarteten Genfrequenzen, was auf eine gerichtete Selektion der SHW-Allele schließen lässt. Die Erzeugung von SHWs bietet somit ein Potential zur Erweiterung der genetischen Variation in modernem Zuchtmaterial.

In Experiment 3 wurden Akzessionen von 36 LC aus verschiedenen Ländern sowie 119 Akzessionen von neun LC-Populationen aus der Türkei und Mexiko mit 76 bzw. 44 SSRs analysiert. Beide Materialgruppen zeigten ein hohes Maß an genetischer Diverstät ($\overline{GD}$ = 0,69 bzw. 0,54). Die 36 LCs konnten nicht entsprechend ihren Ursprungs-kontinenten getrennt werden. Zudem wurde eine bisher unbekannte Beziehung zwischen einer chilenischen und nigerianischen LC aufgedeckt. Mit Ausnahme zweier türkischer LC-Populationen konnten alle neun untersuchten LCs differenziert werden. Der Großteil der genetischen Variation wurde jedoch innerhalb der LC-Populationen ermittelt. Die Ergebnisse dieses Experiments leisteten somit einen wesentlichen Beitrag zur Charakterisierung der LCs und führten zu wichtigen Erkenntnissen für das künftige Management von Genbank–Akzessionen.

In Experiment 4 wurden 112 Weizenhybriden sowie deren Elternlinien in zwei mexikanischen Umwelten für Ertrag, Wuchshöhe, Blütezeit und Reife evaluiert. Die

Heterosiswerte für Ertrag schwankten zwischen -15,3% und 14,1% und waren im Durchschnitt zu niedrig, um die höheren Produktionskosten für das Hybridsaatgut zu amortisieren. Die Korrelationen zwischen Elternmittel und Hybridleistung sowie Eigenleistung und Testkreuzungsleistung der Hybrideltern waren signifikant und für alle Merkmale inklusive Ertrag sehr hoch (r = 0,86 bzw. 0,91). Mit Hilfe von 113 SSRs konnten die Elternlinien in drei Gruppen eingeteilt werden. Die Korrelation zwischen GD und Heterosis war jedoch niedrig und nicht signifikant. Nach diesem Ergebnis sind die Aussichten für den potentiellen Erfolg von Hybridweizen in Entwicklungsländern als gering einzustufen.

In Experiment 1 und 4 waren die Korrelationen zwischen den GD und COP-Werten signifikant aber niedrig. Dies beruht vermutlich auf unrealistischen Annahmen bei der Berechnung des COP, da hierbei Selektion und genetische Drift ignoriert werden. Im Vergleich zu genomischen SSRs spiegelten auch EST-SSRs keine funktionelle Diversität wieder. Die EST-SSRs waren in allen Experimenten weniger polymorph, jedoch einfacher auszuwerten als die genomischen SSRs.

Über alle Experimente hinweg zeigt unsere Studie, dass Pflanzenzüchtung nicht notwendigerweise zum Verlust genetischer Variation führen muss. Sie belegt, dass die am CIMMYT verfolgten Strategien erfolgreich zu einer Verbreiterung der genetischen Variation beitragen. Die Ergebnisse geben den Weizenzüchtern am CIMMYT und in nationalen Zuchtprogrammen wertvolle Hinweise zur Nutzung von Wildarten bzw. Landsorten für die Erweiterung der genetischen Basis bei Weizen. Außerdem bieten sie eine Basis für künftige Assoziationskartierungsstudien zur Identifizierung von wertvollen Allelen und deren Nutzung im Rahmen der marker-gestützten Selektion.

ACKNOWLEDGEMENTS

First and foremost I wish to thank Prof. Dr. A.E. Melchinger who offered me the unique opportunity to conduct the work for my research at CIMMYT, Mexico. His invaluable advice, prompt and insightful suggestions, guided and encouraged me especially in scientific writing during the last year back in Hohenheim.

I'm very greatful to Drs. Marilyn Warburton and Dave Hoisington in order to obtain experience in molecular marker analyses. The skilled technical assistance of Emiliano Villordo, Salvador Ambriz and Leticia Diaz helped me to realize all the lab work.

 I'm also deeply indebted to Drs. Maarten van Ginkel, Bent Skovmand, and Karim Amar for providing me the plant materials and detailed pedigree information of the wheat inbreds as well as for their great inspirations interpreting my results. I'd like to thank Jose Luis and Mario for initiating the field experiments and providing the seed bank accessions.

In Hohenheim, I really enjoyed being part of the 350a working group who created me a pleaseant working environment. In particular I'd like to thank Hans Peter, Jochen, Martin, Matthias, Christian and Zhang for always having a listening ear, some chocolate to consulate, and a helping hand when things got rough.

I'd also like to thank Dr. Martin Bohn for his initial scientific support.

My appreciation is extended to Ms Boesig and Ms Dölle for their friendliness and helpfulness in everyday life at the institute and also Prof. Utz for his statistical and moral support.

Special regards go to my all my friends. In Germany, as well as in Mexico they have consistently helped me keep perspective on what is important in life and shown me how to deal with reality. To name a few: Nadja, Wenceslao, Kirsten, Peter, Rogelio, Sybil, Olaf, Anja, Fernando, Simone, Alex, and Christine.

Fianally I'd like to thank my family, especially my brother and sister, Eva and Meinhard, and my parents, for their unconditional support and understanding and for always being there when needed.

CURICULUM VITAE

Personal Details

Name	Susanne Dreisigacker
Birth	16.01.1974 in Landau

School Education

1980-1984	Elementary school in Siebeldingen
1984-1993	High school, Maria-Ward Gymnasium in Landau
	High school degree „Abitur"

University Education

Oct. 1994 - January 2000	Agricultural Biology at the University of Hohenheim, Main subjects: Plant Breeding, Population Genetics, Plant Physiology and Plant Cultivation
	Diploma in Agricultural Biology
May 2000 – to date	Doctorate candidate in Plant Breeding and Applied Genetics, University of Hohenheim

Practical Experience

Sept. 1993 – Aug.1994	Trainee in a tree nursery in Berg, at the Research Institute of Vine Breeding in Siebeldingen, and at the Research Institute of Agriculture and Vine Cultivation in Neustadt
Feb. 1998 – Sept. 1998	Trainee in the Durum wheat breeding program at the International Maize and Wheat Improvement Center (CIMMYT) in Mexico

Professional Employment

Jan. 1997 – March 2000	Research Assistant working in the improvement of double haploids in maize as well as in the evaluation of maize populations for resistance to sugarcane mosaic virus (SCMV) at the Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim
Jan. 2004 – to date	Senior Research Associate at Pioneer Hi-Bred Northern Europe in Greven